

Mini Review Article

Seven Proposals for the Ethical Advancement of Artificial Intelligence and Humanity's Protection from its Potential Risks

Mohsen Moeini*

Modiran Pishro Research Company, Tehran, Iran

Corresponding Author: Mohsen Moeini, Modiran Pishro Research Company, Tehran, Iran.

Received: 📅 2024 Jul 31

Accepted: 📅 2024 Aug 19

Published: 📅 2024 Aug 30

Abstract

In this article, I propose measures for the positive growth and prevention of potential harms of artificial intelligence (AI) from a philosophical, aesthetic, and sociological perspective on a global scale. Drawing from my extensive background in various technology and AI fields, education, management, and research, I urge thinkers and specialists to reach a consensus on practical solutions to strengthen AI's positive aspects and prevent its risks.

Keywords: Artificial Intelligence, humanism, prohibitions, Ethical Advancement and sociological.

1. Introduction

The issues of AI and ethics, as well as the fear of AI's advancement which has mobilized politicians and legislators worldwide, are among the most critical topics of our time. The fear that Henry Kissinger once spoke of has now become widespread. In this article, I will briefly and practically address the root of this concern and its solution. We have reached a point where instead of excessive discussion, we must identify the roots of the problem and act to resolve it. Undoubtedly, AI is the full manifestation of Descartes' cogito. The thinking self, an entity that can perform actions based on neural networks and deep learning. However, the terrifying aspect of AI is akin to Nietzsche's Übermensch and the perfect human envisioned by Farabi and Avicenna. Ultimately, AI is the offspring of humanity's historical dream and the complete manifestation of reactionary humanism. Whether we like it or not, AI will continue to advance, and in an open society, AI will inevitably progress. The fear that AI might surpass humans and disrupt societal order like in Hollywood movies is precisely what alarms many. The inevitable progress of AI towards its ultimate development is bound to happen, but the concern that it might act against humanity and human societies is naturally worrisome. The fear of approaching a kind of AI that creates a dystopia, as depicted in literary works like George Orwell's 1984 or the works of Kazuo Ishiguro and Aldous Huxley, is valid.

As noted in the book *Character and Dystopia* by Aaron S. Rosenfield, humans are concerned: "The term dystopia in our time reflects our widespread anxiety, anger, and turmoil, which can be seen almost everywhere. We are exactly in the situation that Jill Lepore described in her 2017 *New Yorker*

article, 'The Golden Age of Dystopian Fiction.' In early 2017, when I was riding the New York subway, three people around me were reading 'The Handmaid's Tale,' '1984,' and 'It Can't Happen Here.' Artistic and literary dystopian works are no longer seen as remnants of a marginalized and depressed genre. They have suddenly become valuable works."

What Should Be Done to Address This Concern? How to Prevent the Potential and Ultimate Harms of AI?

Without any preamble, I will explicitly present my practical solutions in a brief manner.

- **Accepting the Issue:** We need to question our relationship with technology and, alongside this question, our relationship with existence, as Heidegger did in *Being and Time*. Just as the ancient Greeks questioned *technê* and *logos*, and AL Khwarizmi, the founder of algorithms, questioned his relationship with existence.
- **Defining Our Desires:** What do we want from AI, and what is the relationship between AI and existence, and between us and existence? This fundamental question leads to answers that result in a turn towards ethics. Here, the second question arises: "When we speak of ethics, which ethical system are we referring to?" The fascists of history also claimed to defend an ethical system. Did Hitler not proclaim ethical slogans at the peak of his crimes? Therefore, merely using the term 'ethics' without clarifying its concrete and practical meaning does not solve the problem.
- **Establishing Ethical Principles:** Based on the questions and answers above, ethical principles and prohibitions must be

technically designed into AI systems at every stage. This can be achieved during the neural network project phase with new coding and public oversight. Ethics in AI does not mean creating restrictions but defining clear distinctions between good and evil, which can easily be implemented in AI system design.

- **Free Information Exchange:** Establish transparency and free exchange of information among countries in the AI field, creating a glass room for activities in AI.
- **International Organization:** Form a powerful international organization comprising all countries to control, monitor, and guide AI activities, similar to the International Atomic Energy Agency.

- **Codified Laws:** Write codified laws based on our needs and relationship with AI, grounded in public ethics, and ensure all AI field participants adhere to these laws. This code should be developed by a group of specialists from various fields with the participation of all cultures.

- **Assistance to Weaker Countries:** Help weaker countries utilize the beneficial capabilities of AI and involve all countries and specialists worldwide in a collaborative effort for positive AI progress. This prevents something akin to a nuclear arms race or covert misuse of AI by less capable countries.

- **Regular Ethical Review:** Periodically review the ethical framework governing AI progress, repeatedly questioning the correctness of its advancement and revising principles as needed, considering that an open society is constantly changing and, according to the philosophy of science, is subject to falsification.

- **Cultural and Psychological Monitoring:** Monitor the cultural differences of countries and the psychological state of citizens and society by a team of sociologists and psychologists, and repeatedly review AI governance laws based on the feedback received.

- **Attention to Aesthetics:** Focus on aesthetics in AI products so that the aesthetic and ethical emotions of AI users can keep an inherent ethical system alive in them and not suppress human emotions. Keeping and strengthening the aesthetic drive of AI users surely forms a kind of human immunity in advancing society.

1.1. Practical Measures

By studying hundreds of theoretical books and examining artistic works, we realize that beauty and aesthetics, from Plato to Baumgarten, and from Baumgarten (the founder of the term aesthetics) to postmodern theorists, have all sought to revive a kind of freedom and liberation through creativity and the creation of the beautiful. A creative, beautiful, and pervasive matter. Thus, when in 2018 the portrait of Edmond de Belamy, produced by AI and signed by an algorithm, was sold at Christie's New York auction for \$432,000, it actually realized all those old theories through the collaborative efforts of the Obvious Group and AI. Therefore, what we see of AI's impact on aesthetics and the art market is indeed the

realization of thousands of years of theorizing by thinkers and artists. Inevitably, paying attention to aesthetics in creating AI products, alongside the growth of human intelligence, pushes AI towards the expansion of beauty, ethics, and human creativity. It also makes beauty pervasive, saving art and beauty from the constraints of time and place, and naturally providing a platform for the dialogue between art and science, a crucial matter often overlooked. As Dostoevsky says in *The Idiot*: "Beauty will save the world" [1-15].

2. Conclusion

Two points are essential considering the above introduction:

- **Time:** We must promptly implement these views in collaboration with specialists from humanities, arts, and technology to avoid a scenario like the chip war or a cold war between a few limited countries.

- **Balancing Progress and Regulation:** The above laws should not hinder the beneficial progress of AI, and radicals should not block the evolution of an open society with false pretenses.

Ultimately, reflecting on the two meanings of logos in ancient Greece teaches us that the combination of logic/language in any scientific activity can be enlightening. Let us not forget that AI is a continuation of technological leaps and, in my view, not merely a science but a pioneering human art. The ancients of Greece pondered the concept of technê. From Aristotle's perspective, the term technê refers to the third sphere of human existence, the realm of creation. Technê, which corresponds to "technique" in English and French, and "art" in Persian in its new and specific sense, according to Plato and Aristotle, refers to constructing, completing, and imitating nature with the power of imagination. Therefore, the essence and truth of technê involve imaginative imitation, and imaginative imitation is accompanied by imitation. From this perspective, in Plato and Aristotle's thought, the metaphysical aspect of technê and its essence is precisely this imitation and emulation of the real world and nature. A thinker engaged in fine arts deals with imaginations that imitate the real world in a way. We, too, need to combine imaginative imitation with AI, just as our forefathers did, and this is the beginning of contemplation. Contemplation to achieve utopia and distance from dystopia, or what Farabi describes in *The Opinions of the Virtuous City* as the land of the blessed.

References

1. Heidegger, M. (1927). *Being and Time*. Translated by John Macquarrie and Edward Robinson. *First English Edition*.
2. Grube, G. M. A. (1992). *Republic (Grube Edition)*. Hackett Publishing.
3. Aristotle. *Nicomachean Ethics*. Translated by W.D. Ross, Oxford University Press, 2009.
4. Descartes, R. (1993). *Meditations on first philosophy*, trans. Donald A. Cress (*Indianapolis, Ind.: Hackett*).
5. Nietzsche, F. (1978). *Thus Spoke Zarathustra*, trans. Walter Kaufmann. *The Portable Nietzsche*, 429-436.
6. Ibn Sina, A. A. A. (2005). *The metaphysics of the healing*

- translated by Micheal E. Marmura.
7. Al-Farabi, A. N. (1985). *Mabadi'Ara'Ahl al-Madinah al-Fadhilah*, ed. Richard Walzer.
 8. Orwell, G. George Orwell bibliography.
 9. Ishiguro, K. (2008). *Conversations with Kazuo Ishiguro*. Univ. Press of Mississippi.
 10. Huxley, A. (2021). *Brave new world and brave new world revisited*. Wildside Press LLC.
 11. Lepore, J. (2017). A golden age for dystopian fiction. *The New Yorker*, 5, 102-106.
 12. Dostoyevsky, F. (1927). *The Idiot: A novel in four parts* (Vol. 2). William Heinemann.
 13. Baumgarten, Alexander. *Aesthetica*. Olms, 2007.
 14. Christie's New York. "Portrait of Edmond de Belamy." 2018.
 15. New York Public Library. "Annual Report 2018."