

Ethical Implications of Biases in AI and Machine Learning Algorithms

Niev Sanghvi

The Bishop's School Pune, Maharashtra,
India

Niel Sanghvi

The Bishop's School Pune, Maharashtra,
India

Naman Sanghvi

The Bishop's School Pune, Maharashtra,
India

Anish Porwal

S.M Choksey Jr.College Pune,
Maharashtra, India

Nikhil Ravindra Gayakwad

Aayan Multitrade LLP

Aastha Abhijeet Sapor

University Of Essex

Abstract— This research paper explores the ethical implications of biases in artificial intelligence (AI) and machine learning systems. The proliferation of artificial intelligence and machine learning techniques into various aspects of today's society has raised concerns about the potential biases inherent in these systems. This research paper examines the impact of biases on decision-making processes, particularly in sensitive areas such as health care, criminal justice and finance. In addition, the root causes of biases in AI algorithms are explored and ethical guidelines and mitigation strategies are proposed to address these issues. By exploring the ethical implications of AI and machine learning biases, this article contributes to the ongoing debate about the responsible and fair adoption of these technologies in society.

I. INTRODUCTION

A. Significance of Artificial Intelligence:

The scientific use of Artificial Intelligence by humans has made many astonishing discoveries and inventions at the present. It serves as the building block of Computer Science acquiring skills, industries, IT departments, healthcare, education, etc. Companies also use AI as an assistant to enhance user experience, smartphones are now built and designed with these concepts for more friendly experiences. AI also aids in doing repetitive tasks such as call service centers so that humans can centralize on other priorities, furthermore AI is also applied in Games by the developers to enhance the immersion of the game, and also the non-playable characters in the game are also under the influence of AI with this they are able to communicate and interact with the player according to the players response in game and also the players interaction with game. AI also assist in the Militaries by providing significant support in the battlefield, combat by its detection of intruders, drones cyber warfare etc. This helps the soldiers in any situation of combat, search or rescue operations. Finance, net banking and all other activities related to banks are handled by AI in Fincorp and other companies. The Industries utilize the aid of AI 24/7 for any financial aid, customer care, bank fraud, stock bond and also the predictive side of AI helps them prevent any future financial

risk. Another significant use of AI is that it averts online frauds, scams etc., in addition it also reduces Human errors with its lightning error detection and fierce accuracy. Forecasting the trajectory of artificial intelligence is a challenging task. In the 1990s, the sole goal of artificial intelligence was to improve human circumstances. But will that be the only objective going forward? Research is focused on building robots or devices that resemble humans. This is a result of scientists' fascination with human intellect and their amazement at attempts to replicate it. The role that humans play will undoubtedly alter if machines begin performing tasks that currently need human labor. One day, researchers' efforts may pay off, and we'll find that robots and computers are doing our work while we stroll about.[1]

A. Significance of Machine Learning:

The Financial industries are mastering the utilization of machine learning for their elevating growth and success rate in the market. American Express has embraced Machine learning in the detection of frauds and other digital threats. Credit Card companies are using Machine learning to access the information of the person needing a loan such as credit score, credit history, rent payment, etc. This information helps the companies prevent risk of losing money and also aids the consumer for his approval of loan unlike the old traditional methods for debt.

The Pharmaceutical companies along with the healthcare industries are using Machine Learning for managing medical information, records etc. This helps the experts to discover new medicines, diseases and also predict and perform a perfect diagnosis. The new updated medical systems can now pull pertinent health information in a blink of an eye. Scientists and drug developers are able to produce drugs at a faster rate with the help of machine learning and other AI tools.

The Retail and E-commerce industries are relying on the use of machine Learning. The analysis of machine learning when the consumer purchases an item from the company helps the

consumer again come back for purchase and thus maintains customer retention rate and boosts the company's sales.

Visual search of the desired product plays a vital role in the sales part of a company, due to the ability of machine learning. A customer's product can be easily identified by the imagery and thus provide a more user friendly experience to the consumer.

Machine learning also keeps a track of the new trends and behavior of people in the media and accordingly provides the supply for the demand. Machine Learning also creates advertisements and other such marketing schemes to highlight the product so that it can be easily available to the customer.

B. Significance of Data Science:

Data Science integrates artificial intelligence, machine learning, mathematics and statistics to analyze and understand the trends and patterns which helps the industries in their growth and succession. The use of data science helps the industries in decision making and also in predicting financial risks in future, the interpretation of trends done by data science also helps in preventing and optimizing malware, cyber threat and any other such activities.

Data Science helps in better analysis of trends and patterns and also aids in providing a better customer experience.

Data Science also helps in generating recommendations companies such as Netflix, Spotify and Amazon use this technique to generate new recommendations, sponsorships and advertisements. This active use of Data Science in everyday life improves the user interaction and also helps mankind prioritize other things.

Data visualization is another asset mastered by data science that helps the non-technical business leaders understand the state of their business.

Cybersecurity is another asset owned by data science. The use of data science prevents any transactional fraud in any institution, etc. and also provides security from any malicious software entering.

Data Science involves a 5 step procedure:

Capture: Capturing raw and unstructured data- data extraction.

Maintain: The maintenance stage includes data warehousing, data cleansing, data staging, data processing and data architecture.

Process: The data is examined to interpret the trends.

Analyze: This stage is when multiple types of analyses are performed on the data

Communicate: This stage is when data scientists showcase their data through reports, charts and graphs.

Considered as the great frontier with the ability to improve nearly every area of our lives, "the Internet of things" (IoT) is a new technological sector that makes every electronic device smarter. Due to its ability to identify patterns in data and build models that forecast future behavior and events, machine learning has emerged as a crucial technology for Internet of Things applications. One of the main applications of the Internet of Things is in "smart cities," which use technology to enhance public services and the quality of life for their residents. For example, with the right data, data science techniques may be used to

forecast traffic patterns in smart cities and calculate the overall energy usage of inhabitants over a given amount of time[2]. On Deep learning-based data science models may be built on top of massive IoT information. Many IoT and smart city services, such as smart governance, home automation, learning, communication, transportation, business, farming, health care, and industry, among others, maybe able to be modeled with the help of data science and analytics techniques.[3]

D. Ethical Implications:

Ethical consequences refer to the possible consequences of actions, decisions or policies on various ethical principles and values. In fields as diverse as, but not limited to, technology, health, business and research, ethical implications play a crucial role in guiding responsible and sustainable practices. The Importance of Ethical Consequences

Rights Protection: Consideration of ethical consequences helps protect the rights and well-being of individuals and communities affected by decisions or actions.

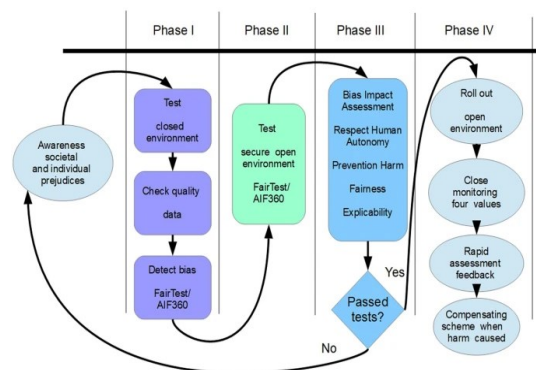


Fig 1. AI bias framework of action (Lorenzo Belenguer)

II. TYPES OF BIAS

A. Algorithm Bias: - The Ethical concerns in AI decision making. Today the most significant issue faced by AI today is Algorithm Bias and this refers to the potential of machine learning to discriminate in terms of race, gender, socioeconomic causes. Algorithmic bias possesses significant authority to favor one candidate over another such as during hiring processes. Similarly to this, law enforcement may use biased algorithms that may unfairly particular civilization or uphold systematic injustice.

B. Automation Bias: Entrusting decision making to Artificial Intelligence can shift human responsibility. Machine learning adds an additional layer of complexity between designers and actions driven by the algorithm.

C. Safety and Resilience: Unethical algorithms can be thought of as malfunctioning software and can be used inappropriately towards mankind.

D. Ethical Auditing: The possible path to achieve interpretability in decision making in new companies is solely dependent on this technology. Ethical policy part in decision making is only of the benefits in business. Many countries in the United States are adversely using this technology for their growth in business unlike hiring a senior manager for that matter. Ethical Auditing can clarify the actual values to which the company operates, provide a baseline which may help in future improvements, identify specific troubles and problems faced by companies and also identify general areas of vulnerability, specifically related to socializing on business grounds.

E. Omitted variable bias: Omitted variable bias is a common problem in statistical analysis, especially regression analysis, where the model does not include a significant variable that correlates with both the dependent variable and one or more independent variables. This omission can lead to biased and inconsistent estimates of model coefficients.

TABLE I

Sr.no	Analysis of bias in Ai algorithms		
	Study	Area	Findings
1.	Buolamwini and Gebru (2018)	Facial Recognition	Higher error rates in dark-skinned and female faces
2.	Mehrabi et al. (2022)	Various Applications	Multiple examples of biases across different domains
3.	Obermeyer (2019)	Healthcare	Racial bias in health management algorithms
4.	Williams et al. (2018)	Data Discrimination	Challenges of algorithms discrimination based on the data gaps

Table 1. Summary of key studies analyzing bias in AI algorithms.

Mechanism of Biases : We witness both the objective function and the inputs and outputs of the algorithm, which gives us a unique view into the mechanisms by which bias occurs. This makes our dataset unique. The algorithm in our situation receives a huge set of raw insurance claims data $X_{i,t-1}$ (features) covering the year $t - 1$ and includes information on diagnosis and procedure codes, drugs, insurance type, demographics (e.g., age, sex), and detailed charges. The algorithm, interestingly, explicitly leaves out race.

The algorithm predicts $Y_{i,t}$ (i.e., the label) based on these data. In this case, the label in the algorithm is the total medical expenses (we call these "costs" C_t for simplicity) for the year t . Therefore, the algorithm's forecast for health demands is actually a forecast for health expenses.[4]

Algorithm Bias:

The lack of impartiality in the outcome generated by Artificial Intelligence is called Algorithm Biasness. Practical Implications : When used by organizations, artificial intelligence (AI) can improve profitability, efficiency, and the range and quality of services offered. However, considering all of AI's advantages and disadvantages is necessary for its ethical application. In terms of how AI affects meaningful employment, we assist in articulating some of those costs and advantages for people. This is significant practically because, according to some authors (Acemoglu & Restrepo, 2020), there is a growing trend in which businesses use AI for complete automation while ignoring the opportunities it presents to improve human labor and inadequately preparing their workforces for the changes this entails (Halloran & Andrews, 2018). We draw attention to the paths that, for organizations, are likely to severely curtail opportunities for meaningful work, such as "minding the machine" work. This suggests that, in order to justify the use of AI, other factors, such as efficiency gains, must far outweigh the potential harm that this type of AI use can cause to workers. We also point out that, when evaluating meaningfulness, it is not enough to concentrate just on the AI itself because the consequences of its application are heavily influenced by the remaining tasks that require human labor, which is something that organizations have direct control over. We provide leaders with particular areas for intervention to enhance meaningful work experiences, and overall, we offer advice on how organizations can preserve or create possibilities for meaningful work when implementing AI. According to Grant (2007), task importance is essential for meaningful work; nevertheless, the methods in which AI can separate employees from recipients jeopardize these experiences. Organizations can address this, though, in a few different ways, like by telling employees about the success stories of their end consumers (Grant, 2008).[5]

There are two major reasons for the biases in Algorithms:

- 1) Personal biases
- 2) Environmental Biases

Factors contributing in Algorithmic bias:

Data Bias:

An AI system's choices may favor the group it was trained on if the data used to train it does not accurately reflect the population.

Discrimination in the design:

Implicit biases held by the AI designers may inadvertently manifest itself in the way the system operates.

Technological and social aspects:

The impact of social, economic, and cultural settings on AI system design, deployment, and usage is one of them, as it has the potential to create bias.

III. EXAMPLES OF BIASED ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING ALGORITHMS

Hiring Algorithms: Amazon once built an AI system to automate its hiring process. The algorithm was trained on resumes sent to the company over a ten-year period, which were overwhelmingly male. As a result, the system began to favor male candidates over women, showing a clear bias.

Facial Recognition Systems: Many studies have shown that facial recognition algorithms, such as those used to track or unlock smartphones, often perform poorly on dark-skinned and female faces. This is mainly due to the lack of diversity in education data.

IV. ETHICAL CONCERNS OF EACH BIAS

A. Algorithm bias: Algorithmic bias raises several ethical issues, particularly in the context of machine learning and artificial intelligence. The main ethical issues related to algorithmic bias are:

Fairness and equality: Algorithmic bias can lead to unfair treatment of certain groups, perpetuating inequality and discrimination. This is especially true in sensitive areas such as employment, lending and law enforcement, where biased algorithms can lead to unfair outcomes.

Transparency and Accountability: The lack of transparency of many machine learning algorithms makes it difficult to understand how they make decisions, leading to challenges holding them accountable for biased results. Lack of transparency can also make it difficult for affected individuals to challenge party decisions.

Privacy and Consent: Biased algorithms can exacerbate privacy concerns by unfairly targeting certain groups for surveillance or using personal data in ways that disadvantage certain populations. Additionally, individuals may not be aware of how their data is being used to train biased algorithms, raising questions about consent.

Social Impact: Algorithmic bias can have far-reaching effects on social trust in institutions and technologies. If biased algorithms are widely used, they can perpetuate social injustice and undermine public trust in the fairness of automated decision-making systems.

Reducing Bias: Addressing algorithmic bias requires significant effort and resources, and there are concerns about whether organizations and developers are doing enough to effectively reduce algorithmic bias. Without proactive efforts to address biases, the risk of harm remains high.

B. Automation bias: Ethical issues arising from automation bias are mostly related to the possibility of over-reliance on automated systems, leading to uncritical acceptance of these decisions. Some of the main ethical issues are:

Human responsibility and control: Automation bias can lead to a reduction in human control and responsibility, as individuals may unquestioningly follow the recommendations or decisions of automated systems, even if they conflict with ethical principles. . or moral considerations.

Responsibility and accountability: When automated systems make critical decisions, such as in healthcare, finance or autonomous vehicles, determining responsibility and liability for faulty or biased results becomes difficult. Lack of responsibility can have profound ethical consequences, especially in cases where harm has been caused.

Human Judgment: Over time, automation bias can impair human judgment and decision-making skills, which can lead to impairment of critical thinking and ethical reasoning. This can have wider social effects by reducing people's agency and moral judgment.

Discriminatory results: Automated systems can inherit and maintain biases in the data used to train them, leading to discriminatory results, as seen in algorithmic bias. This can exacerbate inequality and unfairly disadvantage certain individuals or groups.

Impact on human dignity: Over-reliance on automated systems can undermine the inherent dignity and autonomy of individuals, especially when it comes to critical decisions that deeply affect their lives. Ethical problems arise when automated processes overshadow human functionality and respect for individual choices.

C.Safety and Resilience: Ethical issues arising from safety and resilience are particularly important for technologies and systems that may affect public safety and well-being. Some of the key ethical issues in this domain are:

Security and Risk Mitigation: It is ethically imperative to prioritize security when designing and implementing technologies, infrastructures and systems. Failure to prioritize safety can cause both physical and psychological harm to individuals or communities. This is particularly important in areas such as autonomous vehicles, healthcare technologies and critical infrastructure.

Equity and accessibility: It is important to ensure that safeguards and flexible systems are available to all people, regardless of socioeconomic status, geographic location, or other factors. Failure to do so can increase inequality and leave certain groups more vulnerable to security risks and disruptions.

Transparency and informed consent: ethical considerations related to safety and sustainability include the need to publicly communicate risks and safety measures. People need to have access to clear and understandable information to make informed decisions about their safety. Informed consent is particularly important in areas involving medical treatment, experimental techniques and hazardous environments.

Impact on communities: security and resilience measures should take into account potential impacts on local communities, including environmental impact assessment, consideration of cultural heritage and ensuring that security measures do not disproportionately burden marginalized communities.

Resilience and Preparedness: Ethical issues also raise the resilience of systems and communities to potential disruptions such as natural disasters, cyber attacks or public health crises. There is an obligation to invest in construction resources that can protect vulnerable populations and mitigate the effects of disruptions.

D. Ethical Auditing: An ethical audit, also known as an ethical assessment or ethical evaluation, refers to the evaluation of an organization's operations, processes or technologies from an ethical perspective. Ethical issues can arise in the context of an ethics review and addressing them is important to ensure the effectiveness and integrity of the review process. Some of the ethical issues associated with an ethical audit include:

Independence and objectivity: Ethical auditors must maintain independence and objectivity when making assessments. Concern is when auditors are influenced or biased in their assessment, which can undermine the integrity of the audit process.

Transparency and Disclosure: Organizations must be ethically transparent about the ethics review process, including the criteria used in the evaluation, findings, and potential remedial actions. Lack of transparency can lead to mistrust and weaken the credibility of the review.

Consistency and Standards: An ethics review must be guided by consistent ethical standards and best practices. Inconsistencies in the evaluation of ethical aspects of different audits can raise concerns about fairness and reliability.

Stakeholder engagement: an ethics review should take into account the perspectives and concerns of various stakeholders, including employees, customers, affected communities and relevant stakeholders. Failure to communicate with stakeholders can lead to overlooking important ethical considerations and potential impacts.

Accountability and Remediation: When ethical problems are identified during an audit, concerns arise about whether organizations take sufficient responsibility to address those problems and implement meaningful corrective actions. Lack of responsibility can lead to ongoing ethical lapses.

E. Omitted variable bias: Due to the fact that any variables that remain and have a correlation with the deleted variable still carry information about it, omitted variable bias indicates that just removing a variable does not guarantee

that discrimination will not occur [6]. Similarly, because of correlations in the total data, algorithms may still discriminate based on sensitive personal information even if it is removed from the data [7]. According to a recent study, race-related data must be incorporated into algorithm modeling in order to guarantee that ADM systems do not discriminate, for example, based on race [8]. These results are supported by [9], who not only contend that it is frequently necessary to examine if algorithms discriminate, but also add that this discrimination can be avoided when sensitive information is handled appropriately.[10]

V. PRIVACY CONCERNS

Data Security:

Artificial Intelligence models require a vast amount of data and the protection of this data is necessary for blocking users with unauthorized access, breaches and cyber threats.

Privacy Protection:

Artificial Intelligence models often have the authority to analyze and interpret the personal credentials of the user and this is subject to privacy invasion.

Data Quality:

Inaccurate and biased data can lead to anonymous trouble and furthermore towards infringement.

Consent and Transparency:

Users should be fully informed about their utilization from the data fostering transparency and trust.

Accountability and Governance: Establishing a clear and safe accountability for the utilization of users data helps in user-data safety and also improves trust.

Data Retention:

Artificial Intelligence models should provide the user with the time span of users data possession in order to safeguard individuals data.

Algorithmic Bias:

The unfair use of algorithmic biases in the business and other models may lead to serious discrimination of race and gender and may offend the user creating a non user-friendly experience and major business loss.

Ethical use of data:

It is mandatory for the Artificial Intelligence models to adhere to the legal rights and other privacy policies to ensure the security of the users credentials and data.

Addressing these challenges requires a multifaceted approach involving robust technical, legal, and ethical frameworks to

ensure that AI-driven systems prioritize data privacy while delivering value and innovation.

Algorithm Bias Assessment : Bias in AI systems, especially in systems using ML, has been a significant cause of harm in recent years (Buolamwini and Gebru, 2018; Mehrabi et al.,2022). Identifying and mitigating biases in AI systems is often the first need that prompts companies to seek third-party assessments, and is usually the primary focus of technical reviews and algorithm reviews. In fact, many seem to think that biases are synonymous with "responsible AI". Because of the prominence of such issues, biases have been a central part of our risk assessment framework and we focus on them here. However, it should be noted that other technical judgments that go beyond algorithmic bias can be ethically relevant and important in some contexts (eg judgments about architectural transparency, explainability of results, and hackability).[11]

Privacy concerns : Significant user privacy concerns have been raised by the emergence of conversational text-based AI chatbots. User privacy concerns of conversational text-based AI chatbots remain largely unexplored, despite advancements in chatbot design. In order to comprehend this, 38 pertinent works in this topic were analyzed as part of a literature review using a grounded theory approach in this paper. The main focus of this review was to examine their theoretical and methodological frameworks. It also looked at issues that raise concerns about security and privacy when people communicate with chatbots. Investigating the potential effects that the gathering of user data over time may have on personal privacy and decision-making is a line of inquiry that merits more investigation. Studies have looked at the technical aspects of privacy concerns, but there is still a knowledge gap about how users see and understand these.[12]

VI. METHODS OF MITIGATING RISKS OF EACH BIAS

A. Algorithm bias: Mitigating the risk of algorithmic bias is critical to ensuring fair and just outcomes in automated decision-making systems. Here are some methods to reduce the risk of algorithmic bias:

Diverse and representative data: Use diverse and representative data sets in the training phase to reduce the risk of bias. Ensure that the data is balanced and reflects the diversity of the population to which the algorithm is applied.

Error detection and tracking: Implement bias detection and tracking mechanisms to continuously evaluate algorithm performance and detect potential anomalies. Regularly

evaluate results in different populations to identify and address biases.

Openness and explainability: Ensure transparency in the decision-making process of the algorithm. Make the logic, inputs and outputs of the algorithm understandable to users and stakeholders. This transparency can help identify and correct biased decisions.

Bias mitigation techniques: Apply bias mitigation techniques such as fairness constraints, bias correction algorithms, and post-processing techniques to adjust algorithms and reduce bias. These techniques help ensure that decisions are fair and impartial.

Diverse Development Teams: Build diverse and inclusive development teams to bring diverse perspectives and experiences to algorithm design and implementation. Diverse teams are more likely to identify and effectively address biases.

Regular audits and checks: Conduct regular audits and check algorithm performance to identify and correct biases. Independent audits can help ensure that the algorithm is working fairly and efficiently.

Ethical guidelines and training: Establish clear ethical guidelines for algorithm development and implementation. Provide training to developers and stakeholders on ethical considerations, bias detection and mitigation strategies.

User Feedback and Input: Collect feedback from users and stakeholders about the performance and results of the algorithm. Include user input to identify and correct biases that may affect different user groups.

Bias impact assessments: Conduct bias impact assessments to understand the potential impact of algorithmic decisions on different populations. Use these ratings to eliminate bias and ensure fairness.

Compliance: Ensure compliance with relevant policies and guidelines related to algorithmic fairness and bias. Keep up with emerging standards and best practices to drive the development process to ensure algorithmic fairness.

By adopting these methods, organizations can reduce the risk of misleading algorithms and promote fair and just outcomes in automated decision-making systems. Constant vigilance, openness, and cooperation are necessary to effectively address algorithmic biases.

B. Automation bias: Mitigating the risks associated with automation bias is important to ensure that human decision makers maintain critical thinking and control over automated systems. Here are some methods to reduce the risks associated with automation bias:

Education and Training: Educate users and decision makers about the capabilities and limitations of automated systems. Teach them to effectively integrate automated recommendations or decisions based on their own judgment.

Decision support tools: Develop decision support tools that encourage users to critically evaluate and review automated recommendations. These tools can invite users to consider alternative perspectives and potential biases in an automated system.

Human in the loop systems: Design systems that incorporate a "human-in-the-loop" approach where human decision makers actively participate in the decision-making process alongside automated systems. This arrangement allows people to monitor and intervene when necessary.

Feedback Mechanisms: Create feedback mechanisms that allow users to provide feedback on the performance of automated systems. This feedback can help identify biases or errors, leading to systematic improvements over time.

Algorithmic Transparency: Increases transparency of algorithms and decision-making processes in automated systems. Make sure users can see how decisions are made and what factors the system takes into account.

By implementing these methods, organizations can help reduce the risks associated with automation and promote a balanced approach to decision-making that uses both automated systems and human judgment. Continuous monitoring, transparency and user empowerment are key components of a successful strategy to combat automation bias.

C. Safety and Resilience: Mitigating security and resilience risks is essential to protecting individuals, communities and critical systems from potential harm or disruption. Here are some methods to reduce security and resilience risks:

Risk assessment and management: Conduct an in-depth risk analysis to identify potential security and resilience risks and vulnerabilities. Develop risk management strategies to effectively reduce and manage these risks.

Compliance with standards and regulations: Ensure compliance with relevant safety standards, building codes and regulations to reduce risks associated with structures, systems and operations. Following industry best practices can help prevent security incidents.

Emergency Planning and Response: Develop comprehensive emergency plans that outline procedures for dealing with security incidents, natural disasters or other disruptions. Do regular exercises and drills to ensure emergency preparedness.

Invest in resilience measures: Invest in resilience measures such as redundant systems, redundant power sources and structural reinforcements to increase the ability of critical infrastructure to withstand disruptions and recover quickly.

Security Education and Training: Conduct regular security education and training for employees, stakeholders and the community to increase awareness of security risks and foster a culture of security awareness. Empower individuals to take proactive steps to prevent accidents.

By implementing these methods and taking a holistic approach to security and resilience, organizations can effectively mitigate risk, improve emergency preparedness and protect the well-being of individuals and assets. Prioritizing safety and sustainability as essential elements of operations can promote long-term sustainability and success.

D. Ethical Auditing: Mitigating the risks associated with ethical auditing is central to ensuring the reliability and effectiveness of the review process. Here are some methods to reduce the risks of an ethical audit:

Independence and objectivity: Ensure that ethical auditors maintain independence and objectivity in their assessments. Measures are taken to avoid conflicts of interest and undue influence that may threaten the integrity of the audit process.

Transparency and Accountability: Promote transparency in the ethics review process by clearly stating audit objectives, scope and methods. Ensure that audit findings, recommendations and corrective actions are communicated to relevant stakeholders in a transparent manner.

Qualifications and Training: Ensure that ethical auditors have the knowledge, skills and expertise to conduct comprehensive audits. Provide continuous training to auditors to keep up with new ethical considerations and review best practices.

Stakeholder engagement: Engage with a wide range of stakeholders throughout the review process, including employees, customers, regulators and community members. Solicit feedback from stakeholders to understand their views and concerns about ethical issues in the organization.

Whistleblower Protection: Implement mechanisms to protect whistleblowers who report ethical violations or concerns during an audit. Foster a culture that values openness and encourages people to raise ethical issues without fear of retribution.

By implementing these methods, organizations can reduce the risks associated with ethical auditing and strengthen the integrity and effectiveness of their auditing processes. Prioritizing transparency, accountability, stakeholder engagement and adherence to ethical guidelines is essential to promoting ethical behavior and compliance within an organization.

E. Omitted variable bias: To mitigate the risks of omitted variable bias and get reliable results, it is important to reduce the risks of omitting variable bias in statistical analysis. Here are some methods to reduce the risk of omitted variable bias:

Conduct a thorough literature review: Before conducting an analysis, thoroughly review the existing literature to identify potential omitted variables that have been shown in previous studies to affect the dependent variable. . . Including relevant variables from the literature can help reduce the risk of leaving important factors out of the analysis.

Theoretical Understanding: Develop a strong theoretical understanding of the relationships between the variables under study. By understanding the theoretical basis of the relationships between variables, researchers can identify potential omitted variables and reduce the risk of bias.

Sensitivity analysis: Perform a sensitivity analysis by reestimating the model after including different sets of control variables. This helps to assess the reliability of the results when additional variables are included, reducing the risk of missing variables.

Use of instrumental variables: In situations where there may be unobserved variables that are correlated with included

regressors, consider using instrumental variables to address potential endogeneity and omitted variable bias. Instrumental variables can help control for omitted variable bias if the omitted variable is correlated with the included regressors.

Causal methods: If the goal is to establish cause-effect relationships, consider using causal inference methods such as propensity score matching, difference-in-differences, or regression discontinuity. These methods can help mitigate the risks associated with omitted variable bias by addressing potential confounding variables.

By adopting these methods, researchers can proactively reduce the risks associated with omitted variable bias and improve the validity and reliability of their statistical analyses. Prudent variable selection, careful model specification, and thorough sensitivity analysis are essential components to effectively address omitted variable bias.

TABLE II.

Sr.no	Key Privacy Concerns and Mitigation Measures		
	Privacy	Cause	Mitigation measures
1.	Data Breaches	Unauthorized access	Strong encryption, access controls
2.	Informed Consent	Lack of transparency	Clear communication, simplified consent forms
3.	Data Retention Policies	Indefinite data storage	Defining data retention periods, regular audits
4.	User Data Utilization	Data misuse	Transparent data policies, user control

Table 2. Summary of key privacy concerns associated with AI and possible mitigation measures.

VII. CONCLUSION

This research paper aims to elucidate on the ethical challenges put forth by AI and ML algorithms in Data Science and also offers insights into mitigating risks to ensure responsible and ethical use of these technologies.

Positively, the Ethical Implication related to artificial Intelligence and Machine Learning algorithms in Data Science include the following key findings:

- Algorithmic Bias
- Privacy Concerns
- Transparency
- Accountability and Governance
- Unemployment
- Ethical use of data
- Security and Unfair utilization of Data
- Data Quality

Ethical use of machine learning and artificial intelligence: digital phenotypic data
The four ethical pillars of medicine are autonomy (the right to choose), beneficence (doing good), non-maleficence (doing no harm) and justice (equal access), and these pillars should not be ignored in the democratization of digital phenotypic data. Mulvenna et al. [3] The operational phase of using digital phenotypic data raises important ethical issues involving responsibility, user protection, transparency and informed consent (Martinez-Martin et al., 2018). Intended use and informed consent involve autonomy because patients must know how the application and digital phenotype will be used before agreeing to TandC agreements. Clear and unambiguous language is critical, and the purpose of using TandC in digital phenotyping must be clear to ensure accurate informed consent (Dagum and Montag, 2019). Today, most users use TandCs carelessly and casually due to its complex and dense nature. This concerns that users have not given appropriate informed consent. In medical settings, it is necessary to clearly define to patients how their information is collected, stored and used in connection with medical care. The incorporation of digital phenotypes into a patient's EHR (Electronic Health Record) raises new concerns about possible unauthorized access to the EHR. [13]

The data science community, legislators, and industry stakeholders must work together to define and implement ethical principles, encourage openness, and give the rights and well-being of those impacted by AI and ML algorithms top priority in order to address these ethical concerns.

REFERENCE

1. R. Gupta, "Research Paper on Artificial Intelligence," Valley International, Feb. 18, 2023. https://www.researchgate.net/publication/371426909_Research_Paper_on_Artificial_Intelligence (accessed May 26, 2024).
2. C. Galkaduwa and N. Ranasinghe, "Data Science and Its Importance," unknown, Jan. 04, 2024. https://www.researchgate.net/publication/377169764_Data_Science_and_Its_Importance (accessed May 26, 2024).
3. I. H. Sarker, "Deep Cybersecurity: A Comprehensive Overview from Neural Network and Deep Learning Perspective," *SN Computer Science*, vol. 2, no. 3, Mar. 2021, doi: 10.1007/s42979-021-00535-6.
4. Z. Obermeyer, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, Oct. 2019.
5. S. Banks and P. Formosa, "The Ethical Implications of Artificial Intelligence (AI) For Meaningful Work," *Journal of Business Ethics*, vol. 185, no. 4, pp. 725–740, Feb. 2023, doi: 10.1007/s10551-023-05339-7.
6. K. A. Clarke, "The Phantom Menace: Omitted Variable Bias in Econometric Research," *Conflict Management and Peace Science*, vol. 22, pp. 341–352, 2005.
7. F. Bonchi, C. Castillo, and S. Hajian, "Algorithmic bias: from discrimination discovery to fairness-aware data mining," pp. 1–7, 2016.
8. I. Zliobaitis and B. Custers, "Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models," *Artificial Intelligence and Law*, vol. 24, no. 2, pp. 183–201, 2016.
9. B. A. Williams, C. F. Brooks, and Y. Shmargad, "How Algorithms Discriminate Based on Data they Lack: Challenges, Solutions, and Policy," *Source Journal of Information Policy Journal of Information Policy*, vol. 8, no. 8, pp. 78–115, 2018.
10. <https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/aef8b0cf-8c00-4e96-84fc-b96091a08ae5> (accessed May 27, 2024).
11. A. Hasan, S. Brown, J. Davidovic, B. Lange, and M. Regan, "Algorithmic Bias and Risk Assessments: Lessons from Practice," *Digital Society*, vol. 1, no. 2, pp. 1–20, Aug. 2022, doi: 10.1007/s44206-022-00017-z.
12. E. Gumusel, "A literature review of user privacy concerns in conversational chatbots: A social informatics approach: An Annual Review of Information Science and Technology (ARIST) paper," *Journal of the Association for Information Science and Technology*, May 2024, doi: 10.1002/asi.24898.
13. M. D. Mulvenna et al., "Ethical Issues in Democratizing Digital Phenotypes and Machine Learning in the Next Generation of Digital Health Technologies," *Philosophy & Technology*, vol. 34, no. 4, pp. 1945–1960, Mar. 2021, doi: 10.1007/s13347-021-00445-8
Cite relevant scholarly articles, reports, and ethical guidelines pertaining to AI ethics in Data Science.