SINGULARITY AND THE LIMITS OF THE HUMAN SUBJECT: A KANTIAN PERSPECTIVE ON THE FUTURE OF ARTIFICIAL INTELLIGENCE

Zhiwu Zhang and John Giordano[1]

ABSTRACT

This paper discusses the issue of technological singularity and analyzes its far-reaching impact on the future of mankind in combination with Kant's transcendental philosophy. Technological singularity is often understood as the point at which artificial intelligence surpasses human intelligence, leading to discussions about human free will, ethics, and social structures. However, this paper argues that technological singularity is not necessarily a threat, but a natural result of human reason and technological innovation. Using Kant's critical philosophy, this paper considers artificial intelligence from the perspective of human motivation, moral responsibility, and the limits of the human subject. This paper aims to provide a framework for philosophical reflection on the technological singularity in order to promote a deeper understanding of the relationship between technology and humanity.

Keywords: Technological Singularity; Artificial Intelligence; Kantian Philosophy; Ethics

[1] Graduate Program in Philosophy and Religion, Assumption University of Thailand. Zhiwu Zhang can be reached at: zhangzhiw1995@163.com. John Giordano can be reached at: jgiordano@au.edu.

Prajñā Vihāra Vol. 26 no. 2 July to December 2025, 99-121

DOI: https://doi.org/10.59865/prajn.2025.11

Introduction

Since the concept of Technological Singularity was proposed, it has become a hot topic in scientific and philosophical circles. On the one hand, it raises the idea that AI may surpass human intelligence. Some scholars have expressed deep concerns about the future status of human beings, free will and existential risks brought about by the development of AI. Karamjit S. Gill discussed the debate on AI put forward by the philosopher Harry Collins.[2] Collins criticizes the philosophical view of "equating computers with human brain," arguing that the algorithms of AI is essentially a "top-down" process, and separated from the formations of human society. This indicates that human logic is rooted in its own environment, society, culture and other factors, but AI cannot 'feel' such an environment. This would be the biggest limiting factor in AI's inability to replace the role of the human brain. In terms of AI's potential to surpass humans and endanger the status of humanity, it is mainly human concerns about AI empowerment. Abeba Birhane and Jelle van Dijk argue against granting rights to robots, emphasizing that AI should be viewed as a tool mediating human existence rather than as entities deserving of rights.[3] They conclude that the primary ethical concern should be human welfare, and that the responsibility lies with those who design, sell, and deploy machines, rather than with the machines themselves.

These various concern about AI reflect more on how human being understand their own agency and morality and to a lesser extent on the power of AI itself. This is what the article wishes to address. It will employ Immanuel Kant's critical philosophy to try to investigate issues of the limits of knowledge, about the human being as an ethical subject, and the role of the imagination. It will conclude that AI should function as a human assistant and the anxiety of the danger of the overpowering

---

[2] Karamjit S. Gill. "Artifictional Intelligence: Against Humanity's Surrender to Computers." *AI & Society* 34, no. 2 (January 2, 2019): 391–392.

[3] Abeba Birhane and Jelle van Dijk. "Robot rights? Let's talk about human welfare instead." *In AIES '20: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. Association for Computing Machinery. 2020, 207-213.

of human beings is unjustified, at least if the limits of the Kantian subject are respected.

As Danaher believes, we should either admit that AI will bring apocalyptic results, or program and control AI to conform to people's beliefs.[4] We can reword this for the purposes of this essay to say that we need to program AI so that it will participate equally with autonomous free human beings and nature.

On Technological singularity

On the definition of technological singularity, Vernor Vinge first pointed out that "From the human point of view this change will be a throwing away of all the previous rules, perhaps in the blink of an eye, an exponential runaway beyond any hope of control",[5] this change refers to a shift from quantitative change to qualitative change, he argues that "It is a point where our models must be discarded and a new reality rules. As we move closer and closer to this point, it will loom vaster and vaster over human affairs till the notion becomes a commonplace".[6] AI technology is very much in line with this definition, and now more and more fields and industries are being popularized and restructured by AI. Artificial intelligence is both affecting human intelligence and moving beyond level of intelligence beyond human beings. This will even trigger a tipping point of unpredictable social change. This out-of-control state is also being promoted by economic systems. AI is absorbing the world's economy, which is also in line with the definition of technological singularity by Vinge. Our situation has taken on the character of Plato's allegory of the cave. Human beings now live surrounded by the curtains

---

[4] John Danaher. "Why AI Doomsayers are Like Sceptical Theists and Why it Matters." *Minds & Machines* 25, 2015, 231–246.

[5] Vernor, Vinge. "The Coming Technological Singularity: How to Survive in the Post-Human Era". In Vision-21: Interdisciplinary Science and Engineering in the Era of Cyberspace, 1993, NASA Conference Publication 10129. NASA Lewis Research Center, 12.

[6] Ibid., 12.

of big data, no one knows what it is true or false, and the Internet had become, as explained by Bostrom, "something more like a virtual skull housing an emerging unified super-intellect".[7] Another aspect of this is our involvement in the programming of AI. People inadvertently participate in AI's self-training through data their participation in cyberspace. But while AI uses human beings to learn, it has also created a lot of human beings who are developing skills at controlling or mastering AI. Yet the human mastery of AI does not yet address its philosophical problems and dangers. We still have not reflected adequately on AI's philosophical limits and consequences. They develop only according to the potential for profit and the flow of capital. We are still at a stage where human beings cannot consider the deeper consequences of their technology. They speculate about various future developments at a shallow utilitarian level. Some thinkers believe that technological singularity could cause artificial intelligence to run out of control, threatening the survival of humanity. Others believe that it is an inevitable result of the development of human science and technology and should be approached with a positive attitude. But the problem is much deeper. We need to ask if the human mind is becoming a product of programming? If this is the case, we must develop safeguards by distinguishing the proper limits of the human, the limits of AI and the limits of nature. To make these distinctions, we need to reflect on the character and limits of human knowledge. This is essentially a reflection on Kant's question: 'What can I know, what ought I do, and what can I hope for.'

Kant and AI

Initially, what might draw one's attention to the relationship of Kant's philosophy to AI is the comparison of *apriori* algorithms with the Kantian *a priori*. This leads one to consider AI as a kind of transcendental knowledge base that can be compared and contrasted to Kant's transcendental understanding of knowledge.

[7] Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford : Oxford University Press, 2013, 49.

While there hasn't been much research on Kant and AI compared to other philosophical approaches, the book *Kant and Artificial Intelligence* edited by Hyeongjoo Kim and Dieter Schönecker, provided an important contribution that systematically explored the intersection of Kant's philosophy and artificial intelligence.[8] It covered the theoretical, practical, and aesthetic dimensions of Kant's philosophy and AI.

For instance, in the theoretical philosophy section, Tobias Schlicht, through Kant's theory of cognition, analyzed the theoretical developments in contemporary cognitive science such as functionalism, embodied cognition, and prediction processing models. He particularly focused on the challenges posed by deep learning to Kant's cognitive concepts.[9] Richard Evans proposed the concept of an "apperception engine", applying Kant's priori psychology to the architecture of machine learning systems, exploring the specific implementation of Kant's theory of apperception in modern AI technology.[10] Sorin Baiasu further explored the issue of artificial intelligence's self-awareness. He analyzed through Kantian philosophy how AI constructs meaning and raised philosophical questions about whether AI can truly reach the level of human cognition.[11] Hyeongjoo Kim considered the philosophical basis of AI from the perspective of Kantian philosophy. He particularly focused on the definition of artificial intelligence proposed by John McCarthy and

---

[8] Hyeongjoo Kim and Dieter Schönecker. *Kant and Artificial Intelligence*. Walter de Gruyter GmbH & Co KG, 2022, 1-144.

[9] Tobias Schlicht. "1 Minds, Brains, and Deep Learning: The Development of Cognitive Science Through the Lens of Kant's Approach to Cognition". *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 3–38.

[10] Richard Evans. "The Apperception Engine." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 39–104.

[11] Sorin Baiasu. "The Challenge of (Self-) Consciousness: Kant, Artificial Intelligence and Sense-Making." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 105–128.

considered the philosophical foundation of artificial intelligence from the perspective of Kant's a priori concepts.[12]

The practical philosophy section of the book focused on ethical issues.[13] Lisa Benossi and Sven Bernecker explored robot ethics through Kantian ethics, questioning whether machines could truly become moral subjects, and emphasizing that from Kant's principle of respecting human life and reason, preventing robots from obtaining moral personality is ethically permissible.[14] Dieter Schonecker emphasized Kant's theory of moral emotions and argued that practical reason cannot be replicated or 'artificialized' by artificial intelligence.[15] Elke Elisabeth Schmidt analyzed the 'trolley problem' in autonomous driving ethical decision-making from the Kantian perspective, proposing a Kantian interpretation and solution to the AI ethical dilemma.[16] Ava Thomas Wright explored the issue of machine rights, analyzing whether machines can have some moral or legal status within the Kantian ethical framework.[17] And Claus Dierksmeier further proposed that AI should be regarded as a partner of

---

[12] Hyeongjoo Kim. "Tracing the Origins of Artificial Intelligence: A Kantian Response to McCarthy's Call for Philosophical Help." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 129–144.

[13] Hyeongjoo Kim and Dieter Schönecker. *Kant and Artificial Intelligence*. Walter de Gruyter GmbH & Co KG, 2022, 145-254.

[14] Lisa Benossi, and Sven Bernecker. "5 a Kantian Perspective on Robot Ethics." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 145–168.

[15] Dieter Schönecker. "6 Kant's Argument From Moral Feelings: Why Practical Reason Cannot Be Artificial." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 169–188.

[16] Elke Elisabeth Schmidt. "7 Kant on Trolleys and Autonomous Driving." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 189–222.

[17] Ava Thomas Wright. "8 Rightful Machines." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 223–238.

human autonomy rather than a tool, emphasizing the guiding role of the concept of Kantian autonomy in the development of AI technology.[18]

In the aesthetics section,[19] Larissa Berger, starting from Kant's theory of aesthetic judgment, explored whether AI can possess a truly meaningful aesthetic experience and provided a Kantian philosophical interpretation of artificial intelligence's aesthetic ability.[20]

Overall, this book provided a valuable theoretical resources for the study of philosophy and AI, and also leaves important theoretical space and exploration directions for subsequent research. Although this book has made significant contributions to the interdisciplinary research on Kant's philosophy and AI, it has certain limitations. On one hand, when the book was published, AI technology had not yet reached the level of the recent rapid development of generative AI, which led to which we need to consider in this paper. The field of AI is developing very quickly, and this connection of Kant and AI needs to be constantly considered anew. This is the intention of this paper. Kant's philosophy can contribute to the question of the dangers of technological singularity and how to avoid these dangers by allowing us to reflect on the limits of human knowledge and the preservation of the autonomy of the human subject.

Kant developed began developing his critical philosophy in the *Critique of Pure Reason*. A critical philosophy is one that investigates the conditions of the possibility of knowledge. In this case he is investigating the transcendental conditions which make knowledge of the world possible. Human reason is made possible and limited by *a priori* conditions. This is an attempt to understand the proper foundations and

[18] Claus Dierksmeier. "9 Partners, Not Parts. Enhanced Autonomy Through Artificial Intelligence? A Kantian Perspective." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 239–254.

[19] Hyeongjoo Kim and Dieter Schönecker. *Kant and Artificial Intelligence*. Walter de Gruyter GmbH & Co KG, 2022, 255-282.

[20] Larissa Berger. "10 on the Subjective, Beauty and Artificial Intelligence: A Kantian Approach." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 255–282.

limits of thought. He writes that "such universal cognitions, which at the same time have the character of inner necessity, must be clear and certain for themselves, independently of experience; hence one calls them *a priori* cognitions".[21] Kant goes on to explain this kind of *a priori* knowledge that cannot be acquired through experience, He writes, "if one removes from our experiences everything that belongs to the senses, there still remain certain original concepts and the judgments generated from them, which must have arisen entirely *a priori*."[22]

Kant distinguishes between phenomena and things-in-themselves in his philosophical system, and in his *Critique of Pure Reason*, Kant writes, "we can have cognition of no object as a thing in itself, but only insofar as it is an object of sensible intuition, i.e. as an appearance".[23] Kant believed that we can only recognize the phenomena of things, the way they appear in our senses, and that we do not have the concepts and elements that pure reason can provide for knowing things in the natural world (the "thing-in -itself") unless these concepts are conditioned by the intuition and the categories of the understanding.[24] We can consider this in comparison to the idea of "Heart" in Chinese neo Confucian philosophy and Buddhism and "Dao" in Chinese Taoism which can be understood as a law of the thing-in-itself. That is to say, the laws of things themselves involve the essence of things, and human reason can only recognize the phenomena of things.

AI can only "imagine" the real world through the processing of knowledge and images, just as human cognition is limited by senses and the categories of understanding and cannot fully grasp the nature of the object itself. It echoes the unknown technologies and forms of intelligence involved in the technological singularity, meaning that we need to acknowledge the limitations of our own cognition and explore the unknown with an open mind. The human being is the most important

[21] Ibid., 127.

[22] Immanuel Kant. *Critique of Pure Reason*. Cambridge University Press, 1998, 128.

[23] Immanuel Kant. *Critique of Pure Reason*. Cambridge University Press, 1998, 115.

[24] Ibid., 115.

medium for bridging AI into our real world for the purpose of solving a problem. That is, AI needs to enter the real world through human beings to understand the world. The writer Gregory Bateson reflected on this in his book *Steps to an Ecology of Mind*.

> Now, let us consider for a moment the question of whether a computer thinks. I would state that it does not. What "thinks" and engages in "trial and error" is the man *plus* the computer *plus* the environment. And the lines between man, computer, and environment are purely artificial, fictitious lines. They are lines *across* the pathways along which information or difference is transmitted. They are not boundaries of the thinking system. What thinks is the total system which engages in trial and error, which is man plus environment.[25]

We will return to this, but here we can say in general that we need then to consider AI as involving both a relationship to the human and to nature. And we also need to take a step beyond Bateson; we need to recognize the limits of the human and the limits of nature in their interaction which creates 'thinking'.

Rethinking technological singularity through Kant involves a re-examination of the connection between humans and the world, despite their cognitive limitations. According to Kant, reason has the characteristic of pursuing infinity, and Kant believes that what reason seeks is an unconditional wholeness that cannot be found in experience.[26] Our infinite pursuit of scientific development drives us to constantly explore unknown fields beyond experience. In the process of scientific and technological development, this spirit of exploration is crucial. Through reason and imagination, human beings try to break through the limitations of cognition and discover new scientific principles and technological methods. It was

---

[25] Gregory Bateson, *Steps to an Ecology of Mind,* (University of Chicago Press, 2000), 488.

[26] Immanuel Kant, *Critique of Pure Reason*. Cambridge University Press, 1998, 391.

in this spirit that Kant developed his late work, *Metaphysical Foundations of Nature Science*. It was an attempt to provide a proper foundation for science (in Kant's case, Newtonian physics). If we follow the foundational project of Kant's philosophical thought, the technological singularity is not an inevitable event that poses a threat to mankind, but the culmination of human reason and its creation of new technologies. That is, the development of sciences and technology and the knowledge of *a priori* principles support one another at a historic point of time. This is what Kant recognized in his "Metaphysical Foundations of Natural Science":

> All true metaphysics is drawn from the essence of the faculty of thinking itself, and is in no way fictitiously invented on account of not being borrowed from experience. Rather, it contains the pure actions of thought, and thus a priori concepts and principles, which first bring the manifold of empirical representations into the law-governed connection through which it can become empirical cognition, that is, experience. Thus these mathematical physicists could in no way avoid metaphysical principles, and, among them, also not those that make the concept of their proper object, namely, matter, a priori suitable for application to outer experience, such as the concept of motion, the filling of space, inertia, and so on. But they rightly held that to let merely empirical principles govern these concepts would in no way be appropriate to the apodictic certainty they wished their laws of nature to possess, so they preferred to postulate such [principles], without investigating them with regard to their a priori sources.[27]

Kant argued that true natural science (such as Newtonian physics) must rely on both mathematics (formal structure) and metaphysics (a priori concepts). Mathematics provides a quantifiable formal framework (such as the mathematical expression of laws of motion), while metaphysics

[27] Immanuel Kant, *Kant: Metaphysical Foundations of Natural Science*. Cambridge University Press, 2004, 8.

offers the a priori categories of material nature, such as force, causality, and substance. The technological singularity, as the "possible outcome" brought about by technology, corresponds to Kant's understanding of the development of scientific knowledge – humans integrate empirical materials through a priori categories to expand the boundaries of cognition. When the accumulation of technology and the cognition of a priori principles reach a certain point, a qualitative change, such as a singularity, becomes possible. This is reflected in technological singularity: if technological development deviates from the a priori norms of human ethics, it may become a "threat"; conversely, if it conforms to rational norms, it becomes an advance in civilization. The philosophical connotation of the technological singularity does not lie in its technicality itself, but in how human rationality, through a priori principles, gives meaning and boundaries to technological development. This framework goes beyond simple technological determinism and transforms the singularity issue into a reflection on the mission of human rationality. The writer Philippe Verdoux refers to this as "inflationism." This would involve "amplifying the abilities of philosophers rather than reducing the ambitions of philosophy," thus advancing the study of philosophy in connection with science and technology.[28] In terms of advancing human decision-making (what Kant would call 'judgement') the logic of AI is not yet rigorous, but if it is harnessed to philosophical reflection, it can be controlled and developed. Reid McIlroy-Young et al. combines superhuman artificial intelligence with human behavior, especially discussing the development of AI from the perspective of chess as a model system and argued that "there is substantial promise in designing artificial intelligence systems with human collaboration in mind by first accurately modeling granular human decision-making".[29]

[28] Philippe Verdoux, "Emerging Technologies and the Future of Philosophy." *Metaphilosophy* 42, no. 5 (October 1, 2011): 682–707

[29] Reid McIlroy-Young, Siddhartha Sen, Jon Kleinberg, and Ashton Anderson. "Aligning Superhuman AI With Human Behavior: Chess as a Model System." *KDD '20: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, August 20, 2020, 1677–1687.

We should look at the technological singularity in a rational and prudent manner, actively exploring unknown areas and promoting technological progress, while fully considering possible ethical and social impacts.

Reason and Imagination

Kant's *Critique of Pure Reason* shows that the in act of knowing nature we condition nature; that the understanding does not merely follow natural laws but imposes laws on nature, what is known as Kant's 'Copernican turn'.[30] In the development of science and technology, human beings 'discover' the laws of nature through reason, they are also changing and reshaping the living environment. The creation of artificial intelligence is the extreme extension of this process. By designing algorithms and models, humans enable machines to have a certain level of intelligent logic, which involves the classification and ordering of data. This intelligence, while derived from human programming and training, also shows the great potential of technological development. But there's another way to bridge reality and knowledge, and that's imagination, Kant describes the imagination in his *Critique of Judgment* with reference to the concept of 'genius' as follows:

> The imagination (as a productive cognitive faculty) is, namely, very powerful in creating, as it were, another nature, out of the material which the real one gives it. We entertain ourselves with it when experience seems too mundane to us; we transform the latter, no doubt always in accordance with analogous laws, but also in accordance with principles that lie higher in reason (and which are every bit as natural to us as those in accordance with which the understanding apprehends empirical nature); in this we feel our freedom from the law of association (which applies to the empirical use of that faculty), in accordance with which

[30] Immanuel Kant, *Critique of Pure Reason*. Cambridge University Press, 1998, 260-263, see B159-B164.

> material can certainly be lent to us by nature, but the latter can be transformed by us into something entirely different, namely into that which steps beyond nature.[31]

Imagination is connected to freedom, the power to create new forms and possibilities, completely beyond the intuition of nature. Imagination enables humans to transcend existing experiences and knowledge, allowing for the creation of entirely new concepts and technologies. Human imagination begins to develop in early childhood. In Chinese, there is an idiom that states, "Children's words are unbridled," which can cast light on both the imperfections in early logic algorithms and the visual and cognitive abilities of young children. This phenomenon can also be observed in artificial intelligence when it generates text or videos. In the development process of artificial intelligence, it is human imagination and creativity that promote the continuous breakthrough of technology. From simple algorithms in the early days to deep learning and neural networks today, human imagination provides a steady stream of power for the development of technology, and human imagination is the most critical factor in human development.

But in Kant's *Critique of Judgment* we also see where the imagination reaches its limits. This is what Kant calls the "sublime."

> Natural beauty carries with it a purposiveness in its form, by which the object seems as it were predetermined for our power of judgment, so that this beauty constitutes in-itself an object of our liking. On the other hand, if something arouses in us, merely in apprehension and without any reasoning on our part, a feeling of the sublime, then it may indeed appear, in its form, contrapurposive for our power of judgment, incommensurate with our power of exhibition, and as it were violent to our imagination, and yet we judge it all the more sublime for that.[32]

---

[31] Immanuel Kant, *Critique of Judgment*. Translated W.S. Pluhar. (Hackett, 1987), 182.

[32] Ibid, 98-99,

Where the apprehension of beauty leads to the idea of a purpose of objects in nature to create pleasure, the subline creates a displeasure in the face of the power of nature which leads to 'respect' for the power of nature outside of us. This also allows us to recognize an inner 'purposiveness' of the human subject. In Kant's philosophy the sublime is what draws the limits to the autonomous and moral subject in relationship to nature.

When Kant speaks about beauty as well, he distinguishes between the beauty of art and the beauty of nature. Natural beauty is much deeper that the beauty of art. Even fine art is measured by its imitation of the deeper beauty of nature. And this becomes important for Kant. What he calls the "intellectual interest in the beauty of nature" becomes the sign of a good soul.

> And interest in the beautiful in art... provides no proof whatever that someone's way of thinking is attached to the morally good, or even inclined toward it. On the other hand, I do maintain that to take a direct interest in the beauty of nature.. is always a mark of a good soul; and that, if this interest is habitual, if it readily associates itself with the contemplation of nature, this fact indicates at least a mental attunement favorable to moral feeling.[33]

Kant will even go on to point out that 'beauty is the symbol of morality'. The task of Kant's *Critique of Judgment* was to consider judgements of taste based upon an *a priori* principles. In this case it is, human pleasure and displeasure. AI does not possess this ground for judgement. It cannot at this stage realize its own pursuit of knowledge, it cannot yet draw boundaries between itself and the human, between itself and nature. It does not find 'pleasure' in a world which exists beyond its 'understanding' and information processing. It does not feel 'displease' in its apprehension of nature which would develop a consciousness of its limits. Not only is it not grounded in the sensuality of thinking, but it

[33] Ibid, 165.

is also contaminated by its human programming.

Freedom and Moral Law

We need to now examine this moral subject. Not everything is in the realm of theoretical reason for Kant. For instance, we cannot know with certainty moral law. We can only use reason in a regulative way or a practical way to construct universals and guide us in our moral decisions. Kant proposed in the *Groundwork for the Metaphysics of Morals*, "Act only in accordance with that maxim through which you can at the same time will that it become a universal law".[34] This is the attempt to root moral decisions within the human subject, but it can also be extended beyond the subject.

The formula of the rational foundation for universal law can provide ethical guidelines to harmonize technological innovation, with nature and human freedom. With the rapid development of artificial intelligence technology, relevant ethical issues become more and more prominent. For example, algorithmic bias can lead to unfair treatment of certain social groups, privacy protection concerns the security and respect of personal data, the autonomy of AI may lead to a challenge to human control. These issues require us to be vigilant in the process of technological development and establish sound ethical norms and laws and regulations to ensure that the development and application of artificial intelligence meet ethical standards and serve the common interests of mankind. Kant will therefor write that

> Therefore freedom, even though it is not a quality of the will in accordance with natural laws, is not for this reason lawless, but rather it has to be a causality in accordance with unchangeable laws, but of a particular kind; for otherwise a free will would be an impossibility, freedom is an inherent characteristic of man, expressed as independence from the laws of nature, following only those moral laws which he

[34] Immanuel Kant, *Groundwork for the Metaphysics of Morals (A. W. Wood, Trans.)*. Yale University Press. 2002 (Original work published 1785), 37, See Ak 4:421.

has made for himself.[35]

In the *Critique of Practical Reason*, Kant further examined practical reason and the law of freedom, namely the moral law. This work explores innate moral principles and illustrates how humans act morally guided by free will. It's *a priori* principle is desire.[36] But the question arises for Kant of how freedom is connected to nature. In the introduction to the *Critique of Judgment*, Kant states that "The critique of pure theoretical reason, which was dedicated to the sources of all cognition *a priori* (hence also to that in it which belongs to intuition), yielded the laws of nature, the critique of practical reason the law of freedom, and so the a prior principles for the whole of philosophy already seem to have been completely treated".[37] But Kant realized there must be a bridge between the two.

> Hence an immense gulf is fixed between the domain of the concept of nature, the sensible, and the domain of the concept of freedom, the supersensible, so that no transition from the sensible to the supersensible... is possible, just as if they were two different worlds, the first of which cannot have an influence on the second... Hence it must be possible to think of nature as being such that the lawfulness in its form will harmonize with at least the possibility of [achieving] the purposes that we are to achieve in nature according to laws of freedom.[38]

---

[35] Immanuel Kant. *Groundwork for the Metaphysics of Morals (A. W. Wood, Trans.)*. Yale University Press. 2002 (Original work published 1785), 82, See Ak 4:447.

[36] Immanuel Kant, *Critique of Practical Reason*. Cambridge University Press, 1997, 3-13.

[37] Immanuel Kant, *Critique of the power of judgment*. Cambridge University Press, 2000, 8.

[38] Immanuel Kant, *Critique of Judgment*. Translated W.S. Pluhar. (Hackett, 1987), 14-15.

Through this book, Kant aims to bridge the gap between theoretical reason and practical reason through aesthetic judgment and teleological judgment. And this role of judgment as a kind of mediation is a very important consideration for AI. Kant in his three critiques tried to develop a unified framework of the *a priori* foundations of knowledge. The data (laws of nature) and algorithms (laws of freedom) of AI should also be understood under a unified framework, or a unified grand model of AI world. At present, Professor Li Feifei's team at Stanford University has devoted much research to developing a grand model of AI world. However, the philosophical community has not yet proposed a relatively foundational model of the AI world. Of course, because the development speed of AI is so fast, and it is restructuring the entire human society, the philosophical consequences may lag behind innovation. This leads to a question: if there is an intelligent agent beyond human beings, can human beings still use reason to criticize it? Can we still control it? Kant's philosophical thought tells us that the answer is yes, but we need to consider the motivation of human beings in creating AI, because the motivation of creating AI should not only be for wealth. Kant's critical philosophy needs to consider the condition for the possibility of AI knowledge, it limits where it passes into moral questions, and its motivation. Therefore, Kant's philosophical criticism provides a hard "free algorithm" foundation for AI to construct its own world.

Human motivation in the development of AI

The development of artificial intelligence is essentially the reflection of human beings on their own nature and the pursuit of perfection. Humanity seeks to expand and enhance its capabilities through technological means to better understand and transform the world. This motivation stems from the human desire for knowledge, efficiency and control. However, as some scholars have asked, is the human way of thinking and behavior also a product of programming? If the human mind is also "programmed", then is the development of artificial intelligence just a process of human self-replication and expansion?

In the pursuit of rapid development, humans constantly improve the function of tools to better apply to the objective world. Human beings use subjective imagination and rational powers to unite subjective thought with objective reality. In this process, the emergence of artificial intelligence seems inevitable, and even its possibility of surpassing humans is incorporated into our assumptions. However, this does not mean that humans will be replaced. Instead, AI can be a tool to augment human intelligence and help us solve more complex problems. Therefore, technological singularity is not a fixed and inevitable event. Our attention should instead be on how the two-dimensional input delivered to AI becomes a three-dimensional "phenomenon." As Bateson pointed out, we need to consider this "relationship" between the machine, the human and nature that leads to the process of thinking. The development of artificial intelligence will cause humans to think more deeply about their own uniqueness in the face of its potential threat. How can human beings understand that AI will replace them if they do not understand themselves? When machines can simulate or even surpass some human capabilities, we need to re-examine the position of human beings in nature, and we even need to re-examine the purpose of human existence now. Many scholars believe that AI is a stage in the evolution of human beings. Just like the process of biological evolution toward the development of the rational human being. But just like biological evolution, the new form must compliment the limitations of the previous one. Therefore, the difference between human beings and their previous incarnations is rational ability, and AI does not have a fully developed rational ability at present. Kant emphasized that human beings have rational and moral consciousness, which is the fundamental difference between human beings and other beings. So, although machines can surpass humans in some ways, they cannot replace humans as rational agents. Human emotions, moral judgment, and autonomous consciousness are still things that AI can't fully replicate because they lack *a priori* foundations. This understanding helps us adhere to human-centered principles in the application of technology that would otherwise give AI an introduction to independent thinking

(free algorithms).

Regarding people's concern that the development of artificial intelligence may lead to the emergence of a subject beyond human beings and threaten the survival and interests of human beings, humans should strive to understand the limits of the human subject and the limits of the machine.

This returns us to the core question of this article, where does the human motivation to create artificial intelligence come from? In essence, AI is motivated and trained by humans, and its goals and behavior are based on the human will to program the data. So, the dilemma brought about by artificial intelligence reflects the dilemma of humanity itself. If human beings cannot maintain a "people-oriented" approach to technology, human beings will be become more controlled. For example, in academic research, for those who have not formed strong critical philosophical thinking, AI is leading to the degeneration of critical thinking system.

Conclusion: Philosophical reflection and AI

Technological singularity should not be seen as an inevitable threat to human society, but rather as a natural outcome of human reason and technological accumulation. From Kant's philosophical perspective, reason not only possesses autonomy but also contains creative potential. The emergence of technological singularity is precisely the result of humanity's relentless pursuit of knowledge and innovation. Although the development of artificial intelligence may surpass human capabilities in certain fields, this does not mean that humanity will necessarily be replaced. On the contrary, artificial intelligence should be viewed as a tool to extend human capabilities and help us achieve higher societal goals. It should also motivate us to more deeply question ourselves and our position in nature.

We need to maintain a proactive attitude during the process of technological development, actively participating in and guiding the design and application of artificial intelligence to ensure that it aligns with the core values and ethical standards of human society. Kant's

moral philosophy reminds us that technological development must be constrained by moral laws and fulfill necessary social responsibilities. Only under a sound ethical framework and legal system can technology be applied in a transparent, just, and controllable manner.

The advancement of artificial intelligence undoubtedly brings many serious challenges, but it also presents great opportunities. By strengthening international cooperation and jointly establishing technical standards and ethical norms, we can promote the synergistic development of humanity and technology, advancing towards a sustainable future. Reexamining the relationship between humanity and nature will help us find the appropriate balance between technology and humanity. As suggested by Kantian philosophy, technological singularity is not a threat to humanity, but rather a reflection of the maturity of human thought and technology. We should have the courage to use reason, actively explore the unknown, and reflect on moral principles, thereby guiding technological progress to serve the well-being of all humankind.

REFERENCES

Ava Thomas Wright. "Rightful Machines." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022, 223–238.

Bateson, Gregory. *Steps to an Ecology of Mind,* University of Chicago Press, 2000.

Baiasu, Sorin. "The Challenge of (Self-) Consciousness: Kant, Artificial Intelligence and Sense-Making." in *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022.

Berger, Larissa. "On the Subjective, Beauty and Artificial Intelligence: A Kantian Approach." in *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, *In De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022.

Birhane, Abeba. and Jelle van Dijk. "Robot rights? Let's talk about human welfare instead." In AIES '20: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society. Association for Computing Machinery. 2020. p. 207-213. https://doi.org/10.1145/3375627.3375855.

Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford : Oxford University Press, 2013. https://archive.org/details/superintelligenc00unse/page/48/mode/2up?q=Vernor+Vinge

Dierksmeier, Claus. "Partners, Not Parts. Enhanced Autonomy Through Artificial Intelligence? A Kantian Perspective." in *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022.

Danaher, John. "Why AI Doomsayers are Like Sceptical Theists and Why it Matters." *Minds & Machines* 25, 2015, 231–246. https://doi.org/10.1007/s11023-015-9365-y.

Evans, Richard. "The Apperception Engine." in *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022.

Gill, Karamjit S. "Artifictional Intelligence: Against Humanity's Surrender to Computers." *AI & Society* 34, no. 2 (January 2, 2019): 391–92. https://doi.org/10.1007/s00146-018-0873-1.

Heidegger. Martin. *Being and time*. New York: Harper Publishers, 1962.

Hyeongjoo, Kim and Dieter Schönecker. *Kant and Artificial Intelligence*. Walter de Gruyter GmbH & Co KG, 2022. https://doi.org/10.1515/9783110706611.

Hyeongjoo, Kim. "Tracing the Origins of Artificial Intelligence: A Kantian Response to McCarthy's Call for Philosophical Help." *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022.

Kant, Immanuel. *Critique of Judgment*. Translated W.S. Pluhar. Hackett, 1987.

_____. *Critique of Practical Reason*. Cambridge University Press, 1997.

_____. *Critique of Pure Reason*. Cambridge University Press, 1998.

_____. *Critique of the Power of Judgment*. Cambridge University Press, 2000.

_____. *Groundwork for the Metaphysics of Morals* (A. W. Wood, Trans.). Yale University Press. 2002 (Original work published 1785).

_____. *Metaphysical Foundations of Natural Science*. Cambridge University Press, 2004

Lisa Benossi, and Sven Bernecker. "A Kantian Perspective on Robot Ethics." in *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022.

McIlroy-Young, Reid, Siddhartha Sen, Jon Kleinberg, and Ashton Anderson. "Aligning Superhuman AI With Human Behavior: Chess as a Model System." KDD '20: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, August 20, 2020, 1677–1687. https://doi.org/10.1145/3394486.3403219.

Schönecker, Dieter. "Kant's Argument From Moral Feelings: Why Practical Reason Cannot Be Artificial." in *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022.

Schlicht, Tobias. "Minds, Brains, and Deep Learning: The Development of Cognitive Science Through the Lens of Kant's Approach to Cognition". *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022.

Schmidt, Elke Elisabeth. "Kant on Trolleys and Autonomous Driving." in *Kant and Artificial Intelligence*, edited by Hyeongjoo Kim and Dieter Schönecker, In *De Gruyter eBooks*, Walter de Gruyter GmbH & Co KG, 2022.

Verdoux, Philippe. "Emerging Technologies and the Future of Philosophy." Metaphilosophy 42, no. 5 (October 1, 2011): 682–707. https://doi.org/10.1111/j.1467-9973.2011.01715.x.

Vinge, Vernor. "The Coming Technological Singularity: How to Survive in the Post-Human Era". In Vision-21: Interdisciplinary Science and Engineering in the Era of Cyberspace, 1993, 11-22. NASA Conference Publication 10129. NASA Lewis Research Center. https://archive.org/details/NASA_NTRS_Archive_19940022856/mode/2up.