

Original Article

Artificial Intelligence and Human Values: Balancing Innovation with Ethics

Shubham Yadaorao Gaoli¹ & Akhilesh Ramesh Bramhe²

¹Department of Mathematics,
 Sardar Vallabhbhai National Institute of Technology, Surat
²P. G. Department of Education,
 Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur

Manuscript ID:
 yrj-150202

ISSN: 2277-7911
 Impact Factor – 7.025

Volume 15
 Issue 2
 Apr-May-June- 2026
 Pp. 11 - 18

Submitted: 06 Apr. 2026
 Revised: 22 Apr. 2026
 Accepted: 30 Apr. 2026
 Published: 01 May. 2026

Corresponding Author:
Shubham Yadaorao Gaoli

Quick Response Code:



Web. <https://yra.ijaar.co.in/>



DOI:
 10.5281/zenodo.22258809

DOI Link:
<https://doi.org/10.5281/zenodo.22258809>



Creative Commons



Abstract:

Artificial Intelligence (AI) is one of the most game-changing technological advancements of the 21st century. It has found itself an early adoption in education, healthcare, business, governance, communication and scientific research, bringing a lot of opportunities for human development. Meanwhile, the growing adoption of AI raises serious ethical issues of privacy, human dignity, equality, autonomy, accountability, employment, misinformation, and social justice. While AI systems can handle complex tasks, process vast amounts of data, and assist in decision-making, they lack human consciousness, empathy, and moral responsibility as humans do. Hence, it is essential that technological development is paired with ethic and solid foundation of human values. This paper discusses the connection between the values of humans and the Artificial Intelligence, and also the balance of innovation with responsibility. It addresses the advantages of AI, significant ethical concerns, and the necessity of transparency, fairness, accountability, privacy, human oversight and inclusion. The paper also calls attention to the role of education in fostering responsible AI users and offers recommendations to policymakers, educational institutions, technology developers and society. The paper contends that AI technology should be designed as a human-centred technology, enhancing human capacities and safeguarding dignity, freedom, equality and social well-being.

Keywords: *Artificial Intelligence, Human Values, Ethics, Technological Innovation, Human Dignity, Privacy, Responsible AI, Social Justice.*

Creative Commons (CC BY-NC-SA 4.0)

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits others to remix, adapt, and build upon the work non-commercially, provided that appropriate credit is given and that any new creations are licensed under identical terms.

How to cite this article:

Shubham Yadaorao Gaoli & Akhilesh Ramesh Bramhe (2026). Artificial Intelligence and Human Values: Balancing Innovation with Ethics. Young Researcher, 15(2), 11 – 18.
<https://doi.org/10.5281/zenodo.22258809>

Introduction:

Artificial Intelligence is the name given to systems or machines that are able to perform tasks which are usually done by humans and that involve reasoning, learning, problem solving, language processing, prediction and decision making. The past decade has seen AI come into a rapid evolution due to improved computers, data, machine learning, deep learning, and generative technology. AI is making an impact in nearly all the sectors of society.

AI can be used in education to enhance personalized learning, automate assessment, translate languages, create educational content, and provide academic support. AI can be used in healthcare for medical diagnosis, medical imaging, drug discovery, and patient management. It can be used for crop monitoring, weather forecasting and resource management in agriculture. AI is applied in business for customer support, data analysis, marketing and automation. AI's impact on governance and service provision is also being investigated by governments.

While all these are excellent, AI isn't just a technological problem. It's also a social, philosophical and ethical question. The application of intelligent technologies in human activities have implications about who is in control of automated decisions, who is responsible for them, how personal data should be safeguarded and whether technological systems can generate social inequalities.

Human values serve as the moral compass that guides the people and societies to decide what is good, just, responsible, and meaningful. Honesty, Justice, Compassion, Equality, Freedom, Dignity, Responsibility and Respect towards others are vital values for a healthy society. AI systems, on the other hand, are largely algorithms and data and lack human moral consciousness. As a result, the moral appropriateness of an action cannot be decided by AI alone, since it is technically feasible.

The question then is not one of whether or not to create AI but how to create and use AI without violating core human values. Technological innovation and ethics should not be regarded as opposing forces. Ethical principles offer the basis of trustworthy, inclusive and sustainable technological development.

Objectives:

1. The study is designed to analyze the connection between AI and human values, discuss the ethical issues, explore the beneficial aspects of AI, and propose strategies for responsible and ethical implementation of AI.

Artificial Intelligence and Human Values:

Human values affect people's behavior, social interaction, institutional processes and public decisions. AI, on the other hand, is developed to fine-tune specific purposes based on information and computational methods. Moral issues which arise with human moral judgement and with

machine decision making become more significant in relation to human rights and welfare.

For instance, a recruitment system powered by AI can analyze employment data from previous candidates to assess their suitability. If there are gender or social biases in the historical data, then it will likely replicate those biases. Likewise, if an algorithm is used in an educational context such as an assessment tool, it can inadvertently disadvantage students whose language, learning style or socioeconomic background are not the majority.

This illustrates that AI is not inherently good or bad. Algorithms are designed by humans, trained from humans' data and deployed in social institutions. Therefore, the values and assumptions of developers and organizations can have a bearing on technological outcomes.

Floridi et al. (2018) stated that there is a need for guiding AI in promoting human flourishing and societal well-being. Likewise, UNESCO (2021) highlighted the principles of human rights, dignity, fairness, transparency and accountability as key principles for ethical AI.

Positive Contributions of AI to Human Development:

With responsible use, AI can significantly contribute to human welfare.

1. AI in Education:

AI can help with personalized learning, by identifying the learning needs of students and providing them with resources

that are tailored to their needs. It can also assist teachers in administrative things, assessment, translation, and content development. Artificial Intelligence (AI) can enhance the accessibility of the education for individuals with disabilities.

But AI should be used to assist teachers, not replace the teacher/learner connection. Ethical education, motivational, moral and emotional development, socialization and empathy are areas where human educators are still needed in education.

2. AI in Healthcare:

AI can help healthcare professionals detect trends in medical imaging, forecast health risks, and guide clinical decisions. Can support research and drug development, and can help enhance even delivery of health care in areas with limited resources.

However, personal dignity and information must be taken into account when making health care choices. In this context, AI in healthcare must, therefore, be protected, supervised, transparent and accountable.

3. AI in Economic Development:

AI can help in boosting productivity by automating repetitive tasks and assisting in data-driven decision making. New jobs could be created in the development of AI, data science, cybersecurity and digital services.

Meanwhile, some jobs currently held by humans could be lost. The challenge is to prevent the technological development from inducing harsh economic disparities.

Reskilling and lifelong learning are thus two important pillars of responsible AI adoption.

Ethical Challenges of Artificial Intelligence:

1. Privacy and Data Protection:

AI systems need tons of data. Examples of personal information might be educational information, financial information, photos, online activity or other confidential data. Having such data collected or used out of control can compromise individuals privacy.

In this regard, privacy is a key human value that should be considered when developing AI. Organizations should only gather data that is needed, implement proper security measures, and make it clear to people what they are doing with their data.

2. Algorithmic Bias and Discrimination:

Biases in training data can be perpetuated or exacerbated by AI systems. Discrimination can happen at the time of recruitment, loan applications, education, health care service and elsewhere.

To ensure fairness, it is crucial to systematically test AI systems, use diverse datasets, monitor them on an ongoing basis, and establish processes for detecting and mitigating discriminatory results.

3. Accountability:

One of the major ethical dilemmas is accountability for AI-driven decision-making. It is often hard to establish who is at fault when an automated system fails to deliver

the expected or desired outcome or it causes a negative consequence.

The accountability of human beings and organizations in the design and utilization of systems has to be maintained. The key is to not shift the burden of responsibility from human to algorithm.

4. Human Autonomy:

Over-reliance on AI can lead to a decrease in human decision-making and critical thinking. When people blindly follow the algorithm, they might become dependent on it and stop making their own judgments.

AI should therefore be viewed as a tool for making a decision, and not the deciding factor in a decision-particularly one with a direct impact on an individual.

5. Misinformation and Deepfakes:

Generative AI can create realistic text, photos, audio and videos. They can be a great educational and creative tool, but they can also be utilized for creating misinformation, manipulated media, and deceptive content.

Media literacy and information validation, along with responsible technology use, are critical.

Innovation and Ethics - A Balance Act:

Creating risks is no reason to limit the development of AI. Rather, technological innovations should be created with ethical considerations. The following is part of a balanced approach:

- **AI as a tool for human development:**
AI should enhance human well-being and address human needs, not be a data point or a productive asset.

- **Transparency:** People should be given useful information on important decisions that are made with the help of AI, and be able to understand the factors that went into the decisions, if possible.
- **Fairness:** Assesses for Bias and Designed to be Inclusive for all.
- **Transparency:** Responsibilities for AI outcomes should be clearly defined among developers, institutions, organizations, and policymakers.
- **Meaningful human oversight:** Key decisions around health, education, work, justice, and human rights should include human oversight.
- **Respecting individual privacy and legal protections:** Data collection, processing, storage and sharing should be carried out in a way that respects individual privacy and relevant legal protections.

The OECD (2019) provided a set of principles that are important factors for building trustworthy AI, such as principles of inclusive growth, human-centred values, transparency, robustness, security and accountability. These principles show that good moral behaviour and creativity are not mutually exclusive.

Role of Education in Promoting Ethical AI:

Education is key to equipping people for the AI-powered world. Students must have more than mere technical knowledge. They need critical thinking, digital literacy, ethical thinking, creativity, communication skills and social responsibility. Students

should be educated on understanding AI systems, assessing information created by AI systems, and the impact of algorithmic bias and privacy concerns on society. Professional development should also be provided to the educators to enable them to incorporate AI in an ethical manner in the teaching-learning activities. This is pertinent to the principles of holistic development and critical thinking which are reflected in the National Education Policy 2020 of India. It is important to not only teach programming and technical skills in AI education, but to also delve into the social, ethical, cultural and philosophical implications of technology.

Indian Context:

AI has a lot of potential applications for social and economic development in India. AI can assist in the fields of agriculture, healthcare, education, financial inclusion, governance, transportation, and environmental management. India's linguistic and cultural diversity also presents opportunities for the creation of AI tools that cater to the local needs. But, the challenges of unequal access to digital infrastructure, digital literacy, and data privacy concerns, and socioeconomic inequalities need to be addressed. AI should not exacerbate, but mitigate the digital divide. There is a need to have ethical applications of AI that align with democratic principles, human dignity, equality, justice, inclusion and individual rights in India. The role of policy makers and institutions in using technological

advancement to the benefit of the society as a whole is important.

Recommendations:

The following recommendations are made based upon the discussion above:

- **Create human-centred AI policies:** AI policies should be based on the value of human dignity, safety, equality and well-being in the development of technology.
- **Improve AI Ethics instruction:** AI ethics instruction in schools, college, and universities should be enhanced.
- **Be transparent:** Clear information on the use of the AI, especially when an automated decision has a high impact on the people involved should be provided.
- **Institute accountability systems:** Developers, institutions, and users need to be held accountable for harmful and discriminatory AI outcomes.
- **Respect privacy:** robust data-governance measures must be taken, such as data minimisation, security and informed consent, and responsible data handling.
- **Take proactive steps to mitigate algorithmic bias:** AI systems should be subject to periodic testing and also be subject to independent analysis to detect discriminatory outcomes.
- **Keep human judgment** at the centre of high-stakes decisions such as health, education, employment, law and public welfare.
- **Advance digital inclusion:** Governments and education systems

should ensure that there is digital infrastructure and learning opportunities in the field of AI, for disadvantaged communities.

- **Provide opportunities for reskilling and lifelong learning:** those who are impacted by automation should be offered a chance to acquire new skills and be available in the new industries.
- **Encourage interdisciplinary research:** The development of AI should not be limited to the involvement of computer scientists, but should include educators, social scientists, philosophers, legal experts, policymakers and other stakeholders.

Conclusion:

AI can have a massive impact on human society. It has the potential to enhance the educational, health care, agricultural, scientific research, communication, governance and economic productivity sectors. But the power of technology does not make an automatic guarantee that AI can be beneficial to society.

The real challenge lies in keeping technological progress in line with human values. Developing AI should not be narrowly defined by these categories, but they should be at the heart of its development. The goal of AI systems should not be to take the place of human moral responsibility nor to replace human autonomy. AI systems should support humans and not replace human moral responsibility and/or human autonomy.

The future of AI is a responsible one, and it needs the cooperation of citizens, teachers, researchers, educational institutions, AI technology developers and governments. Education is especially crucial, as future generations need to know how to interact with AI and critically evaluate and responsibly manage it.

Success of Artificial Intelligence should not be judged simply on its computation or economic value, but by the well-being, dignity, equality and social progress it provides for human beings. Innovation and ethics do not contradict. AI, with its human-centric advancement, can prove itself to be a formidable tool in fostering a more inclusive, sustainable, and compassionate future.

References:

1. Bhalla, N., Brooks, L., & Leach, T. (2024). Ensuring a 'responsible' AI future in India: RRI as an approach for identifying the ethical challenges from an Indian perspective. *AI and Ethics*, 4, 1409-1422. <https://doi.org/10.1007/s43681-023-00370-w>
2. Biju, P. R., & Gayathri, O. (2025). Indic approach to ethical AI in automated decision making system: Implications for social, cultural, and linguistic diversity in native population. *AI and Crawford, K., & Calo, R. (2016). There is a blind spot in AI research. Nature*, 538, 311-313. <https://doi.org/10.1038/538311a>
3. 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, and recommendations. *Minds and Machines*, 28, 689-707. <https://doi.org/10.1007/s11023-018-9482-5>
4. Government of India. (2020). National education policy 2020. Ministry of Education.
5. Hajam, K. B., & Purohit, R. (2024). The ethics of artificial intelligence: A critical examination of moral responsibility and autonomy. *Indian Journal of Educational Technology*, 6(2), 398-407. <https://journals.ncert.gov.in/IJET/article/view/513>
6. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1, 389-399. <https://doi.org/10.1038/s42256-019-0088-2>
7. Lal, B., Lucini, B. & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/j.ijinfomgt.2020.101994>

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. Munshi, U. M., Yadav, G., Munshi, A., Swati, R. K., Singh, D., & Kumari, N. (2024). AI & ethics: Charting a responsible future. *Annals of Library and Information Studies*, 71(4), 496-505. <https://doi.org/10.56042/alis.v71i4.14291>
10. NITI Aayog. (2018). National strategy for artificial intelligence: #AIForAll. Government of India.
11. Organisation for Economic Co-operation and Development. (2019). Recommendation of the Council on Artificial Intelligence. OECD.
12. UNESCO. (2021). Recommendation on the ethics of artificial intelligence. UNESCO.
13. UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.
14. World Economic Forum. (2020). The future of jobs report 2020. World Economic Forum.