

How inadequate data governance frameworks lead to unethical outcomes in Artificial Intelligence Systems

Neelam Gupta *

Director, Data and Artificial Intelligence, Avanade, United States.

Abstract

The increasing adoption of artificial intelligence (AI) technology in decision-making has made incredible advances, but it also has significant ethical problems. A crucial, yet often ignored, factor is insufficient data governance practices. This article examines how inadequate data governance practices, such as a lack of accountability, weak privacy protections, a lack of quality control in data management, and weak traceability, contribute to unethical outcomes with AI. Using relevant case studies and promising practices for consideration, we conclude that data governance is at the core of the ethical use of AI. The paper ends with public policy recommendations and organizational approaches to attenuate risks and enhance fairness, transparency, and accountability in AI.

Keywords: Data Governance; Artificial Intelligence Ethics; Data Quality; Algorithmic Accountability; Data Privacy; Responsible AI; Data Stewardship; Transparency in AI; AI Risk Management

1. Introduction

Artificial intelligence (AI) systems are increasingly being used to automate complex decisions in many fields from health care to finance to criminal justice and education. AI systems rely on large amounts of data to perform as intended. The ethical potential of AI is contingent not just on the algorithms but also whether there is strong oversight of the data itself.

Data governance refers to the people, processes, policies, standards, and technologies that enable the appropriate ethical and effective management of data throughout its lifecycle. Data governance is a crucial part of ensuring, specifically in the context of AI systems, that they are transparent, accountable, ethical and fair. We recognize that, ethical AI systems are possible, sustainable and viable even though many organizations are using AI without appropriate governance frameworks that result in opaque, biased and sometimes harmful decision outcomes.

There has been significant discussion regarding the intersection of AI and ethics, however the relationship between the absence of strong data governance frameworks and unethical AI decision-making is under-researched. This could be framed as an identified gap in the research. We discuss the causal mechanisms connecting weak data governance to ethical breaches including bias, privacy violations, discrimination and other ethical issues. Drawing on case examples and our policy recommendations, we provide recommendations for the adoption of ethical governance within the development and deployment of AI.

* Corresponding author: Neelam Gupta

2. Understanding Data Governance in the AI Context

2.1. Definition and Importance

Data governance refers to the separate governing framework around the availability, integrity, usability, and security of data for a given organization. This governing framework is especially paramount in the domain of artificial intelligence (AI) systems, due to the data-centric nature of machine learning algorithms, as the data ingested by AI model is the basis for decisions that impact individuals and society. If the data used to train an AI model is flawed, poorly managed, or ethically obtained, the resultant AI system will inevitably generate unethical, biased, or harmful outcomes.

The value of data governance in AI is that it also provides ethical compliance, public trust, and regulatory standards. Organizations can no longer operate AI under a singular performance-based premise, but must now understand that AI needs to be fair, transparent, and accountable, all of which begins with data governance.

2.2. Key Components of Data Governance in AI

Robust data governance is based on a series of related elements that help control responsible data use throughout its lifecycle. These components include:

- **Data Quality Management:** is used to ensure that the data being processed and used in AI systems is accurate, complete, consistent, and up-to-date.
- **Metadata Management:** describes additional information concerning where data came from, what it was intended to represent, its format, etc.
- **Access Control and Data Security:** describes who may access, edit, or share sensitive or personal data.
- **Data Lineage and Traceability:** describes the journey of data from source through its processing, transformation, and eventual use in AI models.
- **Data Stewardship and Assignment of Accountability:** outlines roles and responsibilities for managing data, and then holds those responsible for the enforcement of the agreed-upon ethical standards.
- **Policy Compliance and Legal Alignment:** describes if, and to what extent, the data governance aligns with existing laws (e.g., GDPR, CCPA) and ethical codes of practice.

These elements are all critical to managing the risks involved with automated decision-making in data-driven AI systems.

Table 1 Key Data Governance Components and Associated Ethical Risks in AI

Data Governance Component	Purpose in AI Systems	Ethical Risk if Inadequate
Data Quality Management	Ensures accurate, complete, and reliable data	Biased or misleading model predictions due to inaccurate or skewed data
Metadata Management	Provides context, origin, and description of data	Lack of transparency; inability to assess data relevance or fairness
Access Control & Data Security	Restricts unauthorized access to sensitive data	Privacy breaches; misuse of personal data
Data Lineage & Traceability	Tracks data journey from source to output	Accountability gaps; difficult to trace errors or unethical outcomes
Data Stewardship & Accountability	Assigns responsibility for data integrity and compliance	No clear responsibility for bias or data misuse
Policy & Legal Compliance	Ensures data use complies with laws like GDPR, CCPA	Legal violations; infringement on individual rights

2.3. AI-specific challenges to governance

AI creates different governance challenges from standard data governance practices.

Dynamic data environments: There are many AI systems that continually learn and adapt using real-time data. In these contexts of continual change, governance processes and mechanisms to monitor these changes and deviations are essential to prevent mission drift towards unethical purpose and behavior.

Opacity and complexity of models: The 'black box' nature of AI, especially deep neural networks, makes it challenging to unpack decision-making. Models can be so opaque, often poorly documented with limited data lineage, that unpacking the source of flawed or questionable outcomes and correcting them can range from difficult to impossible.

Bias amplification: AI systems develop social bias through the datasets that they ingest. If datasets are ingested without audit for representative, fair or equitable purposes, in spite of having equally well-intentioned outcomes, AI systems can reinforce harms across hiring, credit scoring, policing, etc.

Lack of ethical oversight: Data teams and AI developers often work in silos. Siloed culture and lack of cross discipline awareness of ethics can lead overrated AI behavior, behavior that may exhibit harms, which are not exposed until an AI system is up and fully running.

2.4. The role of data governance in AI governance

AI governance is the greater governance framework that data governance actually forms a part of. AI governance encompasses risk management, ethical oversight, and regulatory compliance frameworks to govern AI systems. AI governance is broader and is not fully static, but it includes data governance as a subset.

Without good and reliable data governance, a larger vision of ethical AI will always remain a hypothetical ideal, including frameworks for transparency, fairness, and accountability. Data is the raw material of AI and if it is not good data, the opportunity for AI to have good outcomes is diminished.

3. Ethical Dimensions in AI Systems

3.1. Overview of AI Ethics

AI ethics is the use of moral precepts and social values that inform how artificial intelligence systems are built, used, and governed. AI is making our decisions in part, if not in whole, for several types of systems, including healthcare, hiring practices, criminal justice, and finance, thus making the design and implementation of AI ethical processes of paramount importance.

Some key ethical questions are:

- Is the AI fair and non-discriminatory?
- Does the AI respect the privacy or autonomy of users?
- Did the AI make a decision that is transparent and explainable?
- Who takes the blame for mistakes or harms?

AI ethics stems from many areas philosophy, law, data science, and sociology, and is meant to ensure that artificial intelligence technologies are fair, transparent, accountable, and safe.

Recent examples of unethical AI-based hiring platform algorithms or non-transparent credit scoring make these ethical concerns very real. AI ethics is not just theoretical; it is essential to developing systems that are trustworthy, lawful, and socially responsible.

3.2. Core Ethical Principles in AI

The core notion of ethical AI development is built on a set of principles to ensure technology serves humanity in an equitable, transparent, and responsible manner. The principles are the moral guide for policymakers, developers, and organizations consuming AI, across all industries. We will delve further into these ideas below:

3.2.1. Fairness

Fairness refers to the duty of non-biased, non-discriminatory, and non-favoritism within the design, development, and deployment of algorithms. The expectation of fairness infers reasonable equitable treatment regardless of race, gender, age, socioeconomic status, among many other traits.

3.2.2. Key Challenges

Bias in Historical Training Data: Historical training data captures certain societal inequalities. When that occurs and the historical data, and by proxy, the unfair behaviors or action are unconsciously incorporated in by a system operator or a data scientist, the AI systems may factor in existing discrimination or inequities.

Homogeneous Information System Developer Teams: Developers who belong to social identity bases which are stereotypically overrepresented may not see the fairness problem that adversely impacts underrepresented social identity bases.

Example

A hiring tool that uses AI would train on male-dominated resumes and rank females lower than males, therefore reaffirming gender biases in the hiring system.

3.2.3. Mitigation Strategies

- Regular Audits for fair and bias outcomes
- Embedding in a wrapped environment an Inclusive and equitable training dataset
- Algorithmic fairness requires assistive metrics implemented into AI systems (equal opportunity, demographic parity).

3.2.4. Ethical Dimensions of AI

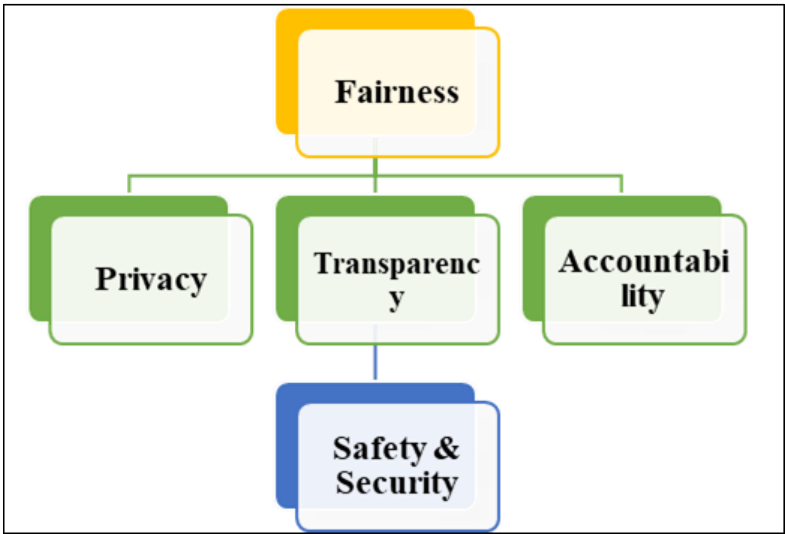


Figure 1 Ethical Dimensions of AI

This SmartArt graphic visually separates the five principles of ethical principles, emphasizing how interconnected they are. Fairness is depicted at the top in recognition of being the foundational principle responsible for establishing equitable treatment across all other principles.

3.3. Ethical Risks from Inadequate Data Governance

Poor data governance is one of the leading causes of ethical failures in AI systems. Below is a table illustrating how weak governance in specific areas leads to violations of core ethical principles.

Table 2 Ethical Consequences of Poor Data Governance in AI Systems

Weak Governance Area	Ethical Principle Affected	Resulting Ethical Violation
Inaccurate or biased datasets	Fairness	Discrimination against marginalized groups
Unsecured personal data	Privacy	Unauthorized data sharing or breaches
Lack of documentation or lineage	Transparency	Inability to explain AI decisions
Absence of clear roles/responsibility	Accountability	No entity held liable for algorithmic harm
Unmonitored algorithm behavior	Safety and Security	Systems prone to adversarial attacks or malfunction

3.4. Real-World Examples

- **COMPAS Algorithm (U.S.):** a risk assessment tool applied to criminal justice that was shown to be racially biased, primarily due to unbalanced and poorly organized training data.
- **Amazon's Hiring AI (2018):** which was trained on resumes submitted over a decade, made predominantly by male applicants, resulted in gender bias against female applicants.
- **Facebook:** Cambridge Analytica Scandal– a large case of privacy violation due to poor control over how our data was collected and subsequently shared.

These instances illustrate the consequences of poor data governance, lending themselves directly to ethical violations and public distrust.

4. Case Studies: Consequences of Inadequate Data Governance in AI

AI systems often take on the properties related to the nature of the data and the design processes of the data. When governance frameworks are weak, lacking in transparency, accountability, quality control, or fairness, the result can be ethically disastrous. In this section, we share real examples, case studies, where weak data governance resulted in demonstrable harm, litigation, or social backlash.

4.1. Case Study 1: Amazon’s AI Hiring Tool

4.1.1. Background

In 2014, Amazon built an internal AI system for automating the hiring process. The system was trained on resumes that have been submitted over the past ten years.

4.1.2. Issue

The dataset was predominantly male resumes, particularly for technical roles, and the AI system had learned to downgrade resumes that contained the word “women’s,” for example, “women’s chess club captain,” or graduates from all-women's colleges.

4.1.3. Governance Gaps

- No bias auditing of the training data.
- No oversight for ethical considerations or human review.

4.1.4. Ethical Foul-ups

- Fairness: The system discriminated against qualified female candidates.
- Accountability: No attribution for responsibility to ensure model was fair.

4.1.5. Outcome

Amazon ultimately scrapped the tool. But the case illustrated the risks of deploying AI without inclusive datasets and an assessment of bias.

4.2. Case Study 2: COMPAS Recidivism Algorithm

4.2.1. Background

The Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) tool has been utilized in U.S. courts to assess the risk of recidivism and assist in making bail and sentencing decisions for offenders.

4.2.2. Issue

In 2016, a ProPublica investigation discovered that even though Black and white defendants had identical or even less serious criminal records, Black defendants were almost twice as likely to be misclassified as high risk.

4.2.3. Governance Gaps

The algorithm was proprietary and had no transparency in decision-making.

No external auditing or fairness audits.

Claimed to be based on historical criminal justice data, which was racially biased.

4.2.4. Ethical Foul-ups

- Fairness: Further entrenched systemic racial biases.
- Transparency: Neither the courts nor the defendants had a right to see or contest the results.

4.2.5. Impact

Significant public backlash and advocacy for legislation with demands for transparency and fairness in AI criminal justice tools.

4.3. Case Study 3: Google Photos Tagging Incident

4.3.1. Background

In 2015, people noticed that the photo identification algorithm in Google Photos incorrectly tagged photos of Black people with "gorilla."

4.3.2. Issue

This came about because the algorithm trained with non-representative image data-- the training dataset was not racially diverse.

4.3.3. Governance Gaps

- Not enough testing with diverse populations
- No ethical review mechanism while deploying
- Poor dataset curation practices

4.3.4. Ethical Failures

- Privacy and Dignity: Offensive and harmful misclassification.
- Safety and Trust: Erosion of trust by the general public in AI systems for image recognition.

4.3.5. Resolution

Google publicly apologized and also deactivated the gorilla tagging feature, but this incident raised industry-wide questions regarding racial bias in computer vision systems.

Table 3 Summary of Case Studies and Ethical Failures

Case Study	Data Governance Failure	Ethical Breaches	Consequences
Amazon AI Hiring Tool	Biased training data, lack of bias audits	Fairness, Accountability	Discriminated against women, system withdrawn
COMPAS Recidivism Tool	Historical bias, lack of transparency or audits	Fairness, Transparency, Accountability	Racial discrimination, public backlash, calls for legal reform
Google Photos Image Tagging	Poor dataset diversity, insufficient testing	Fairness, Privacy, Safety	Racially offensive misclassification, public apology, loss of user trust

4.4. Chain of Ethical Failure in AI Systems

This visual flow illustrates how poor data governance cascades into ethical failure:

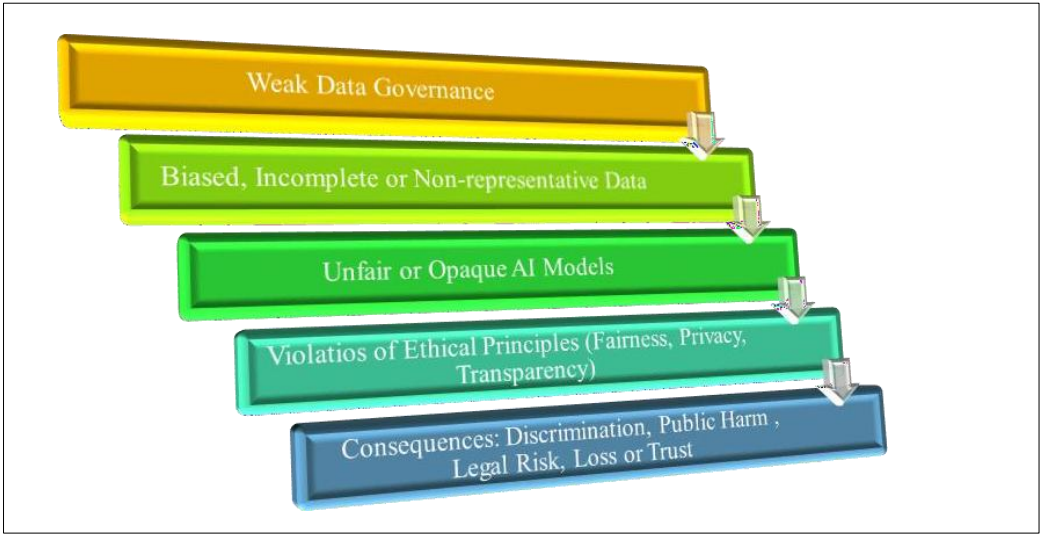


Figure 2 Chain of Ethical Failure in AI Systems

These case studies demonstrate that without effective and strong data governance, ethical AI will not exist. Organizations must pay attention to strong standards at each stage—data collection, model design, deployment, and monitoring—to prevent harm. Without strong due diligence, even well-meaning AI systems will discriminate, harm and erode public trust.

5. Implications for Policy and Practice

The dangers of poor data governance in AI systems can be well beyond being just technical or legal in nature. There are potential risks that are additional because there are societal or ethical implications for decision makers. In this section, we will outline the significant lessons learned from the above case studies and provide some recommendations for implementers, developers, and others who work with AI in the future.

5.1. Regulatory implications

5.1.1. Regulatory or mandated audits and impact assessments

Governments and regulators ought to mandate that organizations conduct regular algorithmic audits, ethical impact assessments, and fairness checks when they are unable to determine if they have a high-stakes decision to make. All stakeholders should be required and can demonstrate oversight over their high-stakes decision-making systems, particularly in healthcare, money, and law enforcement.

Transparency and explainability laws. Policies should require explainability in AI. For authorized AI-driven systems, rationales ought to be provided in a stand-alone format that can be clearly described to humans, particularly when the outcomes affect rights or access to important services.

Privacy regulation. There is now a supplementary risk to the outcomes of AI systems, but we will still have to comply with existing data protection regimes like the GDPR. New AI regulations might also be required to address new and emergent risks related to the new territory of issues and risks posed by machine learning and neural networks.

5.2. Best Practices for an Organization

5.2.1. Establish Ethical Review Boards

Organizations should form multidisciplinary committees to consider the ethical, legal, and social implications of AI projects prior to implementation.

5.2.2. Implement Strong Data Governance Frameworks

Establish clear data stewardship roles, data lineage, and access controls to manage data quality and fidelity, compliance with legislation and regulations, and accountability.

5.2.3. Bias Mitigation Training

AI Development teams should be trained in bias detection, fairness, and principles of responsible AI design.

5.3. Standards and Certification

5.3.1. Standards Bodies

Academia, industry, and governments will need to collaborate to define best practices and set minimum standards for ethical requirements - e.g., IEEE "Ethically Aligned Design," or ISO/IEC standards for AI.

5.3.2. Third-party Certification

Just as products are subjected to quality and safety testing, AI systems should be certified by a third party for ethical readiness, bias mitigation, and data governance maturity.

5.4. Public and Stakeholder Engagement

5.4.1. Inclusivity in design and a mechanism for public engagement

Involve users and affected communities in the AI lifecycle, and incorporate inclusive data and system design regarding cultural, gender, and socio/economic representation.

5.4.2. Transparency portals and disclosures

Organizations should publish transparency reports that record their AI-related decisions, uses, rates of use, bias tests conducted, and provide mechanisms for individuals to seek redress for being harmed due to automated decisions.

Table 4 Key Policy and Practice Recommendations

Category	Recommendation	Objective
Regulatory	Algorithmic impact assessments and audits	Promote transparency, prevent discrimination
Organizational Governance	Ethical AI committees and role-based data stewardship	Enhance accountability and bias control
Industry Standards	Adoption of IEEE/ISO guidelines and ethical certifications	Ensure compliance and stakeholder trust
Public Engagement	Feedback mechanisms and community consultations	Build inclusive and user-centric AI

5.5. Strategic Framework for Ethical AI Governance

To unify the above, organizations can adopt the following **Strategic Governance Cycle**

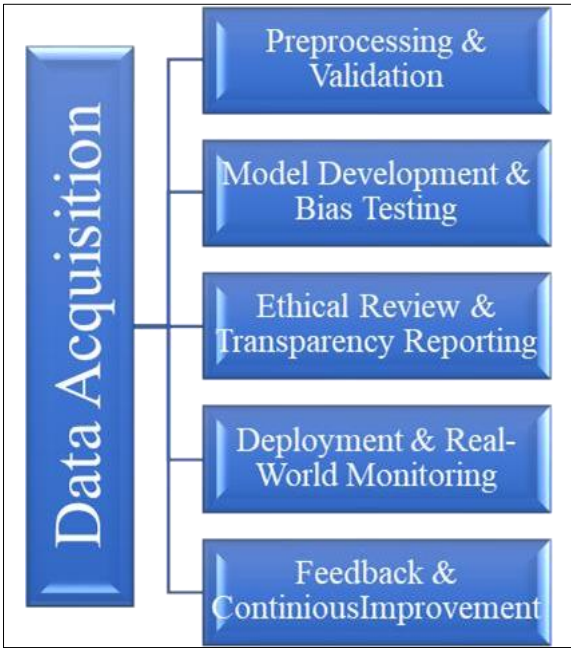


Figure 3 Strategic Framework for Ethical AI Governance

By embedding these governance strategies, policymakers and practitioners can build ethical, transparent, and socially responsible AI systems that protect rights and foster trust across all stakeholders.

6. Recommendations and Future Directions

Creating ethically responsible AI systems requires a broad approach to connect policy, design, and education. Improving data governance, as a starting point, is essential—ensuring data quality, representativeness, and indicators of bias in outcomes can greatly reduce unethical outputs from AI systems. It is also just as important to embed ethical considerations into AI systems rather than determining how to address ethical concerns after the AI systems are designed. This includes not just designing fair and accountable systems, but also including the human component of oversight, especially for critical applications.

From a regulatory perspective, the need for collaboration across borders is essential to develop common ethical standards or certifications that hold AI systems accountable. Other innovations, for example, requiring real-time auditing of AI systems, can also be helpful in exposing and addressing ethical concerns while they are occurring.

Equally important is education for stakeholders in the entire AI lifecycle, from the people building the systems to people applying the systems. Conveying the ethical implications of AI through academic programs, and increasing public awareness of the ethical implications of AI will enable everyone to do their part towards deploying responsible AI.

The journey to ethically responsible AI lies with the continuous improvement through governing, designing, and educating. It can only be done in a concerted effort to hold AI systems to their responsibility to fairness, transparency, and dignity for humanity.

Table 5 Strategic Recommendations by Stakeholder

Stakeholder Group	Recommendation	Impact
Developers & Engineers	Ethics-by-design, Bias detection tools	Ethically robust and inclusive AI systems
Policymakers	Global AI coalitions, Ethical AI certification	Standardization and accountability
Educators & Academia	Embed AI ethics in the curriculum, promote interdisciplinary R&D	Future-ready and responsible AI professionals
Organizations	Real-time audits, Human oversight systems	Governance maturity and stakeholder trust
General Public	Awareness programs, Redress mechanisms	Empowerment and protection of user rights

The table provides a foundational basis for recommendations to various AI ecosystem stakeholders. Developers commit to ethics-by-design and ways to detect bias to develop a fair system. Policymakers support accountability with global regulations and certifications. Educators equip the next generation of AI professionals with a social responsibility, understanding ethical obligations and embedding it as part of their curriculum. Organizations can support trust with audits in consultation with oversight committees. Public awareness, redress, and feedback mechanisms empower users and protect local rights. These recommendations, taken together, serve to diminish unethical outcomes from AI

7. Conclusion

The pace at which artificial intelligence is being incorporated into high-stakes areas vastly exceeds our ability to develop strong data governance frameworks, which creates ethical lapses that result in real consequences. When AI systems are trained on biased, partial, or opaque data and lack actionable accountability mechanisms, the products of these systems can exacerbate inequality, decrease trust, and compromise the values of society.

We have shown that weak data governance and the lack of appropriate legal and ethical protections are a serious vulnerability within the ethical architecture of AI. Weak data governance is not just the root of privacy violations and algorithmic discrimination; it represents a deeply flawed data ecosystem. Creating stronger governance frameworks will need more than simple compliance it will need diligent, multidisciplinary work that focuses on adapting values such as fairness, accountability, and transparency during the complete lifecycle of the AI system.

Going forward, ethical design processes for AI, institutional accountability, and the inclusion of the public should be prioritized in policy and practice. It is going to take a collective effort - inclusive governance approaches and shared ethical standards - to satisfactorily address these challenges and respond to the larger societal question of whether artificial intelligence is (or will) serve the common good or simply replicate or exaggerate existing harms.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare that they have no conflict of interest.

References

- [1] Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy Artificial Intelligence. *Government information quarterly*, 37(3), 101493.
- [2] Guidance, W. H. O. (2021). Ethics and governance of artificial intelligence for health. World Health Organization.
- [3] Guidance, W. H. O. "Ethics and governance of artificial intelligence for health." World Health Organization (2021).
- [4] Attar, A., Raissi, S., & Khalili-Damghani, K. (2017). A simulation-based optimization approach for free distributed repairable multi-state availability-redundancy allocation problems. *Reliability Engineering & System Safety*, 157, 177-191.
- [5] Pike, E. R. (2019). Defending data: Toward ethical protections and comprehensive data governance. *Emory LJ*, 69, 687.

- [6] Bogroff, A., & Guegan, D. (2019). Artificial intelligence, data, ethics an holistic approach for risks and regulation. University Ca'Foscari of Venice, Dept. of Economics Research Paper Series, (19).
- [7] Winfield, A. F., & Jirotko, M. (2018). Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133), 20180085.
- [8] Attar, A., Raissi, S., & Khalili-Damghani, K. (2016). Simulation–optimization approach for a continuous-review, base-stock inventory model with general compound demands, random lead times, and lost sales. *Simulation*, 92(6), 547-564.
- [9] Slota, S. C., Fleischmann, K. R., Greenberg, S., Verma, N., Cummings, B., Li, L., & Shenefiel, C. (2020). Good systems, bad data?: Interpretations of AI hype and failures. *Proceedings of the Association for Information Science and Technology*, 57(1), e275.
- [10] Cheatham, B., Javanmardian, K., & Samandari, H. (2019). Confronting the risks of artificial intelligence. *McKinsey Quarterly*, 2(38), 1-9.
- [11] Attar, A., Raissi, S., Tohidi, H., & Feizollahi, M. J. (2023). A novel perspective on reliable system design with erlang failures and realistic constraints for incomplete switching mechanisms. *IEEE Access*, 11, 51900-51914.
- [12] Floridi, L. (2019). Translating principles into practices of digital ethics: Five risks of being unethical. *Philosophy & Technology*, 32(2), 185-193.
- [13] Agbese, M. (2021). Implementing artificial intelligence ethics in trustworthy systems development: Extending ECCOLA to cover information governance principles.
- [14] Attar, A., Raissi, S., & Khalili-Damghani, K. (2015). Multi-objective reliability-redundancy allocation for non-exponential multi-state repairable components.
- [15] Onoja, J. P., Hamza, O., Collins, A., Chibunna, U. B., Eweja, A., & Daraojimba, A. I. (2021). Digital Transformation and Data Governance: Strategies for Regulatory Compliance and Secure AI-Driven Business Operations. *J. Front. Multidiscip. Res*, 2(1), 43-55.
- [16] Bankins, S. (2021). The ethical use of artificial intelligence in human resource management: a decision-making framework. *Ethics and Information Technology*, 23(4), 841-854.
- [17] Shneiderman, B. (2020). Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 10(4), 1-31.
- [18] Shneiderman, B. (2020). Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 10(4), 1-31.
- [19] Brendel, A. B., Mirbabaie, M., Lembcke, T. B., & Hofeditz, L. (2021). Ethical management of artificial intelligence. *Sustainability*, 13(4), 1974.
- [20] Liaw, S. T., Liyanage, H., Kuziemy, C., Terry, A. L., Schreiber, R., Jonnagaddala, J., & de Lusignan, S. (2020). Ethical use of electronic health record data and artificial intelligence: recommendations of the primary care informatics working group of the international medical informatics association. *Yearbook of medical informatics*, 29(01), 051-057.