

THE ETHICAL IMPLICATIONS OF ARTIFICIAL INTELLIGENCE (AI)

Rev. Fr. Emmanuel C. Umeh, PhD

Seat of Wisdom Seminary, Philosophy Campus, Umuahia,
Nigeria.

Email: dr_umehemmanuel@yahoo.de, Tel: +234 8133714051
&

Kalu Ikechukwu Kalu, PhD

Chair of Environmental Philosophy and Ethics,
Kiel University, Germany.

Email: ikechukwu@philsem.uni-kiel.de

Abstract

Artificial Intelligence (AI) is no longer just a concept in science fiction (sci-fi) movies or a topic for academics; it has embedded itself into daily human lives. From facial recognition unlocking phones to social media feeds curated just for us, AI is working behind the scenes. But this quiet revolution has not come without unintended consequences. Hence, this paper holds the view that as AI systems become increasingly capable and continue to be integrated into people's daily lives, it is pertinent to consider their ethical implications. What happens when machines start making decisions that used to require human judgment? Can AI algorithms be trusted to be fair when human beings cannot even agree on what fairness looks like? And who takes the blame when things go wrong? These are not just questions for experts; they are real issues that need answers as AI continues to advance. They are real issues that are more pressing with each little step taken in the advancement and integration of Artificial Intelligence. This paper employs the philosophical methods of analysis and reflection to address the stated questions. It is of the view that AI is useful but needs a thorough ethical evaluation to minimise the risks and maximise the opportunities involved with its applications.

Key Words: Artificial, intelligence, narrow AI, generative AI, algorithm, ChatGPT, Midjourney and Platforms

Introduction

The convergence of artificial intelligence and ethics has ignited a vital discourse, underscoring the need for a nuanced examination of the far-reaching implications of Artificial Intelligence (AI). Philosophy, as the foundational discipline that has historically nurtured and informed various fields of human inquiry, is uniquely positioned to provide a clarifying framework for understanding the complex interplay between AI, human values, and societal norms. As the influence of AI expands globally, affecting every facet of human existence, it is imperative to subject its development and deployment to a rigorous moral scrutiny, discerning the boundaries between right and wrong, good and bad, and just and unjust.

Throughout history, there have been recurring efforts to illuminate the intersection of technology and morality, as well as to establish frameworks for assessing the consequences of these advancements. While acknowledging that technological creations, including artificial intelligence, stem from human ingenuity and thus possess inherent value, one is compelled to confront the complex ethical dimensions of these innovations and their impact on humanity.

This paper scrutinises the ethical implications of AI, commencing with an exploration of the concept itself, followed by an examination of the imperative for integrating ethics into AI development. Subsequently, it delves into the major ethical concerns surrounding AI, endeavouring to address these challenges and culminating in a conclusive synthesis.

What is Artificial Intelligence?

Defining artificial intelligence is a daunting task, as it is a concept that resists a single and unified definition. This is not surprising, given that language allows for multiple expressions of concept, and AI is no exception. Experts from various fields, like AI research, technology, teaching, or entrepreneurship, have not converged on a universally accepted definition. This lack of consensus speaks to

the multifaceted nature of AI, thereby reflecting its revolutionary and far-reaching impact, which in turn has spawned a diverse lexicon of artificial intelligence.

The term “artificial intelligence” is a deliberate construct, born from meticulous reasoning and foresight. It encapsulates the initial vision, potential, design, and constraints of this ground-breaking technology. Coined by Stanford professor John McCarthy in 1955, the term was crafted to describe a field that merges science and engineering to create intelligent machines. McCarthy, often hailed as a pioneer in the field, defined AI as “the science and engineering of making intelligent machines, especially intelligent computer programs” (McCarthy, 2007), laying the foundation for a new era of innovation and technological determinism. This concept is built around two pivotal words: “artificial” and “intelligence”.

The term “artificial” immediately conveys that the subject is non-human, crafted by human ingenuity, emphasising the technical and engineering prowess that goes into creating these intelligent systems. By placing “artificial” before “intelligence”, the term itself suggests that understanding the “artificial” aspect is key to grasping the nature of these machines – it is the tangible, man-made component that shapes human perception of their intelligence.

The term “intelligence” in the context of artificial systems is indeed complex. Traditionally, intelligence is seen as a hallmark of human cognition, making its application to machines somewhat nuanced. The emergence of AI has expanded the understanding of intelligence, allowing it to encompass not only human beings but also non-human entities such as animals, plants, and machines. In this broader sense, intelligence refers to a distinctive quality that sets these machines apart from others – a sophisticated capacity that mirrors human intelligence. Through machine learning (ML) and deep learning (DL), AI systems can process vast amounts of data,

make autonomous decisions, and exhibit intelligent behaviour. This intelligence can be understood as a structural property, where the parts of the system and interactions yield smart outcomes, regardless of its composition. In practical terms, AI intelligence involves reasoning, learning, analysis, judgment, and inference – capabilities that enable these systems to mimic human-like decision-making.

Artificial intelligence, according to the Oxford Advanced Learners Dictionary, is “an area of study concerned with making computers copy intelligent human behaviour” (Hornby, 2018). More broadly, AI encompasses computer systems designed to mimic aspects of human intelligence, including learning, problem-solving, perception, and communication. Today, the most common form of AI is *narrow AI*, tailored for specific tasks. These systems are not sentient or self-aware, lacking human-like cognition. Still, they are incredibly effective at recognising patterns and making predictions – often faster and more accurately than human beings.

A more advanced subset of AI, *generative AI*, has recently gained prominence, capable of creating content such as text, images, and music based on user prompts. However impressive, these outputs stem from pattern recognition across vast datasets, rather than genuine creativity or comprehension. The apparent intelligence can be misleading, as AI systems do not truly “understand” in the human sense – they lack cognition, meaning, and empathy, operating solely on statistical correlations. This distinction is crucial, as AI’s lack of true understanding can have significant ethical implications in certain contexts. While AI can replicate patterns from its training data, it cannot question, criticise, or apply moral judgment. This limitation creates an ethical blind spot, highlighting the need for careful consideration and oversight in AI development and deployment.

The Need for Ethics in Artificial Intelligence

The growing presence of AI in daily human life underscores its role as a social force, extending beyond mere technical tools. These systems significantly influence decisions in hiring, medical diagnoses, and criminal sentencing – areas that inherently require human empathy and judgment. AI systems operate without conscience, lacking understanding of the societal context and capacity for moral reflection. They rely on statistical patterns in data, which may be flawed or biased, to inform decisions impacting human lives. This absence of ethical awareness poses a significant moral risk that extends beyond technical limitations.

Ethics is a non-negotiable imperative in AI development, ensuring that technology aligns with societal values, prevents harm, and promotes accountability. The absence of ethics risks creating systems perceived as unjust or dangerous, thereby eroding public trust. As AI increasingly shapes the human world, the questions of responsibility, autonomy, and decision-making must be prioritised, examining not just the capabilities of AI but its alignment with human values and the greater good.

Major ethical concerns in artificial intelligence

The deployment of AI in critical domains such as healthcare, finance, recruitment, and justice demands scrutiny, as these applications involve consequential decision-making. Some of the major ethical concerns in the application of AI are as follows:

1. The issues of bias and fairness
2. Privacy and surveillance
3. Autonomy and accountability
4. Job displacement and economic inequality
5. Artificial intelligence in warfare and lethal autonomous weapons (laws)
6. Artificial intelligence and human rights
7. Environmental impact

1. The issues of bias and fairness

Algorithmic bias is a pressing concern, as AI systems learn from historical data that is often marred by inequality, discrimination, and prejudice. This can lead to AI perpetuating systemic biases, masquerading as objective truth. Essentially, the fairness of AI is tied to the quality and biases of its training data, risking the reinforcement of existing social injustices. For example, hiring Amazon's algorithm starkly illustrates an algorithmic bias. Trained on male-dominated industry data, the AI system learned to disadvantage resumes with feminine indicators, like "women's chess club", perpetuating historical biases and limiting opportunities for women in tech (Vincent, 2018). These AI systems did not create bias; instead, they clearly reflected and amplified existing human prejudices. The COMPAS algorithm's disparity in labelling Black defendants as high-risk, despite similar records to white defendants, exemplifies how biased training data can lead to unethical outcomes, hidden behind a mask of objectivity.

2. Privacy and surveillance

AI's voracious data appetite raises significant privacy concerns, as systems rely on vast amounts of personal data, often collected without explicit consent. This has led to the commodification of human activity, fuelling surveillance capitalism and blurring the lines between private autonomy and corporate interests.

The Cambridge Analytica scandal seriously illustrates the risks of opaque AI-driven analytics, where private data is exploited to manipulate emotions and behaviours, eroding democratic norms. By harvesting Facebook data without consent, *Cambridge Analytica* created psychographic models to deliver targeted ads, influencing the 2016 US election and Brexit referendum. This highlights the dangers of unchecked data science and AI, threatening electoral transparency and fairness (Cadwalladr & Graham-Harrison, 2018).

In response to such threats, the *General Data Protection Regulation* (GDPR) in the European Union set a precedent with its emphasis on user consent, data minimisation, and the right to explanation, aiming to mitigate the risks of AI-driven data exploitation. However, inconsistent enforcement and varying global regulations create challenges, leaving many users vulnerable to data misuse and manipulation.

3. Autonomy and accountability

The opacity of AI systems, particularly those utilising deep learning and neural networks, poses significant concerns in high-stakes domains such as finance, healthcare, and justice. These "black boxes" produce outputs without transparent explanations, leaving individuals without recourse to understand or challenge decisions impacting their lives. This lack of accountability undermines fairness and erodes trust in institutions deploying these systems.

In medicine, AI-driven misdiagnoses can be catastrophic if clinicians cannot understand the reasoning or detect errors. Likewise, in criminal justice, opaque AI outcomes can perpetuate biases and exacerbate systemic discrimination, especially against marginalised groups, highlighting the need for transparency and explainability in high-stakes AI applications

4. Job displacement and economic inequality

The rapid advancements of AI are revolutionising labour markets, extending automation into domains previously considered uniquely human, such as creativity, language, and emotional intelligence. AI systems now compose music, create art, write articles, and develop software at unprecedented scales and speeds, transforming industries and redefining the future of work.

Generative AI tools like ChatGPT, Midjourney, and DALL-E are revolutionising industries, but also pose significant job displacement risks for professionals in journalism, marketing,

education, and tech. As AI handles tasks at unprecedented scales, creative and knowledge-based sectors face unprecedented disruption, threatening millions of jobs.

The warning of the World Economic Forum that up to 85 million jobs could be displaced globally by 2025 highlights the urgent need for proactive measures, as low- and mid-skill roles face significant vulnerability to AI-driven automation (World Economic Forum, 2020). The AI-driven disruption exacerbates existing inequalities, as those with digital savvy and capital are poised to thrive, while vulnerable populations in under-resourced regions, with precarious jobs, risk being left behind, thereby widening the economic divide.

5. Artificial intelligence in warfare and lethal autonomous weapons (LAWS)

The integration of AI into military systems sparks concerns about the future of warfare. The introduction of Lethal Autonomous Weapon Systems (LAWS), such as autonomous drones and unmanned vehicles, raises significant ethical, legal, and security concerns, blurs lines and sparks debate about human oversight in life-or-death decisions. The removal of human control from life-and-death decision-making is a major concern with LAWS. Algorithms lack objective understanding, emotional nuance, and moral judgment, sparking fears of malfunction, mistaken identity, or misuse, outweighing the potential benefits of precision and reduced casualties. Autonomous weapons operating at machine speeds risk triggering rapid escalation, leaving little room for diplomacy. The AI arms race also raises proliferation concerns, as advanced technology could fall into an unauthorised node on a network, sparking fears of devastating misuse and destabilisation.

6. Artificial intelligence and human rights

AI-generated deepfakes blur reality, threatening privacy, free speech, and dignity. Malicious use for revenge porn, disinformation, and political manipulation raises urgent concerns,

as ordinary citizens and institutions struggle to cope with this potent tool for deception and harm. Deepfakes have become a serious threat, particularly for women, who are often the target of non-consensual pornography. This fabricated content undermines their dignity and can have severe consequences. Politically, deepfakes are being used to manipulate public opinion, spread disinformation, and delegitimise opponents. The 2024 U.S. elections saw fake audio recordings falsely attributed to candidates, causing voter confusion. Similarly, in Pakistan, fabricated videos sparked hostility against human rights activists (Oxford Internet Institute, 2025).

Deepfakes pollute the information ecosystem, eroding trust in media and undermining democratic participation. This fuels a liar's dividend, where real evidence is dismissed as fake, and truth becomes increasingly indistinguishable from fiction, threatening the very fabric of informed decision-making.

7. Environmental impact

The environmental impact of AI is staggering. Training large models like GPT-3 or Palm requires massive energy, emitting over 550 metric tons of CO₂ per model, equivalent to the annual emissions of 125 vehicles. This highlights AI's hidden carbon footprint and growing ecological toll (Saenko, 2023).

The environmental strain of AI is further exacerbated by data centres' massive energy demands, which consume 1.5% of global electricity and are projected to double by 2026, highlighting the urgent need for sustainable AI practices and eco-friendly infrastructure (International Energy Agency, 2025). The rapid growth of AI will exponentially increase energy consumption, strain global resources, and emphasise the need for sustainable development and eco-conscious innovation.

Addressing the ethical concerns of artificial intelligence

Addressing the ethical challenges of AI is an ongoing process. Algorithmic bias is a significant concern which requires diversified datasets reflecting human experiences across ethnicity, gender, class, and geography. Rigorous auditing and continuous monitoring can help detect and correct bias, ensuring AI systems are fair and equitable.

Transparency is of very crucial importance to AI trustworthiness, particularly in high-stakes applications. Explainable AI models can foster accountability and understanding, allowing users to have a grasp of decision-making processes and ensuring informed decisions which impact lives are transparent and justifiable. The erosion of privacy by AI raises concerns. Unchecked data collection and commodification without consent reflect surveillance capitalism's dark side. Regulatory frameworks such as the General Data Protection Regulation (GDPR) of the European Union are steps forward, but internationally harmonised standards prioritising user agency and data dignity are needed to safeguard individual rights and ensure responsible AI development.

Automation's expanding reach sparks economic justice, which is worrisome. As AI displaces jobs, proactive workforce planning is crucial. Governments and industries must invest in upskilling and reskilling programs, focusing on uniquely human skills like empathy, critical thinking, and ethical reasoning, to equip workers for an AI-driven economy and mitigate inequality.

The militarisation of Artificial Intelligence poses an even graver set of ethical concerns. The advent of lethal autonomous weapons (LAWs) signals a disturbing shift in warfare, stripping human judgment from life-and-death decisions. This raises profound moral and legal concerns, as accountability becomes murky. A global treaty, akin to the chemical weapons ban, is urgently needed to curb

the development and deployment of lethal autonomous weapons, preventing a potentially calamitous future where machines decide who lives or dies (Human Rights Watch, 2023).

AI-fabricated media is warping reality, destabilising societies, and eroding trust. Deepfakes and synthetic media threaten civic discourse, making fact-checking a Herculean task. To combat this, we need a multi-pronged approach: robust verification tools, digital literacy education, and laws targeting malicious actors. Platforms must also step up, authenticating content and flagging manipulated media to prevent the truth from becoming the casualty.

The environmental toll of AI is undeniable, with massive computational demands fuelling carbon emissions. To mitigate this, we must prioritise eco-friendly AI development: invest in energy-efficient technology, optimise algorithms, and transition to renewable energy. Ironically, AI can be a potent tool in this fight, optimising energy grids, modelling climate scenarios, and streamlining resource management to help heal the planet (Department of Energy, 2024). Environmental stewardship must be the guiding ethos of AI, baked into every stage, from design to deployment. Sustainability should not be an add-on, but a core principle driving innovation, ensuring AI's benefits do not come at the expense of the planet.

Conclusion

Artificial intelligence is revolutionising industries and daily life, but its rapid growth raises red flags. Without conscience or moral reflection, AI systems can perpetuate biases, compromise privacy, displace jobs, and make unaccountable decisions. These concerns underscore the need for careful moral scrutiny and responsible development to ensure that the benefits of AI are equitably distributed and its risks are mitigated. Embedding ethics into AI development is crucial for its positive impact. Collaboration among

stakeholders is key to creating transparent, fair, and inclusive AI systems. The goal is to harness the potential of AI while prioritising human values, ensuring its future is shaped by moral responsibility, not just technological prowess.

References

- Bankins, Sarah, and Paul Formosa. "*The Ethical Implications of Artificial Intelligence (AI) for Meaningful Work.*" *Journal of Business Ethics* 185 (2023): 725–740. <https://link.springer.com/article/10.1007/s10551-023-05339-7>.
- Cadwalladr, C., & Graham-Harrison, E. 2018, "Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach," *The Guardian*, March 17, 2018, <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>.
- Capitol Technology University. "*The Ethical Considerations of Artificial Intelligence.*" Capitol Technology University, May 30, 2023. <https://www.captechu.edu/blog/ethical-considerations-of-artificial-intelligence>.
- Department of Energy, "*AI and Energy*," U.S. Department of Energy, last modified July 2024, <https://www.energy.gov/cet/articles/ai-energy>.
- Hornby, A. S. 2018. *Oxford Advanced Learner's Dictionary of Current English* (9th Edition). Oxford University Press, Oxford.
- Human Rights Watch, "*Killer Robots: UN Vote Should Spur Action on Treaty*," December 22, 2023, <https://www.hrw.org/news/2024/01/03/killer-robots-un-vote-should-spur-action-treaty>.
- International Energy Agency (IEA), "*AI Is Set to Drive Surging Electricity Demand from Data Centres While Offering the Potential to Transform How the Energy Sector Works*," IEA, April 2025, <https://www.iea.org/news/ai-is-set-to-drive-surging-electricity-demand-from-data-centres-while-offering-the-potential-to-transform-how-the-energy-sector-works.IEA>
- McCarthy, J. 2007, *WHAT IS ARTIFICIAL INTELLIGENCE?* <https://www-formal.stanford.edu/jmc/whatisai.pdf>
- Microsoft. "*Ethical Implications of Artificial Intelligence.*" Microsoft 365 Life Hacks. Accessed May 12, 2025.

- <https://www.microsoft.com/en-us/microsoft-365-life-hacks/writing/ethical-implications-of-artificial-intelligence>.
- Oxford Internet Institute, *Oxford University Calls for Tighter Controls to Tackle Rise in Deepfakes*, The Times, May 7, 2025, <https://www.thetimes.co.uk/article/oxford-university-calls-for-tighter-controls-to-tackle-rise-in-deepfakes-ljvstnrmj>.
- Pazzanese, Christina. "Ethical Concerns Mount as AI Takes Bigger Decision-Making Role." Harvard Gazette, October 26, 2020. <https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/>.
- Saenko, Kate. "A Computer Scientist Breaks Down Generative AI's Hefty Carbon Footprint." Scientific American, May 25, 2023. <https://www.scientificamerican.com/article/a-computer-scientist-breaks-down-generative-ais-hefty-carbon-footprint/>.
- Stefanini. "The Moral and Ethical Implications of Artificial Intelligence." Stefanini, June 20, 2023. <https://stefanini.com/en/insights/articles/the-moral-and-ethical-implications-of-artificial-intelligence>.
- The Times, "How to Avoid the Pitfalls When Using AI to Recruit New Employees," The Times, June 20, 2024, <https://www.thetimes.co.uk/article/how-to-avoid-the-pitfalls-when-using-ai-to-recruit-new-employees-ml3vnlpg>.
- Vincent, James. "Amazon Reportedly Scraps Internal AI Recruiting Tool That Was Biased Against Women." The Verge, 10 Oct. 2018, <https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report>.
- World Economic Forum, 2020. *The Future of Jobs Report 2020*, October 2020, 26, https://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf