The deconstruction of humanity as an end under Artificial Intelligence

Zhaohan Chen

College of Social Science and Law, University College Dublin, Dublin, Ireland

zhaohan.chen@ucdconnect.ie

Abstract. This paper examines the relevance of Immanuel Kant's principle that humanity must be treated as an end in itself in the context of artificial intelligence (AI). Kant's moral philosophy posits that human beings, by virtue of rationality and autonomy, possess dignity beyond price and must never be reduced to mere instruments. Yet contemporary AI practices—including data-driven profiling, algorithmic decision-making, and predictive governance—risk objectifying persons, reshaping power relations, and legitimizing control under the guise of neutrality. Such developments undermine the status of individuals as autonomous agents and threaten to erode the moral community grounded in respect for human dignity. In response, the paper argues for ethical reconstruction and the establishment of clear moral constraints on AI. These include prohibiting systems that circumvent meaningful consent or reduce persons to data commodities, while also promoting human-centered designs that enhance agency in fields such as healthcare and education. The analysis draws on both Kantian ethics and contemporary discussions of AI governance to highlight pathways for aligning technology with respect for persons. The conclusion affirms that while AI may transform the conditions of human life, it cannot alter the fundamental truth that persons are ends in themselves. Societies thus bear responsibility to ensure that technological progress consistently reflects and safeguards human dignity.

Keywords: Kant's principle of humanity as an end, Artificial Intelligence ethics, ethical reconstruction, human-centered AI

1. Introduction

The concept of "human beings as ends in themselves" originates from Immanuel Kant's Groundwork of the Metaphysics of Morals, where he states: "So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means" [1]. This principle affirms the intrinsic worth and dignity of persons, distinguishing them from mere objects that can be exchanged or replaced. By grounding moral value in human rationality and autonomy, Kant established a universal criterion for evaluating actions and institutions, making his idea foundational to modern ethics and human rights discourse. In the contemporary era, however, the rapid development of artificial intelligence (AI) poses a significant challenge to this anthropocentric framework. Algorithmic decision-making, large-scale data extraction, and automated systems risk reducing individuals to replaceable data points, thereby undermining the respect for autonomy and dignity emphasized by Kant. This tension raises pressing questions: How should the principle of treating persons as ends be applied to contexts where human agency is mediated by intelligent technologies? Does the rise of AI demand a reinterpretation or reconstruction of the Kantian imperative? This paper adopts a theoretical and literature-based approach to analyze the tension between the Kantian principle of "human beings as ends" and the growing influence of AI. By examining the challenges posed by AI to human dignity, autonomy, and moral responsibility, it argues that the ethical landscape must be restructured to ensure that technological advancement remains consistent with the principle of respecting humanity as an end in itself.

2. The Kantian principle of humanity as an end in itself

"Rational beings are called persons because their nature already marks them out as ends in themselves, that is, as something which ought not to be used merely as means." [1] Kant's principle that humanity must be regarded as an end in itself arises from his attempt to ground morality in pure reason rather than in contingent desires or external goals. For Kant, human beings are distinguished by their capacity for rational self-legislation: they are not merely subject to natural impulses but can impose moral law upon themselves through the exercise of reason. This autonomy is what gives each person a dignity beyond comparison, a

value that cannot be reduced to calculations of benefit or exchange. To treat humanity as an end is therefore to recognize this unique status of persons as self-determining agents.

A crucial feature of the principle is its prohibition against reducing persons to objects of manipulation. Kant emphasizes that while it is inevitable to involve others as means in practical life, morality demands that their rational agency never be disregarded. What is condemned is not the use of another's skills or services, but the denial of their capacity to consent and to pursue their own ends. When a person is deceived, coerced, or objectified, their will is subordinated to purposes they cannot rationally endorse, and they are treated merely as a tool. Such conduct fails to recognize the inherent worth that belongs to them as rational beings.

The broader ethical implication of this view is that every individual must be seen as a co-legislator of the moral law. Human beings are not passive recipients of duty; they participate in the very creation of universal moral principles. This idea culminates in Kant's vision of a "kingdom of ends," an ideal moral community in which all persons relate to one another as equal bearers of dignity and as authors of shared norms. The principle of humanity as an end thus serves not only as a guide to personal conduct but also as a criterion for judging the justice of social and institutional arrangements.

3. Artificial Intelligence and the challenge to the principle of humanity as an end

3.1. The objectification of persons: AI and the challenge to human dignity

One of the most pressing challenges that artificial intelligence poses to Kant's principle of humanity as an end is the risk of objectifying persons. In order to function, AI systems rely on vast amounts of personal data—ranging from online behaviors and biometric identifiers to social and economic records. In this process, individuals are often reduced to quantifiable data points or algorithmic profiles. While this abstraction is technically necessary for computational efficiency, it carries the danger of erasing the uniqueness of the individual as a rational agent. Instead of being respected as ends in themselves, persons risk being treated as interchangeable variables within a system optimized for prediction or efficiency.

The objectification problem becomes evident in contexts such as targeted advertising, algorithmic hiring, and credit scoring. In these cases, individuals are not treated as autonomous decision-makers but are processed as data objects to be influenced, filtered, or ranked. For Kantian ethics, this is morally troubling because it disregards the person's capacity for rational self-determination. When systems treat individuals merely as inputs to determine their opportunities, access, or social worth, the principle of respecting humanity as an end is undermined.

Moreover, objectification through AI has a cumulative social effect. By normalizing the reduction of persons to data profiles, institutions may begin to conceive of human beings less as participants in moral and political life and more as manageable resources. This risks not only individual harm but also a broader erosion of the dignity that Kant viewed as the foundation of moral community. Safeguarding against this trend requires critical reflection on how AI systems are designed and deployed, ensuring that human beings remain recognized not simply as data sources but as ends in themselves [2].

3.2. Power, neutrality, and the question of human uniqueness

Artificial intelligence not only raises concerns about objectification but also reshapes the dynamics of power and authority in contemporary societies. The extensive use of AI in governance, commerce, and surveillance redistributes power in ways that often elude democratic oversight. Corporations and governments that control large-scale data infrastructures gain unprecedented capacities to predict, influence, and regulate human behavior. By contrast, individuals are placed in increasingly asymmetrical positions, where their autonomy and decision-making are constrained by systems they neither design nor fully understand. For instance, in the advertising and media industry, creative processes once grounded in human artistry are increasingly replaced by AI-generated images and texts. In this shift, both artists and consumers risk being reduced to instruments for maximizing click-through rates and conversion metrics, rather than being valued as ends in themselves. As Kant mentioned, "Man, and in general every rational being, exists as an end in himself, not merely as a means to be arbitrarily used by this or that will" [1]. This practice of simplifying people into manipulable variables violates Kant's categorical imperative.

The problem is compounded by the widespread perception of AI as a "neutral" tool. Under the guise of neutrality, such systems obscure the normative choices embedded in their design and legitimize the subordination of human beings to technological governance [3]. Yet this alleged neutrality is misleading. AI systems are trained on historical data, shaped by the interests of their designers, and embedded in social structures. Their outputs tend to reproduce and even reinforce existing social biases rather than transcend them. For example, in public security, facial recognition and predictive policing systems are often promoted as efficient and neutral, but in practice they rely on biased datasets that disproportionately target marginalized communities.

This "myth of neutrality" directly challenges the uniqueness of humanity as rational agents. When questions of opportunity, justice, or worth are delegated to algorithms under the claim of neutrality, human beings risk being relegated from recognized

autonomous subjects to mere data objects. To preserve the Kantian imperative, it is crucial to expose the fiction of neutrality and to acknowledge AI as a locus of power that requires ethical scrutiny. Only by doing so can we safeguard equal dignity and autonomy in the age of intelligent systems.

4. Reconstructing ethics in the age of Artificial Intelligence

4.1. Ethical reconstruction and the moral constraints of Artificial Intelligence

As artificial intelligence increasingly shapes human life, Kant's moral philosophy provides a necessary anchor for ethical reconstruction. In the Groundwork of the Metaphysics of Morals, Kant distinguishes between things that have a price and those that possess dignity: "In the kingdom of ends everything has either a price or a dignity. What has a price can be replaced by something else as its equivalent; what on the other hand is raised above all price, and therefore admits of no equivalent, has a dignity" [1]. This insight emphasizes that human beings, as bearers of dignity, can never be treated as interchangeable commodities.

AI itself cannot be regarded as a moral agent in the Kantian sense, as it lacks rational self-legislation. Responsibility therefore falls on human designers, policymakers, and users to ensure that AI systems operate within boundaries that respect human dignity. These constraints must prevent algorithms from bypassing meaningful consent, reducing individuals to mere data resources, or eliminating opportunities for appeal and contestation. Safeguards such as explainability requirements, human-in-command mechanisms, and enforceable legal standards are necessary to embed these values into AI development.

Ethical reconstruction, however, is not confined to prohibitive measures. In line with Kant's conception of imperfect duties, institutions also bear positive obligations to design AI systems that enhance human agency. For instance, in healthcare, AI should support doctors and patients in making informed decisions rather than substituting for their judgment. In education, AI ought to expand self-directed learning opportunities rather than constrain students to algorithmically predetermined pathways. These constructive uses of AI illustrate how technology can be aligned with the Kantian imperative when properly constrained and directed [4].

4.2. Human-centered development and the future affirmation of human dignity

The future of AI development depends on embedding human-centered values into both national policies and international frameworks. The cross-border nature of AI demands not only technical standards but also ethical ones that safeguard human dignity across cultural and political contexts. Initiatives such as the Ethics Guidelines for Trustworthy AI issued by the European Commission [5] demonstrate efforts to institutionalize principles of autonomy, transparency, and accountability. These guidelines illustrate that ethical reconstruction can move beyond theory and into practical regulation, shaping the trajectory of AI in ways that affirm the intrinsic worth of persons.

Looking ahead, the challenge lies in sustaining a global consensus that human beings must remain the reference point of all technological development. Public institutions, industry actors, and civil society must collaborate to ensure that efficiency and profit do not override respect for dignity. Moreover, citizens should be equipped with critical understanding to engage with AI systems that increasingly influence their lives.

Most importantly, Kant's principle remains the cornerstone: persons are ends in themselves and cannot be reduced to instruments of technological progress. AI may transform the conditions of human existence, but it cannot alter this fundamental moral truth. The legitimacy of AI systems is therefore contingent upon their consistency with human dignity. By pursuing human-centered development and affirming this principle in future governance, societies can ensure that technological innovation remains not a challenge to humanity's moral status, but an opportunity to reinforce it.

5. Conclusion

The analysis of artificial intelligence through the lens of Kant's principle that humanity must be treated as an end reveals both profound challenges and important opportunities. AI technologies, from data-driven advertising to predictive policing, risk reducing individuals to objects of management or instruments of efficiency. Yet these developments do not undermine the fundamental moral truth that human beings possess dignity beyond price. Rather, they demonstrate the urgency of reaffirming this truth in new technological contexts.

Ethical reconstruction and human-centered development pathways show that AI can and must be governed in ways that respect autonomy, accountability, and equality. Regulatory frameworks and design choices alike must be shaped by the recognition that persons are never merely tools but always bearers of intrinsic worth. The irreducible uniqueness and plurality of human beings, qualities that no algorithm can replicate or replace.

Thus, the conclusion is clear: artificial intelligence may transform the conditions under which human beings live, but it cannot alter their status as ends in themselves. The responsibility falls to societies to ensure that every application of AI affirms, rather than erodes, the dignity that grounds our shared moral life.

References

[1] Kant, I. (1785/2012). Groundwork of the Metaphysics of Morals (M. Gregor & J. Timmermann, Trans.). Cambridge: Cambridge University Press.

[2] Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1(1).

[3] Brey, P. (2021). The Strategic Role of Ethics in AI Design and Governance. *Ethics and Information Technology*, 23(4), 791–801.

[4] Jobin, A., Ienca, M., & Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, 1(9), 389–399.

[5] European Commission. (2019). Ethics Guidelines for Trustworthy AI. Brussels: European Commission.