

Higher Expectations of Artificial Intelligence in Terms of Morality and Humanity

Junzhe Zhou *

Whitman College, 280 Boyer Ave, Walla Walla, WA, US, 99362

* Corresponding Author Email: zhouj@whitman.edu

Abstract. This article delves into the critical question of whether Artificial Intelligence (AI) should adhere to higher standards of morality and humanity. As AI systems infiltrate our everyday lives, their influence surfaces profound ethical and moral conundrums, primarily due to their intrinsic limitation in exercising moral judgments. The ethically sound deployment of AI requires a comprehensive understanding of these moral challenges and the construction of suitable safeguards. The article further scrutinizes the repercussions of AI on our conception of human nature, highlighting potential modifications in human relationships and the essence of our humanity. It contends that AI systems, given their decision-making prowess, should be governed by robust moral principles and human values to curtail harm and endorse fairness and justice. The piece accentuates the pivotal role of human guidance in AI development and application, asserting that these technologies should align with our values. It concludes with an emphasis on the need for continuous surveillance to ensure that AI technologies benefit society, underlining the importance of aligning AI with our moral and human expectations for sustainable societal progression.

Keywords: Artificial Intelligence (AI), Ethics, Machine Learning, Privacy, Autonomous Systems.

1. First Section

1.1. A Subsection Sample

Today, Artificial Intelligence (AI), a pioneering innovation, is no longer just a concept or a blueprint, but has the immense potential to fundamentally change the technological landscape of the 21st century. The field of AI as it has evolved today encompasses perception, reasoning, learning, and decision-making, all of which are realized through the integration of various sub-disciplines.

Artificial Intelligence is recognized and celebrated for its unparalleled potential to revolutionize industries and improve efficiency in numerous fields. However, along with the immense power comes an equally significant challenge, which stems mainly from the breadth and depth of intelligence possessed by AI systems. As well as whether the level of human understanding of AI can keep pace with its development.

The broad categorization of AI revolves around three main categories, which are narrow AI, general AI, and superintelligence [1]. First, narrow AI is often referred to as weak AI and is specialized for specific tasks, such as speech recognition or automatic responses. Although narrow AI is extremely efficient at performing specified tasks, it is fundamentally limited by its lack of awareness and understanding.

In contrast, generalized artificial intelligence (AGI) or strong AI theoretically mirrors human cognitive abilities. It grasps context and can apply knowledge learned in one domain to another. However, as of 2021, AGI remains a concept largely confined to research papers and theoretical discussions.

Superintelligence is more of a speculation as a future development of AI that can outperform human intelligence in every conceivable cognitive task. It includes machines that have the potential to outperform humans in virtually every economically rewarding endeavor. While there is excitement about superintelligence, we also need to remain vigilant because of the associated existential risks it poses to humanity [1].

2. Classification of AI

Today, AI is manifesting itself in the real world in a variety of breathtaking ways, across a range of different fields. In healthcare, AI heralds a new era of precision diagnosis and customized medicine. Advanced intelligent learning models are carefully designed to autonomously analyze information and examine medical images to diagnose specific disease types. Additionally, AI is anticipated to have a significant impact on the field of customized medicine, for instance by better tailoring medicines to the unique characteristics of each patient.

The importance of artificial intelligence in the financial sector is also starting to be recognized in practice, playing a key role in tasks such as fraudulent activity detection and algorithmic trade execution. The ability of AI to recognize patterns and detect anomalies from massive amounts of data helps to successfully identify fraud. On the flip side, AI can maximize profits by drawing on and analyzing information from real-time assessments of market conditions to help execute trades and give advice on decision-making.

When considering transportation, self-driving cars are a prominent testament to the current applications of AI and a major topic of interest. These self-navigating transportation vehicles use AI to analyze different dimensions of sensory data in order to traverse the terrain with a high degree of safety and efficiency.

However, given the proliferation of AI and its potential to catalyze change, a proposition emerges: should we impose higher ethical and humanistic standards on AI? As AI gradually permeates all aspects of our existence and assumes increasingly important decision-making responsibilities, it becomes crucial to establish an ethical framework to guide its development and utilization. The urgency of this discussion is heightened in light of the increasing penetration of AI into areas such as social care and jurisprudence.

As we discuss the possibilities for the future of Artificial Intelligence, two fundamentally different routes have drawn attention: AI development with limits and AI development without limits. These two paths reflect, in part, how we view and respond to the challenges and opportunities presented by AI. Restricted AI development emphasizes accountability and caution for the technology to prevent AI applications from having an unavoidable negative impact on human society. Unrestricted AI development, on the other hand, emphasizes creativity and openness to support the technology's quick development and wider adoption.

Comparing these two development paths, we can better understand the possibilities and limitations of AI and can develop appropriate response strategies for different situations. Both restricted and unrestricted AI development have advantages and disadvantages, but both must deal with the same challenge: balancing the potential risks and advantages of AI development while also ensuring that it is consistent with our moral and ethical principles. We will delve deeper into the specifics and potential ramifications of these two approaches in the debate that follows.

3. Unconstrained AI development

3.1. Possible limitations of AI

The trajectory of unrestricted Artificial Intelligence (AI) development holds vast implications, ranging from economic productivity to social ethics. As it stands, several impediments curtail AI's potential, among which are data limitations, algorithmic bias, and lack of explainability [2].

AI's effectiveness hinges on the breadth and quality of available data. Lack of sufficient data hampers algorithmic learning and undermines prediction accuracy [3]. Additionally, existing biases in datasets can inadvertently perpetuate discriminatory practices, skewing the AI's decision-making capabilities [4]. Lastly, AI's lack of explainability is a pressing concern. Despite the high performance of certain AI models, their 'black box' nature presents transparency issues, impacting user trust and adoption [5].

AI also grapples with ethical considerations, particularly privacy, security, and potential job displacement. Data privacy is an undeniable concern as AI typically necessitates extensive data collection and analysis, potentially compromising individual privacy [6]. This is because if AI centralizes private information, then security risks will follow. Furthermore, with AI expected to automate numerous tasks, concerns about job displacement and resulting socio-economic inequalities loom large [7].

3.2. Impact of removing restrictions on AI

If these restrictions are lifted, unrestricted AI development could wield profound impacts. On the plus side, unrestricted AI could augment decision-making capabilities, improve efficiency, and catalyze innovation [8]. AI could sift through massive datasets, glean insightful patterns, and contribute to informed decision-making across sectors. By automating routine operations, it might increase productivity and free up human resources for more difficult and imaginative jobs. Additionally, unconstrained AI may produce ground-breaking discoveries that open up previously unthinkable possibilities.

Nevertheless, the moral and ethical implications of unrestricted AI are equally significant. Where responsibility lies and who is accountable in the context of AI is a complex and controversial issue. Should AI systems cause harm, attributing responsibility can be intricate given their autonomous nature. Furthermore, the potential impact of unrestricted AI on human dignity is profound. The likelihood of AI systems replacing humans in various spheres could erode human self-worth and autonomy, fostering existential anxiety [8].

Unrestricted AI also raises apprehensions around autonomous weapons, deep fakes, and privacy infringement. Autonomous weapons, guided by AI, could alter the dynamics of warfare, escalating risks, and intensifying ethical quandaries [9]. Deep fakes, powered by AI, could manipulate audiovisual content, casting shadows over truth and fostering misinformation [10]. Lastly, unrestricted AI development, unless appropriately regulated, could potentially lead to invasive surveillance and significant infringements on personal privacy [6].

In conclusion, while the notion of unrestricted AI development is tantalizing, bearing the promise of unprecedented advancements, it also carries with it significant ethical considerations that demand careful contemplation.

4. Restrained AI development

4.1. Limitations that exist in AI today

The path of restrained AI development embodies a delicate balance: the simultaneous pursuit of technological advancement and the mitigation of corresponding risks. The limitations currently imposed on AI are multifaceted, encompassing privacy laws, data protection statutes, and ethical norms [11].

Legal protections for individual privacy serve as a crucial defense against potential abuses of AI in data harvesting, safeguarding the rights of individuals from invasive technologies [12]. The General Data Protection Regulation (GDPR) in the European Union and other data protection laws specify the circumstances under which AI systems may collect, store, and use data, protecting user privacy and rights [13]. Additionally, ethical guidelines for AI strive to ensure that its evolution remains aligned with humanistic values and societal norms [14].

These restrictions play a pivotal role in confronting inherent challenges in AI, including issues related to bias, fairness, and transparency. Unconscious bias in AI algorithms, often originating from prejudiced training data, can perpetuate discrimination. This underscores the importance of balanced data handling regulations to promote equity and fairness [4]. Ensuring fairness in AI, a task partially facilitated by regulatory measures, is critical to prevent exacerbation of societal inequalities [15]. Additionally, AI's 'black box' nature, often cited as a major barrier to transparency, makes explainability indispensable for AI systems to cultivate trust and comply with ethical standards [2].

Both collaborative industrial efforts and regulatory bodies play essential roles in enforcing these AI constraints. Industry standards, mutual research initiatives, and open dialogues encourage a cooperative stance to tackle AI's ethical and technical quandaries [16]. Meanwhile, regulatory bodies provide the structural foundation necessary for the practical implementation and enforcement of these norms and statutes [11].

4.2. Restrictions after high AI development

As AI technology continues to evolve, growing in complexity and capacity, it becomes imperative to have strategies for updating and amending existing restrictions. Striking a balance between preserving societal interests and fostering innovation presents a complex task, yet it remains a necessary endeavor.

This dynamic landscape calls for an interdisciplinary approach to conceptualize and implement updated restrictions. Policymakers, ethicists, and technologists need to work in collaboration to navigate the intricate challenges posed by AI, integrating diverse perspectives to devise robust, forward-looking policies [16].

In conclusion, a measured approach to utilizing AI's capabilities while assiduously limiting its possible negative effects is indicated by the trajectory of restricted AI development. This route, albeit complex, provides a pragmatic approach to fostering AI development that remains aligned with human values and promotes societal well-being.

5. AI and Ethics

The pursuit of ethically robust AI systems presents a plethora of challenges. A prominent concern is that AI has a huge deficit in executing judgment. When humans make decisions, they refer to factual factors while taking into account ethical guidelines, and it so happens that this part of the ethical guidelines is crucial. Consequently, an AI-operated system, bereft of appropriate safeguards or ethical standards, could inadvertently inflict harm [17].

The unconstrained evolution of AI has the potential to intensify these moral quandaries. The dearth of explicit responsibility in AI-mediated decisions can engender moral conflicts. The attribution of blame in the aftermath of a harmful decision by an AI system – should it be allocated to the developer, the user, or the AI system itself? [18]. The obscured boundaries of moral judgment in AI pose a formidable obstacle in the ethical deployment of these technologies [19].

Probing this conundrum further, an intriguing debate surfaces around the role of human intervention in mediating these AI-related concerns. On the one hand, a persuasive argument can be made for humans to adopt a stewardship role. This entails carefully planning AI systems with built-in ethical frameworks and making sure that strict regulatory monitoring governs the creation and application of AI. On the other hand, such a human-centric control should not inhibit the potential benefits that AI systems can deliver, given their immense capabilities in processing voluminous data and making complex decisions in a fraction of the time humans would require.

5.1. AI's impact on human nature

Unrestricted AI development also holds implications for our understanding and appreciation of human nature. AI systems, particularly those designed to mimic human behavior, may change the dynamics of human relationships. With sophisticated AI, human interactions could be reduced or even replaced, leading to potential emotional deficits in our society [20].

Furthermore, AI's inability to experience genuine emotions or empathy may pose a challenge to our understanding of these quintessentially human traits. Could a machine ever genuinely understand human emotions if it does not possess feelings itself?[21]. This raises profound questions about the essence of being human and the unique values we bring to our society and relationships.

5.2. The reason for higher moral and human expectations in AI

Given these concerns, it is pertinent to hold AI to higher moral and human expectations. As bearers of intelligence and decision-making capabilities, AI systems ought to be guided by moral principles and human values [22]. And it's not just about the simple process of programming an AI to perform a task, it's also about making sure that these workflows and operational logics stay within our moral and ethical boundaries.

The responsibility falls on us, humans, to ensure that AI technologies respect and uphold our values [23]. As creators and promoters of these advanced technologies, it is our responsibility to ensure that they are used to benefit the development of human civilization and to promote equity and justice.

Exploring the impacts of AI on society and its sustainable development underscores the importance of aligning AI with our moral and human expectations. As AI continues to transform various aspects of our lives, we must remain vigilant in ensuring that these technologies serve us, not the other way around [24].

6. Conclusion

The accelerated progression of Artificial Intelligence (AI) has ignited pivotal debates concerning its ethical ramifications and the subsequent influence on our interpretation of humanity. The ethical quandaries emerging from this debate are multifold and intricate, encompassing AI's inherent incapacity for moral decision-making and its potential to modify human interactions and values. These ethical dilemmas extend beyond theoretical discourse, invoking substantial practical implications that accentuate the imperative for an ethically grounded strategy in AI evolution.

The proposition that AI should comply with elevated moral and human expectations underscores the indispensable role humans must assume in orchestrating AI's growth. As progenitors of AI, humans bear the responsibility to integrate ethical principles within AI systems and ensure rigorous regulatory supervision. Our AI creations should be reflective of societal norms and values, operating within an accepted moral framework.

Expectations for AI should transcend mere functionality, with an emphasis on equity, transparency, and accountability. Our creations should serve us and mirror our values, not vice versa. We are vested with the power to steer AI's trajectory, signifying the profound implications of our decisions today.

In summation, despite the complexity of the ethical dilemmas enveloping AI, these challenges offer an opportunity to mold a future wherein AI is not merely intelligent but also ethically robust. The task is arduous and necessitates multi-disciplinary collaboration. Nevertheless, the stakes are high, making it imperative that we succeed in ensuring AI adheres to stringent moral and human standards, thereby contributing positively to societal well-being and sustainable development.

References

- [1] Bostrom, N. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press. 2014.
- [2] Doshi-Velez, F., & Kim, B. Towards a rigorous science of interpretable machine learning. *arXiv.org*. <https://arxiv.org/abs/1702.08608> 2017.
- [3] Shalev-Shwartz, S., & Ben-David, S. *Understanding machine learning: From theory to algorithms*. Cambridge University Press. 2022.
- [4] Barocas, S., & Selbst, A. D. (n.d.). Big Data's disparate impact - JSTOR.
- [5] Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature News*. <https://www.nature.com/articles/s42256-019-0048-x> 2019.
- [6] Taylor, L. Group privacy: new challenges of data technologies. SPRINGER. 2019.
- [7] Chui, M., Manyika, J., & Miremadi, M. Where machines could replace humans-and where they can't yet. McKinsey & Company. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights%20/where-machines-could-replace-humans-and-where-they-cant-yet> 2016.

- [8] Brynjolfsson, E., & McAfee, A. (n.d.). Apa PsycNet. American Psychological Association. <https://psycnet.apa.org/record/2014-07087-000>.
- [9] CNAS autonomous weapons and operational risk - 160219. (n.d.). https://www.stopkillerrobots.org/wp-content/uploads/2021/09/CNAS_Autonomous-weapons-operational-risk.pdf.
- [10] Citron, D. K., & Chesney, R. (n.d.). Deep Fakes: A looming challenge for privacy, democracy, and national security. Scholarly Commons at Boston University School of Law. https://scholarship.law.bu.edu/faculty_scholarship/640/.
- [11] BUITEN, M. C. Towards intelligent regulation of artificial intelligence: European Journal of Risk Regulation. Cambridge Core. <https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/towards-intelligent-regulation-of-artificial-intelligence/AF1AD1940B70DB88D2B24202EE933F1B> 2019, April 29.
- [12] Schwartz, P. M., & Solove, D. J. The PII problem: Privacy and a new concept of personally identifiable information. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1909366 2011.
- [13] AXEL., V., PAUL. VON DEM BUSSCHE, EU General Data Protection Regulation GDPR: A practical guide. SPRINGER INTERNATIONAL PU. 2018.
- [14] Jobin, A., Ienca, M., & Vayena, E. The global landscape of AI ethics guidelines. Nature News. <https://www.nature.com/articles/s42256-019-0088-2> 2019.