



# Artificial Intelligence and Question of Ethics in the Digital Age

**Gurdeep Singh**

Research Scholar, Department of Philosophy  
Punjabi University Patiala  
Gurdeep9206@gmail.com

## How to Cite this Article:

Singh, G. (2026). Artificial Intelligence and Question of Ethics in the Digital Age. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(03).  
<https://doi.org/10.55041/ijcope.v2i3.171>

## License:

This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.



<https://doi.org/10.55041/ijcope.v2i3.171>

## Abstract—

Artificial Intelligence (AI) is a disruptive technology that is defining the decision-making process in social, economic, and institutional settings. Whereas, AI is efficient, innovative, and predictive, its fast growth has spawned serious ethical issues. This paper gives a theoretical analysis of the most important ethical aspects of AI, such as algorithmic discrimination and bias, transparency and Explainability, the risk of privacy and data protection, accountability, the disruption of employment, manipulation of behavior, and the ethical consequences of surveillance technologies.

Based on the available ethical viewpoints, the paper will highlight the necessity of incorporating the principles of fairness, transparency, accountability, human control, and respect of fundamental rights into AI creation and regulation. It concludes that an anthropocentric and ethically sound approach is needed to make sure that AI will improve the welfare of the society with a minimal number of risks and unforeseen consequences. **Keywords—** Artificial Intelligence, AI ethics, Algorithmic bias, data privacy, Accountability



## I. INTRODUCTION

The 21st century has seen tremendous development in one significant area: AI (Artificial Intelligence). AI, in terms of digital computing technologies, enables computers to learn and understand (among other things) human cognition and think about opportunities and threats (or both), understood through language, probability, logic, and pattern recognition. With this capability, AI is now part of our day-to-day life, and as we have seen recently, the extensive use of virtual assistants, the creation of personalised recommendation systems, the use of AI in both diagnostic medicines, the operation of unmanned (or self-managed) vehicles, and in the decision-making abilities of a modern-day CEO are just some examples. While we can imagine the many benefits of AI (improved efficiency, enhanced innovation, and growth of the economy), there are also many ethical issues that arise from the rapid growth and adoption of AI technologies. Among these ethical concerns is the moral obligation of the creators of AI to properly manage the ethical implications of their creations (responsibilities when designing and deploying their intelligence). Unlike traditional technology, AI systems can learn from data and make major decisions about a large population of humans.

The use of data to train AI systems and develop advantages and disadvantages of this information leads to an increased potential of harm on a societal scale, which needs to be addressed through a comprehensive understanding of the ethical implications of AI.

In this paper, we explore the broad spectrum of ethical issues that need to be addressed regarding AI, such as bias and discrimination, transparency, privacy, accountability, effects on employment, manipulation, and surveillance; and moral decision-making by machines. Bias and fairness within AI are among the most discussed ethical issues when it comes to artificial intelligence (AI). AI makes predictive decisions and recommendations by learning from prior experience through analysing large amounts of information (data).

When used in this manner, an AI can learn the patterns associated with prejudice and inequity that have historically existed within society. Therefore, an AI could reproduce and even exaggerate any type of biases that are reflected in the data it uses to learn, whether those patterns are unintentional or intentional.

## II. LITERATURE REVIEW

The last ten years have witnessed much academic interest in the ethical ramifications of Artificial Intelligence (AI), especially with the integration of AI systems into the decision-making process of high stakes. The literature has a wide view of issues concerning bias, transparency, privacy, accountability, governance, and the long-term societal impact.

Algorithmic bias and discrimination are one of the issues about AI ethics that has been discussed the most. To support the argument, Barocas and Selbst (2016) claim that data-driven algorithms will have unequal effects in the cases when the overt discrimination is not evident. They note that historical data may be biased, which in turn can be used to produce unequal results among groups of people. Equally, O'Neil (2016) emphasizes the discriminatory nature of predictive algorithms used in hiring, lending, or criminal justice and calls those types of systems weapons of math destruction. The works highlight the importance of machine learning that is fair and constant audits of algorithms. The theme of transparency and Explainability has also come out as a key theme in the literature. According to Burrell (2016), there are three types of machine learning system opacities, such as intentional secrecy, technical illiteracy, and inherent complexity. As per researchers, the lack of accountability and trust in the deep learning models is a weakness of the models due to their black box nature. Floridi et al. (2018) support the idea of explainable AI as a solution that can make sure that automated decisions can be interpreted and justified in accordance with ethics. Another area of dominance in discussion is privacy. The article by Zuboff (2019) presents a notion of surveillance capitalism, stating that AI-based data mining allows engaging in novel methods of behavioural surveillance and manipulation. Researchers highlight that mass surveillance poses a risk to individual privacy and it



must be ensured that there is a strong regulatory framework to ensure the privacy of the information.

Challenges to accountability and governance are also common issues which are scrutinized. According to Mittelstadt et al. (2016), delegating accountability toward algorithmic harm is complicated because of the decentralized nature of AI development and implementation. Lack of clear liability systems may give rise to ethics and legal ambiguity. Therefore, scholars recommend the concept of human regulation and systematic governance patterns.

The effect of AI on jobs has created a controversy in economic and ethical literature. According to Frey and Osborne (2017), a large share of the jobs can be automated, and this raises the issue of inequality and social disruption. Researchers state that the ethical implementation of AI should be supported by a strategy to adapt the workforce and inclusive policies. Also, the role of AI in autonomy and democratic procedures is discussed in literature. Algorithms in content curation researches indicate that personalization systems can form echo chambers and reinforce fake news. It has resulted in demands of ethically designed standards that would put the welfare of society first and engagement measures second. On the policy side, AI development has been suggested to adhere to ethical principles developed by international organizations, focusing on fairness, accountability, transparency, and human rights. Nevertheless, researchers mention that it is also a significant difficulty to convert ethical codes into enforceable rules.

All in all, it can be stated that the literature proves the increasing belief that AI technologies cannot be considered ethically neutral and that they should be proactively controlled. Although a lot of theoretical advancement has taken place, a gap still exists between theoretical and practical application of ethics, which requires continued research.

### III. RESULTS AND DISCUSSION

#### 1. Algorithm Bias and Fairness

The bias that exists in AI can be due to any of the following: unrepresentative datasets, inaccurate data collection methods, or assumptions made by AI developers. An example of this could be classic AI - when an AI model is trained to predict who would be hired for a job based on hiring data from companies.

If some of the demographics had less representation in that dataset, the AI will learn to predict and recommend people from that demographic group more frequently than those from other demographic groups.

To implement ethical AI for companies and institutions, AI will need to undergo continual review. To ensure that we reduce the pervasiveness of discrimination, we need priorities to improve data quality, diversity, and regular audits of created algorithms. Ethical AI defines Fairness in many different ways. There is much debate as to how Fairness should be defined because countries have different standards in areas of equality and merit.

#### 2. Transparency and Explainability

Transparency of AI models means being able to find out how the models function and work through their algorithms and decision-making. The majority of AI models used today are very complex systems that have limited interpretability (usually based on deep learning). As a result of the lack of ease of interpretation, these models create a "black box" problem—while the outcomes generated by these systems are visible, the methods used to arrive at that outcome are hidden or non-existent. The absence of an adequate explanation of how a system makes its predictions creates a growing amount of ethical concern associated with the use of AI in industries where the decision-making process may have serious consequences. For example, if a loan is denied, if a person is predicted to be a criminal, or if someone is recommended a drug, the individual affected cannot challenge or contest the result of the decision. This situation creates fairness and accountability concerns.

One type of AI system is called "Explanatory" or "Explainable AI" (XAI), which tries to create models that can provide information that is clear and understandable about how the system operates and what it is doing. Ethical AI models are designed to be easily interpretable by the users and/or regulators and/or third-party stakeholders, etc. Transparency ensures users can trust AI systems, track mistakes made, and give informed consent.



### 3. Privacy and data protection

AI uses huge amounts of data for machine learning training and development, and much of this data is Personal and Sensitive. Examples are Biometric data, Browser history and behaviour, Health records, Voice recordings, and Location data. The collection and use of this type of data creates significant ethical challenges around Privacy, Consent, and Security.

AI generated profiles will create detailed, digital versions of their Digital Identity that can be used for Advertising, Political Messaging, and accessing Services. In addition, people might not know who is using their information and how; examples include sharing data without permission, Hacking into systems, and Misuse of data.

Ethical Artificial Intelligence should address these issues by providing Transparency with regard to how data is collected and used, require Informed Consent, have a minimal data collection policy, and provide Strong Cybersecurity protections. Individuals have a right of Access, Correction and Deletion of their Own Data and to have their Privacy Respected to Maintain Autonomy and Human Dignity in a Digital Society.

### 4. Accountability and Responsibility

Attributing culpability for the effects of artificial intelligence has proven to be one of the most challenging issues around ethical questions. The actions of an AI may implicate several different actors when something goes wrong: developers, users, data suppliers/owners, companies and corporations who may use the algorithm, etc. It can be difficult for courts to determine who should take culpability or face justice because of these multiple potential parties responsible for an AI's actions.

An example of this would be a self-driving vehicle involved in an accident. A court must consider numerous variables when determining what went wrong — software error, problematic sensors, driver error, or inadequate regulatory oversight of the vehicle. The same is true for an ethical dilemma regarding an AI making biased decisions: how is one able to determine where culpability should fall?

Ethics from many different perspectives emphasize that AI should never function without the oversight of humans and that humans must ultimately be responsible for the outcomes of AI.

Therefore, there needs to be appropriate governance structures to ensure that the lines of responsibility are clear and designated. More recently, legal systems have begun developing ways to establish legal theories for liability for harm caused by AI.

### 5. Impact on Employment and Economic Structures

The growing use of automation powered by Artificial Intelligence affects labor markets throughout much of the world. There are currently many types of repetitive or routine jobs that are often done by machines, which decreases the amount of human labor utilized in those industries. While automation can increase an industry's productivity, it presents challenges to job security for thousands of workers. The rapid change in technology has also raised many ethical issues concerning economic inequality and social justice. Specifically, those with lower skills and who are more vulnerable will probably feel the impact of being displaced from their jobs; in contrast, those who possess the skills required to perform higher levels of responsibility will benefit from the creation of new opportunities as AI technology increases.

For society to progress ethically through the advancement of technology, there must be a balance of efficiency with being socially responsible by providing the necessary resources for workers to make the transition to technological advancements. Governments and institutions must invest in vocational training and education and develop economic policies that include all citizens so that every worker has the ability to adapt to new technologies. Ultimately, AI should not limit the potential of people, but should augment their capacity to create value.

### 6. Manipulation, Persuasion and Autonomy

AI systems increasingly affect how people think and act by providing highly personalized recommendations, targeted ads and curated content. By assessing individual preferences (e.g., what users would like or enjoy most), the technology develops extremely individual insights to help people form opinions, feelings and make decisions about products/services they may eventually purchase.



While personalization makes it easier for people to use a service or product, there are also ethical implications of being manipulated by personalized algorithms. In most cases, algorithms use data to focus on materials that generate the highest amount of engagement (e.g., get people to react), including sensationalism, fake or distorted information and content with an emotional component. This ultimately contributes to the proliferation of misinformation, increased political divisiveness and skewed views of reality.

In order to manage the ethical challenges of algorithmic influence, it is important that individuals be empowered with the ability to control their personal/personalized choices. Transparency of how the technology recommends content and tools to manage the personalization process are essential to keeping individuals independent and self-reliant.

### **7. Surveillance and Social Control**

Use of AI for surveillance technologies (such as facial recognition systems, behavioral analysis systems, and predictive monitoring systems) has grown significantly since their introduction and now is being utilized for governmental (security), corporate (marketing), and institutional (management) use by many different organizations. While surveillance technology enhances safety and efficiency, it poses considerable threats to civil liberties by intimidating individuals into silence regarding their views/thoughts and creating an environment that allows for the empowered discrimination of individuals. Errors generated through surveillance can also result in wrongful identification and create feelings of violation within the incorrectly identified individuals. Ethical utilization of surveillance technology requires the establishment of a regulatory framework that provides strict oversight, demonstrates proportionality between the means utilized that provides for the protection of individual rights, and establishes the need for transparency and accountability among those employees responsible for monitoring the activities of individuals. Society must weigh the need for security against the desire for personal freedom through a careful evaluation of all associated factors.

### **8. Moral decision making by Machines**

For some uses of AI include areas like autonomous driving, which makes choices between differing levels of harm based on different factors, or medical systems deciding how to prioritize patients for treatment based on ethical trade-offs. These situations raise philosophical questions about the way that moral values can be programmed into machines. There are very different approaches to how humans make moral decisions (e.g., utilitarianism, deontological ethics, virtue ethics, etc.) that depend on either an individual or a particular culture. Thus, producing executable code representing these intricate human morals using an algorithm is extremely difficult, not to mention the chance of dramatically simplifying them.

In addition, cultural differences create additional complications when it comes to determining priority critically due to differing moral principles between societies. It follows that in the area of moral decision-making, the use of AI should serve to support human judgment rather than to substitute for it in situations that require moral sensitivity.

### **9. Ethical Framework and Global governance**

Ethical Guidelines and Principles to Address the Ethical Issues Associated with AI. Various organisations & governments have developed Ethical Guidelines & Principles for AI, with commonalities among Many of these Guidelines have Promised Themes of Respect for Human Rights, Fairness, Transparency, Accountability, and Providing for the Social Benefit.

International Collaboration has Been Developed for the First Time Since the Global Nature of AI Technology has Emerged. Regulatory Frameworks Aim to create Responsible Innovation whilst preventing Future Use / Misuse of AI Technology. Ethical Review Boards, Impact Assessments and Algorithm Auditing are some of the New Tools for Governance Over AI Technology. However, Translating the Ethical Principles into Practice still presents a Significant Challenge to Implementing AI Technology. If Ethical Guidelines do not include a mechanism for Enforcing Them, or if No Commitment Exists to Implement Ethical Guidelines, they will only serve to Create a Symbolic Representation of a Professional Organization or Government's Ethical Obligations.



## 10. Long term Societal Implications

The introduction of A.I. will raise questions about our long-term and future ethical responsibility for other humans, machine reliance, and power distributions in society. As A.I. systems continue to evolve and become more autonomous, we as humans will rely on A.I. to make decisions for us. Relying too much on A.I. systems to make decisions has the potential to weaken our ability to think critically, decrease our ability to act, and consolidate power into a few people who design and manufacture the A.I. Therefore, ethical evaluation and discourse must include consideration of "current" risks, as well as future social changes. As all A.I. systems there must be constant communication between those developing these technologies (technologists), those setting policy (policymakers), ethicists who evaluate the technologies' effect on society, and we as the public who will be affected by these technologies.

## IV. CONCLUSION

The advent of artificial intelligence has brought with it many exciting opportunities. However, AI also presents a number of ethical dilemmas that affect our society, economy, and how we govern ourselves. This paper has focused on several of the major ethical issues surrounding the development and use of AI, such as algorithmic bias and discrimination, opacity of the decision-making processes of AI, privacy, data security, ambiguity surrounding accountability, the effect on employment through automation, manipulation of behaviour by AI, and the expanding use of surveillance. As a result, we see that AI systems are embedded in a multitude of societal contexts and provide both benefits and harm. Bias continues to exist in algorithmic models, revealing the need to develop fair practices in collecting and sharing data as well as create processes for ongoing surveillance for the purpose of avoiding technological reproduction of current inequities. Moreover, with their complexity and inability to provide transparency regarding the processes behind automated decision-making, people may not feel comfortable trusting a complex AI system or holding it accountable for the outcomes of that process because they lack visibility into how the decisions were made (or if they were even made intelligently). This results in an increased reliance on ethics, and the need for laws to protect privacy as well as the potential for mass collection of personal information through artificial intelligence

technologies. Finally, our current governance frameworks do not clearly define who would be considered liable for damages when it's due to an artificial intelligence system, which shows us just how inadequate our current governance frameworks are in regard to providing an effective system to facilitate accountability of artificial intelligence systems.

## REFERENCES

- [1] Afreen, J. (2025). Systematic literature review on bias mitigation in generative AI. *AI and Ethics*. <https://doi.org/10.1007/s43681-025-00721-9>
- [2] Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671–732. <https://doi.org/10.15779/Z38BG31>
- [3] Batool, A. (2025). AI governance: A systematic literature review. *AI and Ethics*. <https://doi.org/10.1007/s43681-024-00653-w>
- [4] Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12. <https://doi.org/10.1177/2053951715622512>
- [5] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles,
- [6] Hanna, M. G. (2025). Ethical and bias considerations in artificial intelligence and machine learning. *Journal of Big Data & Ethics Review*. [https://doi.org/10.1016/S0893-3952\(24\)00267](https://doi.org/10.1016/S0893-3952(24)00267)
- [7] Joseph, J. (2025). Algorithmic bias in public health AI: Challenges to equity. *Public Health AI Journal*. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC12325396/>
- [8] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1–21. <https://doi.org/10.1177/2053951716679679>
- [9] Nouis, S. C. E. (2025). Evaluating accountability, transparency, and bias in AI: Evidence from clinical settings. *BMC Medical Ethics*. <https://doi.org/10.1186/s12910-025-01243-z>
- [10] O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.



- [11] Radanliev, P., & Roure, D. (2025). Privacy, ethics, transparency, and accountability in AI systems. *Frontiers in Digital Health*. <https://doi.org/10.3389/fdgth.2025.1431246>
- [12] Saeidnia, H. R. (2024). Ethical considerations in artificial intelligence interventions. *Sociological Perspectives on Emerging Tech*, 13(7), 381. <https://doi.org/10.3390/soc13070381>
- [13] Sargiotis, D. (2024). Ethical AI in information technology: Navigating bias, privacy, transparency, and accountability. *Advances in Machine Learning & Artificial Intelligence*, 5(3), 1–14.
- [14] Singhal, A., & Jamison, K. (2024). Toward fairness, accountability, transparency, and ethics in AI applications on social platforms. *Journal of Medical Internet Research*. <https://doi.org/10.2196/50048>
- [15] Yan, Y. (2025). A systematic review of AI ethics in education: Trends and challenges. *Education Technologies & Ethics Review*. <https://doi.org/10.4018/978-1-7998-8478->
- [16] Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.