The Academic
International Journal of Multidisciplinary Research

An Online Peer Reviewed / Refereed Journal
Volume 4 | Issue 10 | October 2026
ISSN: 2583-973X (Online), SJIF IF(2026): 5.6
Website: www.theacademic.in

From Human Intelligence to Artificial Intelligence: An Investigation of Human Autonomy and Moral Judgement in the Age of AI

Sibin Thomas

II M.A. Philosophy, Department of philosophy, Kerala University, Kariavattom Campus.

DOI : https://doi.org/10.5281/zenodo.24492343

ARTICLE DETAILS

Research Paper

Received: 01-11-2026

Accepted: 07-11-2026

Published: 20-11-2026

Keywords:

Artificial Intelligence, Human Intelligence, Human Autonomy, Moral Judgement, Moral Responsibility, Meaningful Human Control.

ABSTRACT

Artificial intelligence (AI) is an important development in the history of human intelligence. Human beings have used imagination, invention, language, and cooperation to solve problems and improve their lives. AI extends some of these human abilities by processing information, identifying patterns, and supporting decisions. However, its growing use raises important questions about human autonomy and moral judgement. This paper examines how dependence on AI may influence the way people think, make decisions, and judge right and wrong. It asks whether AI simply supports human intelligence or changes the conditions under which people exercise their autonomy. The paper also explores the problem of moral responsibility when people rely on AI recommendations. It distinguishes between the morality of an action and the responsibility of the person who performs it. Drawing on Aristotle's account of voluntary action and the concept of meaningful human control, the paper argues that responsibility depends partly on a person's knowledge and control over a decision. AI assistance does not automatically remove human autonomy or responsibility. However, people must critically examine AI recommendations rather than accept them without question. The paper concludes that technological progress should be evaluated not only by AI's performance but also by its effects

Disclaimer: The opinions, interpretations, conclusions, recommendations, analyses, and statements expressed in this research article are solely those of the author(s) and do not reflect the views, policies, or positions of the Publisher, Chief Editor, Editors, Editorial Board, Reviewers, or their affiliated institutions. The author(s) are solely responsible for ensuring the originality, authenticity, and integrity of the manuscript and for ensuring that it is free from plagiarism, AI-generated content, copyright infringement, data falsification, and any other form of academic misconduct.

on independent thought, moral judgement, and responsible human action.

Introduction

One of the significant developments in human history is the capacity for imagination, invention, and collective cooperation. In *Sapiens: A Brief History of Humankind*, Harari (2014) explains how the Cognitive Revolution enabled human beings to develop sophisticated language and share imagined realities, including myths, nations, laws, and social institutions. This capacity for collective imagination facilitated large-scale cooperation and contributed to the development of human civilisation. Artificial intelligence can be understood within this broader history of human invention as a technological development through which human beings seek to extend their intellectual capabilities and perform tasks that previously required human effort. Today, AI is increasingly used in various fields, including medicine, agriculture, business, art, and education. AI systems can support diagnosis, analyses information, generate creative content, and assist decision-making. Although these applications offer significant practical benefits, they also raise questions about the relationship between technological assistance and human intellectual activity. As individuals increasingly rely on AI to obtain information, evaluate alternatives, and make decisions, AI may influence not only what they decide but also how they arrive at their decisions. Research on intelligent decision-support systems highlights the importance of examining how such technologies affect human ethical decision-making.

This development brings human autonomy and moral judgement into philosophical focus. If individuals increasingly depend on AI to interpret situations and recommend actions, does this dependence merely extend their capabilities, or does it transform the conditions under which they exercise independent judgement? Although technological assistance does not necessarily diminish autonomy, uncritical reliance on AI may weaken a person's ability to evaluate recommendations independently. This concern also raises the problem of moral responsibility: when an AI-assisted decision leads to a morally questionable outcome, who should be held responsible, and under what conditions? It is therefore necessary to distinguish the moral evaluation of an action from the responsibility of the person who performs it. This article critically examines the relationship between AI reliance, human autonomy, moral judgement, and responsibility. It argues that AI assistance does not automatically eliminate human autonomy, but preserving meaningful autonomy requires individuals to retain sufficient knowledge and control over their decisions. Ultimately, the article investigates whether human beings can remain independent moral agents when the technologies they create increasingly influence how they think, judge, and act.

From Human Intelligence to Artificial Intelligence

Human intelligence can be understood through the relationship between human needs, problem-solving, imagination, and invention. In the struggle for survival, early humans faced practical challenges such as obtaining food, which required observing their surroundings, anticipating animal movements, and developing effective hunting strategies. Such activities illustrate how practical challenges can encourage learning, planning, and cooperation. Archaeological research on early stone-tool production suggests that toolmaking involved complex, goal-directed actions and cognitive processes associated with technological planning. These findings also suggest connections between the cognitive processes involved in toolmaking and those involved in language, although the precise evolutionary relationship remains debated. Language and communication enabled humans to share experiences, transmit knowledge, and coordinate their activities, while imagination allowed them to consider possibilities beyond their immediate circumstances. Human intelligence is therefore expressed not only through thinking but also through the ability to devise new ways of responding to practical difficulties. It develops not only through individual problem-solving but also through the ability to share ideas, coordinate activities, and transform individual knowledge into collective action. An idea becomes more influential when it can be communicated, evaluated, and developed through cooperation. Harari, in Chapter 2 of *Sapiens: A Brief History of Humankind*, emphasizes how the human capacity to imagine and share realities such as myths, nations, laws, and social institutions enabled cooperation among large numbers of people. This capacity for collective imagination is significant for technological development because invention does not occur in isolation: knowledge, practical skills, and ideas can be transmitted across individuals and generations, allowing new innovations to build upon existing achievements. The development of technology can therefore be understood as part of a cumulative process through which human beings preserve and extend their intellectual capacities.

Artificial intelligence can be situated within this broader history of human invention. Human beings have continually developed tools and technologies to extend their capabilities, and AI represents a significant development because some systems can perform tasks involving pattern recognition, information processing, prediction, and content generation. However, successful performance of these tasks does not establish that artificial systems possess intelligence identical to human intelligence, nor does it establish that they possess human-like consciousness. The distinction between task performance and human cognition is therefore important when examining the development of AI. From this perspective, AI can be understood as a technological development through which certain functions previously performed by humans can be assisted or partially undertaken by machines. It represents a further development in the cumulative history of human invention, emerging from accumulated knowledge, scientific research, and engineering. By

performing tasks involving information analysis, communication, content generation, and decision-making, AI systems can extend or partially undertake functions previously performed by humans. However, the significance of AI lies not merely in its origin as a human creation but also in its potential influence on the people who use it. As humans increasingly rely on AI to acquire information, evaluate alternatives, and guide decisions, a philosophical question emerges: does AI merely extend the intellectual capacities of human beings, or can increasing dependence on these systems also transform the conditions under which humans exercise their intelligence, autonomy, and moral judgement? This question provides the conceptual bridge between the collective development of human intelligence and the examination of human agency in the age of AI.

AI and the Problem of Human Moral Judgement

Moral judgement refers to the evaluation of human actions as morally right or wrong according to particular moral norms or criteria. A moral norm functions as a standard for assessing actions, intentions, and motives and determining their moral significance. Moral norms are commonly discussed in terms of subjective and objective standards: subjective norms relate to the individual's conscience, whereas objective norms appeal to standards that are not determined solely by individual feelings or preferences. For a moral norm to provide a reliable basis for judgement, it should ideally demonstrate consistency, universality, and applicability across different circumstances. Conscience occupies a central place in subjective moral judgement because it enables individuals to reflect on their conduct and recognize what they ought or ought not to do. Mahatma Gandhi associated conscience with an inner moral voice. Joseph Butler, in his account of human nature and morality, distinguishes the reflective function of conscience, through which a person evaluates actions and motives, from its authoritative function, through which conscience exercises a governing role in determining conduct. Thus, moral judgement involves both the assessment of an action's moral character and the practical consideration of how one ought to act. Autonomy is the capacity to govern one’s own actions through reasons that one accepts as one’s own. Kant describes this as acting on principles that one gives oneself. Frankfurt describes it as endorsing, on reflection, the desires that move one to act. Both views suggest that autonomous agents can evaluate their options for themselves. This evaluation must not be wholly controlled by an outside source. This idea connects to the knowledge and control conditions discussed below.

The morality of artificial intelligence concerns whether AI systems can make morally appropriate decisions and how their actions are evaluated according to moral norms. Research suggests that people may apply different moral expectations to human and artificial agents. Malle et al. (2015) found that people

sometimes considered utilitarian actions, such as sacrificing one person to save several others, more permissible when performed by a robot than by a human. Voiklis et al. (2016) also investigated how people judge the moral behavior of human and robot agents. However, findings are not uniform: Komatsu et al. (2021) found that judgements of blame towards reluctant robots can parallel those directed at humans in moral dilemmas. Furthermore, Bigman and Gray (2018) found that people are often reluctant to delegate moral decisions to machines, particularly when human life is at stake. These findings indicate that public evaluations of AI morality are complex and context-dependent. Nevertheless, judging an AI decision as morally acceptable does not establish that the system genuinely understands moral principles or possesses moral agency.

The problem of responsibility when humans rely on AI

The problem of responsibility in human–AI interaction requires us to distinguish between the morality of an action and the responsibility of the person who performs it. Moral evaluation asks whether an action is right or wrong, while responsibility asks whether a person can be held accountable for that action. Crompton discusses this problem through the idea of a decision point. Consider Carol, who receives an output (x) from an AI system and forms her own conclusion (y). She then acts on her conclusion. Although Carol makes the final decision, the AI may influence how she understands the situation and evaluates the available choices. Therefore, identifying Carol as the decision-maker does not automatically establish that she has acted independently. This problem can be examined through Aristotle's distinction between voluntary and involuntary actions. Aristotle argues that voluntariness is important when determining whether a person deserves praise or blame. In human–AI interaction, two conditions become important: control and knowledge. The control condition asks whether Carol genuinely makes her own decision or whether excessive reliance on AI significantly influences her judgement. The knowledge condition asks whether she understands the relevant situation, the limitations of the AI's output, and the possible consequences of following its recommendation. This is related to the idea of meaningful human control discussed by Santoni de Sio and van den Hoven (2018), who argue that human involvement must be more than a formal requirement. However, even if Carol has sufficient knowledge and control, her decision may still be morally wrong. For example, she may knowingly make an unjust decision after carefully considering the AI's recommendation. In this case, she may be responsible for her action even though the action is morally unacceptable. Similarly, an action may be morally wrong without Carol being fully blameworthy if her knowledge or control was seriously limited. My central argument is that the morality of an action and the responsibility of the agent are two different questions. The first concerns whether the action is right or wrong. The second concerns whether the person can reasonably be held responsible for it. Therefore, in the

age of AI, identifying the human being as the final decision-maker is not enough. We must also examine whether the person has sufficient knowledge and control to make a responsible decision. We must also consider the roles of AI developers and institutions, since they may share responsibility for the outcome.

Critical Discussion

AI reliance creates a tension between human autonomy and moral responsibility. Although humans may retain the final decision, their autonomy can be weakened when AI significantly influences their judgement or replaces independent evaluation. However, reduced autonomy does not automatically eliminate responsibility, just as making an autonomous decision does not guarantee that an action is morally right. Aristotle's account of voluntary action highlights the importance of knowledge and control in attributing responsibility. Therefore, the morality of an action and the responsibility of the agent must be evaluated separately. The critical issue is whether humans can remain morally responsible when their decisions are increasingly shaped by AI systems. AI may assist human judgement, but meaningful autonomy requires individuals to retain the ability to question, evaluate, and reject AI recommendations. AI assistance may not always provide an adequate moral conclusion when it applies universal rules to situations that require contextual judgement. Moral decisions often involve competing values, cultural differences, and circumstances that cannot be resolved by applying a single rule mechanically. As Serafimova argues, the ability of an AI system to produce apparently rational ethical judgements does not establish that it qualifies as an autonomous moral agent. However, moral diversity does not necessarily mean that all moral judgements are relative or that universal principles are impossible. Rather, AI recommendations must be critically evaluated in relation to the particular circumstances and the values involved. Therefore, humans should not accept AI-generated moral conclusions as universally valid without examining their justification and context.

Conclusion

Artificial intelligence represents a significant development in the history of human intelligence and invention, offering new possibilities for extending human capabilities. However, its increasing influence on human decision-making raises fundamental questions about autonomy, moral judgement, and responsibility. Although AI can provide information and recommend actions, its outputs cannot always resolve moral questions that require contextual understanding and critical evaluation. The central challenge is therefore not simply whether AI can make decisions, but whether human beings can preserve their capacity to judge, act, and take responsibility while relying on it. This article argues that AI assistance does not necessarily undermine human autonomy; rather, the extent to which autonomy is preserved depends on whether

individuals retain meaningful control over their decisions and critically evaluate AI-generated recommendations. Ultimately, technological progress must be assessed not only by how effectively AI performs human tasks, but also by how its use affects the human capacity for independent thought, moral judgement, and responsible action.

References

Aristotle. (n.d.). *Nicomachean ethics* (Book III, Chapter 1).

Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition, 181*, 21–34. https://doi.org/10.1016/j.cognition.2018.08.003

Butler, J. (1726). *Fifteen sermons preached at the Rolls Chapel*.

Catena, E. (2026). AI and human autonomy: A literature review. *AI and Ethics, 6*, Article 126. https://doi.org/10.1007/s43681-025-00958-4

Crompton, L. (2021). The decision-point-dilemma: Yet another problem of responsibility in human-AI interaction. *Journal of Responsible Technology, 7–8*, Article 100013. https://doi.org/10.1016/j.jrt.2021.100013

Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy, 68*(1), 5–20. https://doi.org/10.2307/2024717

Gandhi, M. K. (n.d.). *The mind of Mahatma Gandhi*. Section: “The Inner Voice” / “Voice of God.”

Harari, Y. N. (2015). *Sapiens: A brief history of humankind*. Vintage.

Harris, J., & Dubljević, V. (2025). Navigating the ethics of artificial intelligence. *Encyclopedia, 5*(4), Article 201. https://doi.org/10.3390/encyclopedia5040201

Kant, I. (1998). *Groundwork of the metaphysics of morals* (M. Gregor, Trans.). Cambridge University Press. (Original work published 1785)

Komatsu, T., Malle, B. F., & Scheutz, M. (2021). Blaming the reluctant robot: Parallel blame judgments for robots in moral dilemmas across US and Japan. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 63–72).

Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015). Sacrifice one for the good of many? People apply different moral norms to human and robot agents. In *Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction* (pp. 117–124). https://doi.org/10.1145/2696454.2696458

Moore, G. E. (1903). *Principia ethica*. Cambridge University Press.

Poszler, F., & Lange, B. (2024). The impact of intelligent decision-support systems on humans' ethical decision-making: A systematic literature review and an integrated framework. *Technological Forecasting and Social Change, 204*, Article 123403. https://doi.org/10.1016/j.techfore.2024.123403

Rachels, J., & Rachels, S. (2021). *The elements of moral philosophy* (9th ed.). McGraw Hill.

Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Santoni de Sio, F., & van den Hoven, J. (2018). Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI, 5*, Article 15. https://doi.org/10.3389/frobt.2018.00015

Serafimova, S. (2020). Whose morality? Which rationality? Challenging artificial intelligence as a remedy for the lack of moral enhancement. *Humanities and Social Sciences Communications, 7*, Article 119. https://doi.org/10.1057/s41599-020-00614-8

Stout, D., & Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1585), 75–87. https://doi.org/10.1098/rstb.2011.0099

Stout, D., Toth, N., Schick, K., & Chaminade, T. (2008). Neural correlates of Early Stone Age toolmaking: Technology, language and cognition in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1499), 1939–1949. https://doi.org/10.1098/rstb.2008.0001

Voiklis, J., Kim, B., Cusimano, C., & Malle, B. F. (2016). Moral judgments of human vs. robot agents. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 775–780).

[1] Harari, Y. N. (2015). Sapiens: A brief history of humankind. Vintage.

[1] Russell, S. J., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.

[1] Poszler, F., & Lange, B. (2024). The impact of intelligent decision-support systems on humans' ethical decision-making: A systematic literature review and an integrated framework. Technological Forecasting and Social Change, 204, Article 123403. https://doi.org/10.1016/j.techfore.2024.123403

[1] Catena, E. (2026). AI and human autonomy: A literature review. AI and Ethics, 6, Article 126. https://doi.org/10.1007/s43681-025-00958-4

[1] Poszler, F., & Lange, B. (2024).

[1] Stout, D., & Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1585), 75–87. https://doi.org/10.1098/rstb.2011.0099.

[1] Stout, D., & Chaminade, T. (2012).

[1] Harari, Y. N. (2015). Sapiens: A brief history of humankind. Vintage.

[1] Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

[1] Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

[1] Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

[1] Moore, G. E. (1903). Principia ethica. Cambridge University Press.

[1] Rachels, J., & Rachels, S. (2021). The elements of moral philosophy (9th ed.). McGraw Hill.

[1] Kant, I. (1785/1998). Groundwork of the metaphysics of morals (M. Gregor, Trans.). Cambridge University Press.

[1] Gandhi, M. K., “The Inner Voice,” in The Mind of Mahatma Gandhi, section “Voice of God.”

[1] Butler, J. (1726). Fifteen sermons preached at the Rolls Chapel.

[1] Kant, I. (1998). *Groundwork of the metaphysics of morals* (M. Gregor, Trans.). Cambridge University Press. (Original work published 1785)

[1] Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy, 68*(1), 5-20. https://doi.org/10.2307/2024717

[1] Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015). Sacrifice one for the good of many? People apply different moral norms to human and robot agents. In Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction (pp. 117–124). https://doi.org/10.1145/2696454.2696458

[1] Voiklis, J., Kim, B., Cusimano, C., & Malle, B. F. (2016). Moral judgments of human vs. robot agents. In Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (pp. 775–780).

[1] Komatsu, T., Malle, B. F., & Scheutz, M. (2021). Blaming the reluctant robot: Parallel blame judgments for robots in moral dilemmas across US and Japan. In Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (pp. 63–72).

[1] Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. Cognition, 181, 21–34. https://doi.org/10.1016/j.cognition.2018.08.003

[1] Crompton, L. (2021). The decision-point-dilemma: Yet another problem of responsibility in human-AI interaction. Journal of Responsible Technology, 7–8, Article 100013. https://doi.org/10.1016/j.jrt.2021.100013

[1] Aristotle. Nicomachean Ethics, Book III, Chapter 1, 1109b30–1111b3.

[1] Santoni de Sio, F., & van den Hoven, J. (2018). Meaningful human control over autonomous systems: A philosophical account. Frontiers in Robotics and AI, 5, Article 15. https://doi.org/10.3389/frobt.2018.00015

[1] Crompton, L. (2021)

[1] Aristotle. Nicomachean Ethics, Book III, Chapter 1.

[1] Santoni de Sio, F., & van den Hoven, J. (2018).

[1] Serafimova, S. (2020). Whose morality? Which rationality? Challenging artificial intelligence as a remedy for the lack of moral enhancement. Humanities and Social Sciences Communications, 7, Article 119. https://doi.org/10.1057/s41599-020-00614-8

[1] Harris, J., & Dubljević, V. (2025). Navigating the ethics of artificial intelligence. Encyclopedia, 5(4), Article 201. https://doi.org/10.3390/encyclopedia5040201