Review Article

Statistical Data-Driven Decision-Making Considering Bias, Fairness, and Transparency in AI

[1,*]**T. G. Vasista**

[1] Department of Artificial Intelligence & Data Science, Kakinada Institute of Technology and Science (KITS), Divili, India

* Corresponding Author: tgvasista@gmail.com

Abstract: Bias, fairness, and transparency are critical issues in Artificial Intelligence (AI). These problems can arise from sources such as biased training data, algorithmic bias, and reinforcement learning bias. Bias may lead to unintended consequences while attempting to correct bias. The use of the black-box model, along with proprietary and confidentiality constraints, can further obscure decision-making processes. Regulatory challenges complicate the governance of AI systems. Unfairness can arise when the algorithm uses inappropriate features in AI-based decision-making. Lack of transparency in AI-based computation leads to reduced trust, accountability issues, and difficulty in understanding or challenging automated decisions. Addressing bias, fairness, and transparency in AI is crucial to ensure ethical, responsible, and inclusive technology. Governments, organizations, and researchers must work together to create AI systems that serve humanity without reinforcing discrimination. Without addressing these problems, AI will have to risk inequalities and lose public trust. For example, "if you tell an AI image tool to create a man writing with his LEFT hand, the AI will create a man writing with his right hand" India's PM Modiji pointed it out in a Paris speech. Unfairness can arise when the algorithm uses inappropriate features or a biased training data set to make a decision.

Keywords: addressing AI challenges, bias in AI, challenges in AI, fairness in AI, transparency in AI.

Citation: Vasista, T. G. (2025). Statistical Data-Driven Decision-Making considering Bias, Fairness, and Transparency in AI. *Iota*, 5(2). https://doi.org/10.31763/iota.v5i2.905

Academic Editor: Adi, P.D.P

Received: April 05, 2025

Accepted: April 16, 2025

Published: May 07, 2025

Publisher's Note: ASCEE stays neutral about jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Artificial Intelligence (AI) is revolutionizing statistical-based data-driven decision-making across various fields (Balbaa & Abdurashidova, 2024). Data-driven decision-making is becoming the cornerstone of developing modern organizational strategy (Alimi et al., 2024). Data-driven decision-making uses statistical data to interpret and validate a course of action toward providing decision-making capability (HBS Online, 2021). By leveraging machine learning and deep learning algorithm techniques, AI can analyze large volumes of data quickly and accurately (Pires, 2025).

Key benefits of statistical data-driven decision-making are: (i) Enhanced accuracy (Hossain, 2024) (ii) Statistically based Predictive insights (Hossian, 2024) and forecasting (Li et al., 2022) (iii) Improved efficiency (Li et al., 2022) and Productivity (Sarker, 2021) (iv) Risk management with probabilistic risk assessment (Parhizkar, 2020) and uncertainty reduction (Scholes, 2025), (v) Better customer understanding and market analysis (Haleem et al., 2022) (vi) Cost reduction and revenue optimization (Haleem et al., 2022) (vii) Real-time decision making (Tien, 2017) (viii) Scalability and adaptability (Michael et al., 2024).

While statistical-based data-driven decision-making offers numerous benefits, it also comes with several challenges that must be addressed by organizations to ensure accuracy, reliability, effectiveness, and trust.

They are: (i) Data quality (Aldoseri et al., 2023) (Alimi, 2024) and Statistical reliability (Takeuchi, 2025), (ii) Data/Information overload (Lega et al., 2024) and complex data/information processing (Katipoglu & Keblouti, 2024) (iii) Bias in data and analysis (Aldoseri et al. 2023) (iv) Interpretation and misuse of statistics (Biswas, 2020) (Kate & Emily, 2023) (v) Model accuracy and assumptions (Zumwald, 2021) (vi) Data security and privacy (Aldoseri et al., 2023) (vii) Integration of semantic data (Al-Sudairi & Vasista, 2011) derived out of pattern warehouse for effective decision making (Vasista, 2023) (viii) Legal and ethical considerations (Aldoseri et al. 2023) (ix) Promoting Fairness in AI (Aldoseri et al., 2023) (x) Promoting Transparency in AI (Vidjikant, 2023) (xi) Managing multiple data sources (Alimi, 2024) (xii) Lack of skilled personnel and expertise (Alimi, 2024).

However, the scope of this chapter is limited to focus more on addressing primarily the three considerable challenges in AI such as: (i) Bias in AI (ii) Fairness in AI, and (iii) Transparency in AI, leading to statistically data-driven based decision-making. Thus, three objectives of this chapter are to discuss on (i) What is bias and how it can be addressed in AI, (ii) What is fairness and how it can be addressed in AI, and (iii) What is transparency and how it can be addressed within the context of statistical data-driven decision making. Correspondingly, the chapter is organized to deal with 1. Introduction, 2. Underlying theories for statistical data-driven decision-making, 3. Conceptual underpinning on Bias, Fairness, and Transparency (Definitions and Correlations), 4. Addressing Statistical Data-Driven Bias in AI, 5. Addressing Statistical data-driven fairness in AI, 6. Statistical data-driven transparency in AI, 7. Implications of bias, fairness, and transparency in AI towards data-driven decision making. 8. Conclusion

2. Theory

2.1 Underlying theories for Statistical Data-driven Decision-making

Statistical data-driven decision-making is based on several foundation theories that guide how data is collected, analyzed, and interpreted to support informed decision-making. In this section, some of these relevant theories are discussed.

2.1.1 Probability theory: It is the foundation for statistical inference by modeling randomness and uncertainty (Ramachandran & Tsokos, 2015). The outcome of a random event may be any one of the several possible outcomes, which are determined by chance (Morimer & Blinder, 2024). It is used to measure the uncertainty phenomena. It helps in making predictions based on likelihood using Bayesian theory and inference, which is a modern approach in data science for updating beliefs with new data and making informed predictions (Fofanah, 2024). Bayesian theory is commonly applied in machine learning, decision science, and risk assessment. Further, Regression theory and analysis models, analyze relationships between dependent and independent variables to understand dependencies and are used to make predictions (Mohr, Wilson & Freund, 2022).

2.1.2 Decision Theory: It is the study of choices to make a decision. Decision theory focuses on rationalized decision-making (Peterson, 2011) and outcomes of decisions based on a judgment made referring to pre-determined criteria (Hansson, 2011). Furthermore, decision theory is often described as normative or descriptive. While normative decision models are prescriptive, descriptive decision models are empirical to bring predictive outcomes. With the intervention of AI and emerging technologies, principles of decision theory have been extended towards including information theory, game theory, and system theory with time-based dynamics (e.g. time series analysis) as cited in Elgendy, Elragal & Palvarinta (2021) in passing. These theories when used collectively, enable organizations to make informed and evidence-based choices in data-driven decision-making.

3. Discussion

3.1 Conceptual Framework of Bias, Fairness and Transparency

Bias is defined as a tendency to promote prejudiced results due to erroneous assumptions (dictionary.com) Mavrogiorgos (2024). Bias occurs when systematic errors are introduced into sampling by selecting one outcome over others (Merriam-Webster). The sources of bias can occur during questionnaire design, targeting audience, data collection, data analysis, publication, or during the design of conceptual framework design or research framework design (Pannucci & Wilkins, 2010).

Bias in statistics refers to the tendency to produce systematic errors in data collection, analysis, or modeling that lead to inaccurate or misleading conclusions (Simundic, 2013). It can occur due to sampling bias (e.g. presentation of certain groups), measurement bias (e.g. flawed data collection), or confirmation bias (e.g. selecting data that supports preconceived notions) (dovetail.com, 2024).

Table 1. Types of Statistical Bias, Occurring Stage, occurs due to, How to Avoid (CAS, 2022)

Sl. No.	Type of Bias	Occurring Phase	Occurs due to	How to Avoid
1.	Sampling Bias (Selection bias) (Supervision bias)	Planning	Non-representative data selection & Supervision	Clearly outline the Target audience or beneficiaries, Random sampling for resulting highest abstraction level generalization.
2.	Response bias	Planning	Only considering successful cases (avoiding failure cases). Giving only socially desirable answers	Link the Report to refer to relevant Grounded Theory and Fundamental Principles & Standards.
3.	Recall bias (Participant bias)	Conducting Study	Differences in Accuracies of past collections of events are experiences	Use objective records Shorten recall period
4.	Observation bias (Participant bias)	Conducting Study	Hawthorne effect, intentionally creating bias, More subjectivity in data Inconsistency in recording	Make observers unaware of the participants Use scales, sensors, and software
5	Interviewer bias (Researcher bias)	Conducting Study	Interviewer tone, body language or phrasing of question, facial expressions, demographic differences	Encourage blindfold survey studies, encourage neutral phrase communication, and employ computer-assisted interviews.
6	Measurement bias (Researcher bias)	Conducting Study	Using faulty instruments, inconsistent measurement methods, lack of rationality, data input mistakes, and errors	Calibration of instruments, setting standard operating procedures, mapping subjective measures to objective measures using the Likert scale, use double-entry systems and cross-verifications

Sl. No.	Type of Bias	Occurring Phase	Occurs due to	How to Avoid
7	Confirmation bias (Researcher bias)	Conducting Study	Matching individuals seeking interpretations, favors, beliefs	Awareness, acknowledgment, encourage critical thinking, use of diversified sources, encourage peer and double-blind reviews
8.	Outcome bias	Reporting	Partial results are reported	Focus on results over process, overestimation of predictability after outcomes are known, tied emotions to results,
9.	Publication bias	Reporting	Preference for positive results, selected reporting matter, editorial policies, peer pressures, funding influence	Pre-registering study protocols, encouraging raw data sharing, promoting research platforms that publish valid research, encouraging the publications of null or negative results too, avoiding selected reporting
10.	Spin or Rotated bias	Reporting	Highlighting only favorable results, misrepresentations of results, use of misleading language in abstract, conclusion	Adhering to transparent report guidelines, independent peer review, and normalizing the acceptance of studies.
11.	Reader bias	Post Publication	Perceiving only what the reader wanted, ignoring its semantic integrity	Read relevant well-grounded Theories, Fundamental Principles & Standards

3.2 Addressing statistical data-driven bias in AI

Statistical bias in AI occurs when models produce systematic errors due to imbalanced, incomplete, or misrepresented data (Jui & Rivas, 2024).

3.2.1 Bias in Artificial Intelligence and Machine Learning (AI and ML) Systems

AI and ML systems automate decision-making through complex algorithms trained on large data sets. So, it means conceptually they become additional layers of complexity. From the perspective of machine learning of AI, label bias (Jiang & Nachum, 2020) and algorithm bias (Akter et al. (2022) can occur, where label bias refers to inaccurate or subjective labeling that reflects human prejudices (Chadha, 2024). Algorithm Bias refers to favoring certain patterns that reinforce inequalities from model design and optimization. In AI-ML, bias manifests in training data (due to historical biases and imbalanced datasets in passing to Johnson & Khoshgoftaar, (2019)), model architecture (in the form of inherent assumptions made), and deployment (through feedback loops reinforcing biases). Examples include Amazon's Biased Hiring Algorithm resulting in gender bias (e.g. Women's Chess Club), racial bias in hiring algorithms (Chen, 2023), and facial recognition models (Black defendants and White defendants in court or legal systems related to criminal justice) performing poorly on certain demographics in the context of hiring them for sports, criminal justice and other activities.

3.3 Sources of bias

The cause or source of bias can be caused by the following factors:

3.3.1 *Bias in Historical data*: The historical data that is used for prediction may contain bias

3.3.2 *Biased Questionnaire design during data collection*: The way questions are asked will introduce biased answers for further processing.

3.3.3 *Biased Training Data*: AI models reflect these biases such as (i) Underfitting of bias data and (2) Overfitting of bias data.

3.3.4 *Flawed Algorithmic data*: Due to the faulty design of algorithms bias can be introduced even with unbiased data.

3.3.5 *Human Intervention Bias*: Human professional in various capacities working for AI development may introduce their respective work-related bias.

3.4 Strategies for Mitigating Bias in AI

3.4.1 *Diverse and Representative Data Collection*: Ensure datasets include all relevant demographics to reflect real-world diversity.

3.4.2 *Algorithmic Transparency and Explainability***:** Implement explainable AI techniques to assess the decision-making process (Gonzalez-Sendino & Serrano & Bajo, 2024)

3.4.3 *Bias correction Techniques***:** Use reweighting or equal weighting (Dave, 2023), auditing training data (Chadha, 2024), and adversarial debiasing to minimize bias in model training (Gonzalez-Sendino, Serrano & Bajo, 2024).

3.4.4 *Human-in-the-loop approaches*: Engage domain experts in reviewing and correcting biases in AI development at various stages such as data extraction, data cleansing, preprocessing, integration, annotation, labeling, training, and inferencing (Chai & Li, 2020; Wu et al., 2022; Memarian & Doleck, 2024).

3.5 Addressing statistical data-driven fairness in AI

Fairness in Artificial Intelligence and Machine Learning (AIML) is emerging as a critical concept, which influences diverse perspectives of society such as healthcare, legal judgments, etc. (Uddin, Lu, Rahman & Gao, 2024).

Fairness is a concept, where every person should get equal opportunities and treatment (Chadha, 2024), in AIML, equal opportunities refer to ensuring a true positive rate (sensitivity) and equal false-positive rates across groups. Demographic parity refers to equal outcomes across groups (Ferrara, 2024).

Fairness in statistical decision-making means ensuring that models and analyses do not disproportionately disadvantage certain groups (Rabaey, De Schutter, De Brant & Derudder, 2019). Statistical discrimination describes a set of informational issues that can induce a rationale of say a Bayesian decision-making leading to unfair outcomes event in the absence of discriminatory intent (Patty, & Penn, 2022).

Fairness involves techniques such as stratified sampling, fairness-aware modeling, and ensuring that decisions are equitable across different demographic segments (Mehrabi, Huang & Morstatter, 2020).

Fairness in Artificial Intelligence and Machine Learning ensures that models do not spread or magnify discrimination (Mensah, 2023). Fairness in AIML can be achieved by correcting algorithmic bias during the automated process models of machine learning. Unfairness can arise when the algorithm uses inappropriate features or biased training data sets to make decisions (Jui & Rivas, 2024). Fairness can be assessed statistically using t-tests and evaluated based on k-fold cross-validation (Uddin, Lu, Rahman & Gao, 2024)

3.6 Types of Fairness

3.6.1 *Individual Fairness*: Individual fairness refers to intuitive principles and similar treatment. Similar treatment refers to having individuals be treated similarly (Fleisher, 2021). Characterizing individual fairness requires having measures that assess fairness or equivalence and bias (Anderson, Visweswaran, 2025).

3.6.2 *Group Fairness***:** It is the fairness concern by evaluating and mitigating measures of group discrimination (Krasanakis & Papadopoulos, 2024) by AI systems. Group fairness refers to equal treatment of different groups by AI systems (Ferrara, 2024).

3.6.3 *Demographic parity/Statistical parity*: It is a fairness metric, which says that if the composition of people selected by the model matches the group membership of the applicants, then the model is said to be fair (Kaggle) or to better understand another way is that demographic parity requires equal proportions of positive predictions in each group (GitHub). Equal Opportunity fairness is when the proportion of people selected by the model is the same for each group (i.e. indicated by the true positive rate or sensitivity of the model). Equal Accuracy Fairness ensures that the model has equal accuracy for each group. Group unaware fairness requires the removal of all group membership information from the dataset (e.g. removing the gender label). While the fairness of demographic parity, equal opportunity, and equal accuracy can be measured using a confusion matrix, group-unaware fairness cannot be detected from the confusion matrix in passing to (Kaggle). Example: Suppose the admissions model accepts 32 candidates from a majority group and 8 candidates from a minority group. The model's decision satisfies the demographic parity, as the acceptance rate for both majority and minority candidates is 40 percent (from a total majority group of 32/80 and minority group of 8/20) (developers.google.com-fairness1).

3.6.4 *Counterfactual Fairness***:** Counterfactual fairness is derived from Pearl's causal model. It considers that the model is fair when the prediction of a particular individual or group in the real-world domain is the same as that in the counterfactual world domain, however, the individuals have to belong to different demographic groups (Wu, Zhang & Wu, 2019). AI would have made the same decision for an individual regardless of their group membership i.e. majority group and minority groups (Ferrara, 2024). For some people who had written UPSC exam group-1 and group-2 then we have two broad groups one is people accepted two is people rejected. The rejected people are many. These rejected people may contain applicants from different religions with different demographics but still one attribute that qualifies the same value i.e. result status = rejected. It is fair that they can be rejected irrespective of religion, caste reservation category, and general category. While the demographic features of these groups can be different they all belong to one group called the rejected group (developers.google.com-fairness2).

3.6.5 *Procedural Fairness*: Procedural fairness is about the fairness of the procedures used by a decision-maker while making a decision. When a fair procedure is followed, the decision maker will make a correct and fair decision (ombudsman.was.au). It emphasizes the importance of aligning fair decision-making procedures to its underlying theories of relational justice (Decker, Wggner, & Leicht-Scholten, 2025).

3.7 Sources of Unfairness

One of the underlying causes of unfairness is bias in training data (Balayn, Lofi & Houben, 2021). Various kinds of bias that represent a source of unfairness are as follows (O'Sullivan, 2022).

3.7.1 Data Related sources of unfairness**: -** Sampling bias, Historical bias, Measurement bias, label bias

3.7.2 Algorithmic and Modelling Bias: **-** Feature Selection bias, Model Overfitting bias, Objective Function Bias.

3.7.3 Interpretation and Evaluation Bias: **-** Confirmation bias, Ignoring Subgroup Performance bias
3.7.4 Deployment and Feedback loop Bias: - Automation bias, Feedback loop bias
3.7.5 Socio-Technical Bias: - Institutional bias, Lack of diversity in development teams' bias

3.8 Strategies for mitigating unfairness

(1) Use fair-ness algorithms (2) Evaluated with fairness metrics like demographic parity, equal opportunity, etc. (Medda, 2024) (3) Include diverse data sources and participatory design (4) Conduct bias audits at every stage

3.9 Strategies for Achieving Better Fairness

Fairness in AIML can be addressed and achieved by using fairness-aware algorithms, debiasing techniques, and enforcing ethical considerations and guidelines for avoiding bias and unfair means, whereas fairness-aware algorithm focuses on developing algorithms and debiasing techniques that ensure fairness and mitigate bias in machine learning models (Palvel, 2023) and ethical considerations guidelines such as promoting and transparency, authenticity, accountability, trust, privacy and protecting through intellectual property rights (Al-fairy, Mustafa, Kshetri, Insiew & Alfandi, 2024).

3.10 Addressing statistical data-driven transparency in AI

Transparency from Statistics Perspective: Transparency refers to the clarity and openness of statistical methods, data sources, and assumptions used in decision-making. It involves documenting methodologies, trustworthiness, and reproducibility.

Transparency from AIML Perspective: Transparency referred to as explainability, involves making models interpretable so that users and regulators understand why a system makes certain predictions. This is critical in high-stakes domains like healthcare, finance, and criminal justice. Techniques like SHAP values, LIME, and model documentation help improve transparency.

3.11 Types Transparency

3.11.1 Data Transparency: - It provides visibility of data to train AI systems
3.11.2 Consent Transparency: - It informs users, how their data might be used across AI systems
3.11.3 Algorithmic Transparency: - This means making the data, logic, and rationale used by the AI system understandable and accessible for producing insights and decisions.
3.11.4 Model Transparency: - It reveals how the AI system functions, possibly by explaining decision-making processes

3.12 Sources of Opacity

3.12.1 Content Opacity: It occurs when an AI system information prevents stakeholders from grasping its semantics and using it for their purpose

3.12.2 Inferential Opacity: It occurs when an AI system prevents stakeholders from better understanding the sense of the reasoning path (Facchini & Termine, 2021).

3.13 Strategies for mitigating or reducing opacity

3.13.1 Technical Solutions: When automation acts like a black box, for people working with an algorithm-based system, the opacity of the black box can undermine adequate trust in system outputs, thus weakening the decision-making capability (Tintarev & Masthoff, 2007) as cited in passing to Langer & Konig, 2023).
3.13.2 Education & Training: Education &Training data disclosure is critical for monitoring financial management and pedagogical accountability.

Imparting training to the school management committee, faculty, parents and selected community groups is important in how data can be used in maintaining accountability. Introducing legal grievance redressal mechanisms helps in better managing the varied data gathered (UNESCO, 2018).

3.13.3 Regulations & Guidelines: Opacity can serve both as a control mechanism and as a conflict mechanism with legal regulations (Goodman & Flaxman, 2017). It is important to outweigh the benefits of transparency with the existing opacity that is acting as a control mechanism as sometimes protecting information in the form of privacy and confidentiality becomes important to safeguard data from unauthorized disclosure, misuse exploitations, and theft (Chapple, 2019).

3.14 Strategies for improving transparency

3.14.1 Data Sources & Modelling: Maintain data sources, model architecture, and preprocessing steps and hyper-parameters (Pillai, 2024)

3.14.2 Develop Explainable AI Techniques: Identify the most influential features in a model's predictions; employ visualization to better understand the model's decision-making process; generate human-understandable explanations to AI predictions; improve interpretability based developing and using the combination of both simple and complex model development (Pillai, 2024).

3.14.3 Stakeholder Engagement: Keep discussing transparent communication with stakeholders; develop feedback mechanisms, focus on user/customer-centric design; participate in developing ethical standards and best practices (Coleman, Manyindo, Parket & Schultz, 2019).

3.14.4 Continuous Monitoring and Improvement: Develop dashboards for continuously monitoring the model performance and identifying potential issues; focus on bias detection and mitigating it; establish auditing AI systems to ensure accuracy, fairness, and reliability; use iterative techniques to keep refining the model towards better optimization (Thanasas, Kampiotis & Karkantzou, 2025).

3.15 Implications of bias, fairness, and transparency in AI towards data-driven decision making

The overall impact of bias, fairness, and transparency on data-driven decision-making include:

3.15.1 Improves ethical integrity, regulatory compliance, and public confidence (Radanliev, 2024)

3.15.2 Ensures decisions are justifiable, inclusive, and socially responsible (Singhal, Neveditsin, Tanveer & Mago, 2024).

3.15.3 Encourages the adoption of responsible AI practices in governments and organizations (Alabi, 2024).

3.16 Implications of Bias in AI towards data-driven decision making

3.16.1 Unfair outcomes: Biased AI models can lead to discrimination against certain groups (e.g. in hiring, lending, and health care) (Ferrara, 2023)

3.16.2 Loss of Trust: Users and stakeholders lose confidence in systems perceived as biased (Fancher, Ammanath, Holdowsky & Buckley, 2021).

3.16.3 Legal Risk: Bias may lead to violation of anti-discrimination laws and result in regulatory penalties (Hilliard, Gulley, Koshiyama & Kazim, 2024)

3.17 Implications of Fairness in AI towards data-driven decision making

3.17.1 Equitable Access: Fair AI ensures equal treatment and benefits for diverse user groups (Gonzlez-Sendino, 2024) (Willie, 2024).

3.17.2 Inclusive Innovation: Encourages design of solutions that cater to the needs of all not just the majority (Persson, Ahman, Yngling & Gulliksen, 2014).

3.17.3 Social Good: Promotes ethical technological development and reduces structural inequalities (Capraro et al., 2024).

3.18 Implications of Transparency in AI towards data-driven decision-making

3.18.1 **Accountability:** Transparent models make it easier to trace, audit, and correct errors or unintended outcomes (Cheong, 2024).

3.18.2 **Explainability:** Helps stakeholders understand and trust AI decisions, especially in important areas like healthcare and criminal justice (van der Veer et al., 2021).

3.18.3 **Improved Governance:** Enables better monitoring and regulation of automated decision-making systems (Margetts, 2022).

4. Conclusion

In the era of AI-driven systems, statistical data-driven decision-making holds immense promise to transform sectors through efficiency, scalability, and precision. This kind of potential can only be realized when bias, fairness, and transparency become integral parts of the design and implementation of AI systems. Bias, whether stemming from data, algorithms, or human oversight, can skew outcomes, reinforcing social inequalities. Fairness demands that models account for diverse perspectives and treat all individuals and groups equitably. Transparency enables stakeholders to understand, trust, and hold AI systems accountable for their actions. For data-driven decision-making to be trustworthy and ethical, it is essential to adopt practices that minimize bias, ensure fair and equitable treatment, and promote clarity and accountability. When these principles are embedded into AI development, it not only enhances the reliability of decisions but also safeguards public trust and ensures that technological advancement contributes to inclusive societies.

Acknowledgments: Thanks to all the staff, faculty, and academics of the Department of Artificial Intelligence & Data Science, Kakinada Institute of Technology and Science (KITS), Divili, India, hopefully this review paper can be a good reference in understanding and building AI now and in the future.

Author contributions: The authors are responsible for building Conceptualization, Methodology, analysis, investigation, data curation, writing—original draft preparation, writing—review and editing, visualization, supervision of project administration, funding acquisition, and have read and agreed to the published version of the manuscript.

Funding: The study was conducted without any financial support from external sources.

Availability of data and Materials: All data are available from the authors.

Conflicts of Interest: The authors declare no conflict of interest.

Additional Information: No Additional Information from the authors.

References

[1] Akter, S., Dwivedi, Y. K., Sajib, S. S., Biswas, K., Bandara, R. J. & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. Journal of Business Research, 144, 201-216, https://doi.org/10.1016/j.busres.2022.01.083

[2] Al-Kafairy, M., Mustafa, D., Kshetri, N. Insiew, M. & Alfandi, O. (2024). Ethical challenges and solutions of Generative AI: An Interdisciplinary Perspective, Informatics, 11 (3), 58, https://doi.org/10.3390/informatics11030058

[3] Al-Sudairi, M. T. & Vasista, T. G. K. (2011). Semantic data integration approaches for E-Governance, International Journal of Web & Semantic Technology, 2(1), 1-12, DOI: 10.5121/ijwest.2011.2101

[4] Alabi, M. (2024). Ethical Implications of AI: Bias, Fairness and Transparency, [Available at] https://www.researchgate.net/publication/385782076_Ethical_Implications_of_AI_Bias_Fairness_and_Transparency Retrieved on 11-04-2025.

[5] Alimi, A., Grace, C., Oladele, S. & Mia J. (2024). Data-driven decision-making: Challenges and Opportunities in Business analytics. [Available at] https://www.researchgate.net/publication/387365368_Data-Driven_Decision-

Making_Challenges_and_Opportunities_in_Business_Analytic Retrieved on 13-03-2025
[6] Anderson, J. W. & Visweswaran, S. (2025). Algorithmic individual fairness and healthcare: a scoping review, JAMIA Open, 8(1), https://doi.org/10.1093/jamiaopen/00ae149
[7] Balbaa, M. & Abdurashidova, M. (2024). The impact of Artificial Intelligence in Decision making: A comprehensive review, EPRA International Journal of Economics, Business and Management Studies, 11 (2): 27-38, DOI: 10.36713/epra15747
[8] Balayn, A., Lofi, C. & Houben, G. (2021). Managing bias and unfairness in data for decision support: a survey of machine learning and data engineering approaches to identify and mitigate bias and unfairness within data management and analytics systems. The VLDB Journal, 30, 739-768, https://doi.org/10.1007/s00778-021-00671-8.
[9] Biswas, A. (2020). The misuse of Statistics. Opinion Statistics [Available at] https://pub.towardsai.net/misuse-of-statistics-c2b78b67a3c Retrieved on 13-03-2025
[10] Caprero, V. et al., (2024). The impact of generative artificial intelligence on socioeconomic inequalities and policy-making, PNAS Nexus, 3 (6), 191, https://doi.org/10.1093/pnasmexus/page191
[11] CAS (2022). Biases at Different Stages of Research – and ways to avoid them. Charlesworth Author Services. [Available at[https://www.cwauthors.com/article/avoiding-biases-at-different-stages-of-research Retrieved on 19-03-2025.
[12] Chadha, K. S. (2024). Bias and Fairness in Artificial Intelligence: Methods and Mitigation Strategies. International Journal for Research Publication and Seminar, 15(3), 36-49.
[13] Chai, C. & Li, G. (2020). Human-in-the-loop Techniques in machine learning. IEEE Data Engineering Bulletin, 43 (31), 37-52.
[14] Chapple, M. (2019/10). Privacy vs Confidentiality vs Security: What's the difference? EdTech Magazine, [Available at] https://edtechmagazine.com/higher/article/2019/10/security-privacy-and-confidentiality-whats-difference#:~:text=Confidentiality%20controls%20protect%20against%20the,maintains%20and%20shares%20with%20others. Retrieved on 10-04-2025.
[15] Chen, Z. (2023). Ethics and discrimination in artificial intelligence-enabled recruitment practices, Humanities and Social Sciences Communications 10, 567, https://doi.org/10.1057/s41599-023-02079-x
[16] Cheong, B. C. (2024). Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision making. Frontiers Human Dynamics, 6, https://doi.org/10.3389/fhumd.2024.1421273
[17] Coleman, E. A., Manyindo, J., Parker A. R. & Schultz, B. (2019). Stakeholder engagement increases transparency, satisfaction, and civil action. PNAS, 116 (49) 24486-24491, https://doi.org/10.1073/pnas.1908433116.
[18] Decker, M., Wegner, L. & Leicht-Scholten, C. (2025). Procedural fairness in algorithmic decision-making: the role of public engagement. Ethics and Information Technology, 27(1), https://doi.org/10.1007/s10676-024-09811-4
[19] developers.google.com-fairness1, Fairness: Demographic Parity, [Available at] https://developers.google.com/machine-learning/crash-course/fairness/demographic-parity Retrieved on 07-04-2025
[20] developers.google.com-fairness2, Fairness: Counterfactual fairness, [Available at] https://developers.google.com/machine-learning/crash-course/fairness/counterfactual-fairness Retrieved on 07-04-2025.
[21] Dictionary.com, Bias definition and meaning - https://www.dictionary.com/browse/bias retrieved on 19-03-2025.
[22] Dovetail.com (2024), Understanding different types of bias in research (2024) guide. [Available at] https://dovetail.com/research/types-of-bias-in-research/ Retrieved on 27-03-2025.
[23] Elegendy, N., Elragal, A. & Palvarinta, T. (2022). DECAS: a modern data-driven decision theory for big data and analytics, Journal of Decision Systems, 31 (4), 337-373.
[24] Facchini, A. & Termine, A. (2021). Towards a Taxonomy for the opacity of AI systems. [Available at] https://philsci-archive.pitt.edu/20376/1/Facchini_Termine.pdf Retrieved on 07-04-2025
[25] Fancher, D., Ammanath, B., Holdowsky, J. & Buckley, N. (2021). AI model bias can damage trust more than you may know. But it doesn't have to. [Available at] https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/ai-model-bias.html Retrieved on 11-04-2025.
[26] Ferrara, E. (2024). Fairness and Bias in Artificial Intelligence. A Brief Survey of Sources, impacts and Mitigation Strategies., Scie, 6 (1), 3; https://doi.org/10.3390/sci601003
[27] GitHub. 5. Parity Measures. [Available at] https://afraenkel.github.io/fairness-book/content/05-parity-measures.html Retrieved on 07-04-2025.
[28] Gonzlez-Sendino, R., Serrano, E. & bajo, J. (2024). Mitigating bias in artificial intelligence: Fair data generation via causal models for transparent and explainable decision-making, Future Generation Computer Systems, 155, 384-401.
[29] Goodman, B. & Flaxman, S. (2017). European Union Regulations on Algorithmic Decision Making and a Right to Explanation, AI Magazine, 38 (3), 3-112.
[30] Hansson, S. O. (2011). Decision Theory: An overview in M. Lovric (Ed.) International encyclopedia of statistical science, 349-355. Springer. http://doi.org/10.1007/978-3-642-04898-2_22
[31] HBSOnline (2021), Business Insights. [Available at] https://online.hbs//.edu/blog/post/data-driven-decision-making Retrieved on 13-03-2025
[32] Hilliard, A., Gulley, A., Koshiyama, A. & Kazim, E. (2024). Bias audit laws: how effective are they at preventing bias in automated employment decision tools? International Review of Law, Computers & Technology, 1-17,

https://doi.org/10.1080/13600869.2024.2403053
[33] Hossain, M. A. (2024), Data-driven decision making: Enhancing Quality management practices through optimized MIS frameworks. Innovatech Engineering Journal, 1 (01): 117-135, AIM International LLC Publisher.
[34] Jiang, H. & Nachum, Ofr (2020). Identifying and Correcting Label Bias in Machine Learning, Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics, PMLR 108:702-712.
[35] Johnson, J. M. & Khoshgoftaar (2019). Survey on deep learning with class imbalance, Journal of Big Data, 6 (27), https://doi.org/10.1186/s40357-019-192-5
[36] Jui, T. D. & Rivas, P. (2024). Fairness issues, current approaches, and challenges in machine learning models. International Journal of Machine Learning and Cybernetics, 15, 3095-3125
[37] Kaggle. AI Fairness, [Available at] https://www.kaggle.com/code/alexisbcook/ai-fairness Retrieved on 07-04-2025.
[38] Kate, W. & Emily, B. (2023). AI tools: A powerful new weapon to combat the misuse of statistics. Statistical Journal of the IAOS, 39 (2), 431-437.
[39] Katipoglu, O. M. & Keblouti, M. (2024). Comparative analysis of data-driven models and signal processing techniques in the monthly maximum daily precipitation prediction of ElKarma station Northeast of Algeria. Neural networks, 28, 10751-10765
[40] Krasanakis, E. & Papadopoulos, S. (2024). Evaluating AI Group Fairness: a Fuzzy Logic Perspective. Journal of Artificial Intelligence Research, 1-32. [Available at] https://arxiv.org/pdf/2406.18939 Retrieved on 07-04-2025.
[41] Langer, M. & Konig, C. J. (2023). Introducing a multi-stakeholder perspective on opacity, transparency, and strategies to reduce opacity on algorithm-based human resource management, Human Resource Management Review, 33(1), 100881, https://doi.org/10.1016/j.hrmr.2021/100881
[42] Lega, M. et al. (2024). Reducing information overload in e-participation: A data-driven prioritization framework for policymakers. International Journal of Information Management Data Insights, 4(2), 100264, https://doi.org/10.1016/j.jjimei.2024.100264
[43] Li, L., Lin, J., Ouyang, Y. & Luo, X. (2022). Evaluating the impact of big data analytics usage on the decision-making quality of organizations. Technological Forecasting and Social Change, 175, 121355, https://doi.org/10.1016/j.techfore.2021.121355
[44] Margetts, H. (2022). Rethinking AI for Good Governance, https://doi.org/10.11162/DAED_a_01922 [Available at] https://www.researchgate.net/publication/360304761_Rethinking_AI_for_Good_Governance Retrieved on 11-04-2025.
[45] Mavrogiorgos, K. et al. (2024). Bias in Machine Learning: A Literature Review. Applied Science, 14(19), 1-40. https:/doi.org/10.3390/app/14198860
[46] Medda, G. (2024). Unfairness Assessment, Explanation and Mitigation in Machine Learning Models for Personalization, Ph.D. D. Thesis submitted to University of Cagliari, Italy.
[47] Mehrabi, N., Huang, Y. & Morstatter, F. (2020) Statistical Equity: A Fairness Classification Objective, https://doi.org/10.48550/arXiv.2005.07293.
[48] Memarian, B., & Doleck, T. (2024). Human-in-the-loop in artificial intelligence in education and entity-relationship (ER) analysis. Computers in Human Behavior: Artificial Humans, 2(1), 100053, Jan-Jul 2024.
[49] Mensah, G. B. (2024). Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency and Accountability in AI Systems, DOI: 10.13140/RG.2.2.23381.19685/1
[50] Merriam-Webster. Bias Definition & Meaning. https://www.merriam-webster.com/dictionary/bias Retrieved on 19-03-2025.
[51] Michael, C. I. et al. (2024). Data-driven decision making in IT: Leveraging AI and Data Science for business intelligence, World Journal of Advanced Research and Reviews, 23 (01), 472-480.
[52] Mortimer, R. G. & Blinder, S. M. (2024). Chapter 16- Probability, Statistics, and Experimental Errors. In Mathematics for Physical Chemistry, Fifth Edition, 197-212, Elsevier. https://doi.org/10.1016/B978-0-44-318945-6.00021-1
[53] Ombudsman.wa.au (2019, April), Ombudsman Western Australia, Guidelines: Procedural fairness (natural justice), [Available at] https://www.ombudsman.wa.gov.au/Publications/Documents/guidelines/Procedural-fairness-guidelines.pdf Retrieved on 07-04-2025.
[54] O'Sullivan, C. (2022). Unfair Predictions: 5 Common sources of bias in Machine Learning, [Available at] https://medium.com/data-science/algorithm-fairness-sources-of-bias-7082e5b78a2c Retrieved on 08-04-2025.
[55] Ramachandran, K. M. & Tsokos, C. P. (2015). Chapter 2: Basic Concepts from Probability Theory. In Mathematical Statistics with Applications in R. Second Edition, Academic Press. https://doi.org/10.1016/B978-0-12-417113-8.00002-3
[56] Sarker, I. H. (2021). Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-making and Applications Perspective. SN computer science, 2(5), 377, https://doi.org/10.1007/s42979-021-00765-8
[57] Palvel, S. (2023). Introduction to fairness-aware ML, [Available at] https://subashpalvel.medium.com/introduction-to-fairness-aware-ml-327df1b61538 Retrieved on 01-04-2025
[58] Pannucci, C. J. & Wilkins, E. G. (2010). Identifying and Avoiding Bias in Research, Plastic and Reconstructive Surgery 126(2), 619-625. DOI: 10.1097/PRS.0b013e318de24bc
[59] Parhizkar, T. et al. (2020). A data-driven approach to risk management and decision support for dynamic positioning systems. Reliability Engineering & System Safety, 201, 106964.
[60] Persson, H. Ahman, H., Yngling, A. A. & Gulliksen, J. (2014). Universal design, inclusive design, accessible design, design for

all: different concepts goal? On the concept of accessibility-historical, methodological and philosophical aspects, Universal Access in the Information Society, 14(4), DOI: 10.1007/s10209-014-0358-z

[61] Peterson, M. (2011). Decision theory: An introduction. In M. Lovric M. (Ed.), International Encyclopaedia of Statistical Science, 349-356, Springer, https://doi.org/10.1007/978-3-642-04898-2_23

[62] Piers, J. (2025). The role of AI in Data-driven decision making. Blog @ https://rauva.com/blog/role-ai-data-driven-decision-making Retrieved on 13-03-2025).

[63] Pillai, V. (2024). Enhancing the transparency of data and ml models using explainable AI (XAI), World Journal of Advanced Engineering Technology and Sciences, DOI: 10.30574/wjaets.2024.13.1.0428

[64] Rabaey, P., De Shutter, M., De Brant, T. & Derudder, M. (2019). Fairness in Data Science, [Available at] https://www.researchgate.net/publication/352021231_Fairness_in_Data_Science Retrieved on 01-04-2025.

[65] Radanliev, P. AI Ethics: Integrating Transparency, Fairness, and Privacy in AI Development, Applied Artificial Intelligence, 39(1). https://doi.org/10.1080/08839514.2025.2463722

[66] Scholes, M. S. (2025). Artificial Intelligence and Uncertainty, Risk Sciences, 1, 100004, https://doi.org/10.1016/j.risk.2024.100004

[67] Simundic, A-M. (2013). Bias in research, Biochemia Medica, 23 (1).

[68] Singhal, A., Neveditsin, N., Tanveer, H. & Mago, V. (2024). Toward fairness, Accountability, Transparency, and Ethics in AI for social media and Health Care: Scoping Review, JMIR Medical Informatics, 3(12:e50048), doi: 10.2196/50048.

[69] Takeuchi, I. (2025). Statistical reliability of data-driven science and technology. IEEJ Transactions on Electrical and Electronic Engineering. https://doi.org/10.1002/tee.24262

[70] Tien, J. M. (2017). Internet of Things, Real-time decision making and artificial intelligence, Annals of Data Science, 4, 149-178, https://doi.org.10.1007/s40745-017-0112-5

[71] Thanasas, G. Kampiotis, G. & Karkantzou, A. (2025). Enhancing transparency and efficiency in Auditing and regulatory compliance with disruptive technologies. Theoretical Economics Letters, 15, 214-233, Doi: 10.4236/tel.2025.151013.

[72] Tintarev, N. & Masthoff, J. (2007). A Survey of Explanations in Recommender Systems. In 23rd International Conference on Data Engineering Workshop, IEEE Explore Proceedings, DOI:10.1109/ICDEW.2007.4401070

[73] UNESCO (2018), 10 ways to promote transparency and accountability in education. [Available at] https://etico.iiep.unesco.org/en/10-ways-promote-transparency-and-accountability-education Retrieved on 10-04-2025

[74] van der Veer, S. N. et al. (2021). Trading off accuracy and explainability in AI decision-making: findings from 2 citizens' juries, J Am Med Inform Assoc. 28 (10), 2128-2138, doi: 10-1093/jamia/ocab127.

[75] Vasista, T. G. (2023). Knowledge on Demand Approach using Business Intelligence and Ontology, Scope, 13 (3), 785-796.

[76] Vidjikant, S. (2023). Data-Driven Decision Making, Softjourn [Available at] https://softjourn.com/insights/data-driven-decision-making Retrieved 13-03-2025.

[77] Willie, A. (2024). Fairness in AI: Developing Methods to ensure equitable Treatment across diverse demographic groups. [Available at] https://www.researchgate.net/publication/387381784_Fairness_in_AI_Developing_Methods_to_Ensure_Equitable_Treatment_Across_Diverse_Demographic_Groups Retrieve on 11-04-2025.

[78] Wu, X. Xiao, L., Sun, Y. Zhang, J., Ma, T. & He, L. (2022). A Survey of human-in-the-loop for machine learning, Future Generation Computer Systems, 135, 364-381.

[79] Wu, Y., Zhang, Lu & Wu, X. (2019). Counterfactual fairness: Un-identification, Bound, and Algorithm. Proceedings of the 28th International Joint Conference on Artificial Intelligence Main track, 1438-1444. https:/doi.org/10.24963/ijcai.2019/199

[80] Zumwald, M. et al. (2021). Assessing the representational accuracy of data-driven models: The case of the effect of urban green infrastructure on temperature. Environmental Modelling & Software, 141, 105048, https://doi.org/10.1016/j.envsoft.2021.105048