

ETHICAL IMPERATIVES IN AI DEVELOPMENT: ALGORITHMIC FAIRNESS AND ITS ROLE IN PROMOTING SOCIAL JUSTICE

*Seema Sharma

**Dr Dev Ras Pandey

Paper Received: 12.11.2021 / **Paper Accepted:** 04.12.2021 / **Paper Published:** 07.12.2021

Corresponding Author: Seema Sharma; doi:10.46360/globus.met.320212020

Abstract

This paper explores the moral obligations posed by the advancement of artificial intelligence (AI), concentrating on the idea of algorithmic fairness and how important it is to the advancement of social justice. Talks about AI ethics have shifted to focus on issues of justice, accountability, and transparency as AI technology become more and more integrated into society. As a guiding concept, algorithmic fairness aims to reduce prejudice and discrimination in AI systems, guaranteeing fair results for every person, irrespective of their gender, colour, or socioeconomic background. The intellectual underpinnings of algorithmic fairness are examined in this work, along with its definitions, difficulties, and moral implications. It also looks into the real-world applications of algorithmic fairness in AI development, emphasising how important it is to include a range of viewpoints, reliable assessment techniques, and legal frameworks to prevent unintended consequences and advance equitable results.

Keywords: Artificial Intelligence, Ethical, IT, Technology, Social.

Introduction

It investigates the issues that arise when technology, ethics, and the impact on society are brought together within the context of artificial intelligence. This study investigates the vital relevance of developing algorithms that are fair and just, while also considering the ethical responsibilities that are inherent in the creation of artificial intelligence.

It is imperative that algorithmic biases, prejudice, and injustices be addressed as soon as possible because of the widespread involvement of artificial intelligence systems in many different parts of society. The purpose of this research is to provide light on the ways in which biases inside algorithms might contribute to the perpetuation of societal injustices, hence making existing gaps more pronounced across gender, ethnicity, socioeconomic position, and other categories.

By investigating various ethical frameworks and approaches, the purpose of this research is to plot a road that will lead to the development of artificial intelligence that is fair, accountable, and transparent. The purpose of this project is to provide insights into the process of building and deploying algorithms that prioritise social justice, protect against biased outcomes, and promote diversity.

This investigation highlights the ethical imperatives that are incumbent upon those who develop artificial intelligence (AI), those who conduct research on AI, those who set policy, and those who have a stake in the field. These imperatives will ensure that AI technologies not only foster innovation but also respect fundamental ideals of fairness and social equity.

Literature Review

Andersen (2019) [1] Artificial intelligence (AI) is poised to become the focal point of the international development sector in the near future. In the middle of this shift, the international development sector and the expanding area of AI and ethics do not adequately explore the distinct ethical challenges that AI efforts encounter in the context of development. This paper contends that the presence of multiple stakeholder layers in international development projects, along with the involvement of

*Research Scholar, Kalinga University, Naya Raipur, Chhattisgarh, India.

**Supervisor, Kalinga University, Naya Raipur, Chhattisgarh, India.

third-party AI vendors, gives rise to specific ethical concerns pertaining to fairness and inclusion, transparency, explainability and accountability, data limitations, as well as privacy and security. The text finishes by presenting a set of principles that expand upon the Principles for Digital Development of the information communication technology for development (ICT4D) community. These principles aim to provide guidance to international development funders and implementers in the responsible and ethical execution of artificial intelligence (AI) efforts.

Piano (2020) [2] The delegation of decision-making in various parts of our everyday lives is being transferred to machine-learning (ML) algorithms and artificial intelligence (AI), driven by the desire for quick and efficient decision-making. ML techniques, which are one of the classifications of algorithms that support artificial intelligence, are commonly designed as opaque systems. ML code scripts are often not thoroughly examined, prioritising usability and effectiveness over interpretability. Issues related to programme development have been identified in various aspects, such as fairness, correctness, accountability, and transparency. There is room for improvement in these practices. This contribution focuses on the creation of guidelines and specialised publications related to these topics. The subsequent instances of AI-powered decision-making are delineated: The text discusses two topics: risk assessment in the criminal justice system and autonomous vehicles. It specifically focuses on the conflicts that arise between ethical ideals in these areas. The examination of potential approaches for implementing AI governance is now being conducted.

Analysis

This study highlights how urgent it is to put into action measures that give fairness and social justice the highest priority in artificial intelligence. It places an emphasis on the necessity of interdisciplinary collaboration, which includes the participation of ethicists, social scientists, technologists, legislators, and communities that are affected in order to develop complete frameworks that reduce biases, guarantee fairness, and advance social equity.

Moreover, it is of the utmost importance to incorporate these ethical considerations into the entire process of developing artificial intelligence, beginning with the collection of data and the design of models and continuing with deployment and ongoing review. Through the use of this proactive strategy, trust is fostered, responsible innovation is encouraged, and technology breakthroughs are brought into alignment with the values and ambitions of society.

The pursuit of social justice is inextricably linked to the ethical imperatives that the development of AI must adhere to. The acceptance of these imperatives not only improves the quality and dependability of artificial intelligence systems, but it also makes a substantial contribution to the development of a future that is more inclusive and equal for everyone.

Compliance with Laws and Regulations: As artificial intelligence becomes more pervasive in everyday life, legal and regulatory frameworks are developing to control the ethical application of this technology. When organisations adhere to fairness principles, they are better able to comply with newly enacted legislation and eliminate the possibility of facing legal repercussions.

Creativity and Market Potential: The development of ethical artificial intelligence has the potential to stimulate creativity by promoting a variety of perspectives and working to lower entry barriers for communities that are underrepresented. Artificial intelligence that is both fair and inclusive has the potential to enter new industries and produce systems that are more robust and adaptive.

Long-term Sustainability: Addressing algorithmic biases helps to promote long-term sustainability by reducing social tensions, building a more egalitarian society, and ensuring that artificial intelligence systems serve everyone rather than perpetuate disparities.

When it comes down to it, giving algorithmic fairness and social justice the highest priority in the creation of artificial intelligence is not just a question of ethical obligation, but it is also a key necessity for the construction of AI systems that are strong, dependable, and socially beneficial, and that have a positive influence on both individuals and society itself.

There are a variety of sectors that are affected by the implications of ethical imperatives in the development of artificial intelligence, particularly with regard to algorithmic fairness and social justice. These implications have severe consequences:

- The impact that artificial intelligence systems have on society is shaped by ethical considerations. The judgements that are made about work, housing, finances, and criminal justice are impacted by concerns over fairness in algorithms, which in turn affects both individuals and communities.
- In order to promote trust in artificial intelligence systems, it is necessary to

address biases and ensure that they are fair. Ethical development procedures boost user confidence and acceptability, which in turn encourages wider adoption of artificial intelligence, or AI, technology.

- **Legal and Regulatory Frameworks:** The formulation of rules and regulations that regulate artificial intelligence is frequently influenced by ethical imperatives. Their influence extends to the formulation of laws that protect against discriminatory practices and provide accountability in the process of developing and deploying artificial intelligence systems.
- **Unchecked biases in artificial intelligence** have the potential to perpetuate inequality, which in turn can have an impact on the economic opportunities available to marginalised people. Ethical artificial intelligence work strives to reduce these prejudices, which will ultimately lead to more inclusive economic growth.
- **Ethical imperatives in the creation of artificial intelligence** are not limited by national borders, it is important to note. They stimulate conversations and partnerships on a worldwide scale in order to establish universal standards and guidelines for the responsible application of artificial intelligence.
- **Ethical issues should be incorporated into the innovation process** since doing so stimulates responsible innovation. The development of artificial intelligence can advance while simultaneously minimising the adverse effects on society if fairness and social justice are given priority.
- **Workforce and Skills Development:** The development of ethical artificial intelligence may have an impact on educational programmes, highlighting the importance of skills in data ethics, justice, and accountability. This will help to prepare the workforce for the responsible application of artificial intelligence technologies.
- **Confronting Ethical Challenges and Making Decisions:** As artificial intelligence (AI) technologies grow more widespread, ethical considerations encourage conversations about difficult conundrums. These conversations necessitate the development of ethical

decision-making frameworks in order to handle difficulties with prejudice, privacy, and interactions between humans and AI.

- **Effects on Society Over the Long Term:** Ethical imperatives have long-term repercussions on their respective societal structures. It is possible to contribute to a more equitable society by addressing biases and encouraging fairness in artificial intelligence. This will help reduce structural inequalities and foster inclusivity.

Not only does the understanding and implementation of ethical imperatives in the development of artificial intelligence shape the technology itself, but it also has far-reaching ramifications across various parts of society, determining how AI impacts individuals, communities, and the planet as a whole.

Conclusion

It is of the utmost importance to sustain ethical standards in the field of technology: ethical responsibility. It is of the utmost importance to make certain that artificial intelligence systems do not perpetuate or exacerbate pre-existing biases and inequities, as these systems have the potential to have far-reaching effects on both individuals and society. Impact on Society Biased algorithms have the potential to magnify and reinforce societal biases, which can result in discriminatory consequences in a variety of domains, including healthcare, employment, finance, and the criminal justice system. The way in which algorithmic fairness is addressed has immediate implications for societal equality and justice.

In order to establish trust in artificial intelligence systems, it is necessary to demonstrate a dedication to maintaining justice and transparency. Accountability among developers is improved by the implementation of fairness, which in turn reduces the risks associated with biased decision-making.

When it comes to the creation of artificial intelligence, the investigation of algorithmic fairness is not only a technological goal; rather, it is an ethical obligation that is profoundly connected with social justice. When it comes to the creation of artificial intelligence, addressing biases, maintaining justice, and preserving ethical norms are all essential steps towards establishing a society that is more egalitarian.

We recognise the tremendous influence that these technologies have on individuals and communities when we acknowledge the ethical imperatives that

are involved in the development of artificial intelligence. In the absence of proper design and monitoring, it is of the utmost importance to comprehend that algorithms are not impartial and have the potential to perpetuate societal biases. Therefore, it is of the utmost importance to incorporate the values of fairness, accountability, and openness into artificial intelligence systems.

Conflicts of Interest Disclosure

The authors declare that there are no significant competing financial, professional, or personal interests that might have influenced the performance or presentation of the work described in this manuscript.

References

1. Andersen, L. (2019). Artificial Intelligence in International Development: Avoiding Ethical Pitfalls. *Journal of Public & International Affairs*.
2. Lo Piano, S. (2020). Ethical principles in machine learning and artificial intelligence: Cases from the field and possible ways forward. *Humanities and Social Sciences Communications*, 7(1), 1-7.
3. Ntoutsis, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejd, W., Vidal, M.E., ... & Staab, S. (2020). Bias in data-driven artificial intelligence systems - An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1356, 2020.
4. Kumar V, Rajan B, Venkatesan R and Lecinski J (2019). Understanding the role of artificial intelligence in personalized engagement marketing. *California Management Review*, 61, 135-155. <https://bit.ly/3xh40Qa>
5. Dignum, V (2018). Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20, 1-3. <https://bit.ly/3rRUO3W>
6. Favaretto, M, De Clercq, E and Elger, BS (2019). Big Data and discrimination: perils, promises and solutions. A systematic review. *Journal of Big Data*, 6. <https://bit.ly/3BZcmzw>
7. Rai, A (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, 48, 137-141. <https://bit.ly/3xh3x0m>
8. Xu, F, Uszkoreit, H, Du, Y, Fan, W, Zhao, D, et al. (2019). Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11839 LNAI, 563-574. <https://bit.ly/3lh6uM3>