

CHALLENGING THE MORAL CHARACTER OF ARTIFICIAL INTELLIGENCE: A PHILOSOPHICAL EXAMINATION OF NATURAL LAW AND ETHICAL RISKS

Ndubuisi Friday Nwankwo and Orji, Okpara Paul

Abstract

The notion that humans possess unique characteristics such as thought, moral awareness, self-regeneration, and emotions has been challenged by computer scientists who develop AI systems that seem to display these traits. This raises questions about the moral status of Artificial Intelligence (AI) and its ability to fulfill legal and moral responsibilities. This paper examines fundamental natural law and ethical issues in AI, focusing on moral autonomy, biases, cyber-terrorism, and digital warfare. Using a qualitative research approach, it analyzes relevant terms and examines moral conundrums underlying AI operations. The study argues the thesis that despite AI's sophistication, it lacks the moral, spiritual, and metaphysical dimensions of human existence. It maintains that while AI can perform some impressive cognitive feats, it lacks the moral autonomy, metaphysical depth, and ethical orientation that is required by natural law traditions to undertake in full moral or legal agency. To address these concerns, the study suggests that policymakers should establish clear regulations on AI development and deployment, while developers should prioritize safety, human autonomy, transparency, accountability, and human values in AI design.

Key words: Artificial intelligence, morality, natural law, reflective consciousness, moral autonomy

1. 0. Introduction

Artificial Intelligence (AI) has grown immensely during the past couple of decades and the division that once separated human cognitive abilities and machine capabilities have become non-existent. From natural language processing to self-driving cars and predictive policing, AI systems now carry out tasks once thought to require human intelligence alone. These advancements have ignited fierce discussions surrounding the ethical and moral considerations of AI, particularly with respect to decision-making and accountability (Boddington, 2017, pp. 12–14). As machines begin to behave and act seemingly autonomously, scholars debate whether artificial intelligence should be afforded the rights of moral or even legal personhood. Such questions address the philosophical essence of humanity and moral responsibility (Gunkel, 2012, pp. 18–20).

A crucial issue in this ethical discussion involves whether machines can or could have — or merely simulate — moral autonomy, the capacity for reflective judgment, moral awareness, and rational choice. AI systems, which can process data efficiently, act via algorithms and ruleset, and have no metaphysical entity or intrinsic consciousness (Coeckelbergh, 2020, pp. 23–25). Drawing particular attention to this idea were theorists of natural law, who assert that morality cannot be reduced to outcomes or rules, but is rather rooted in human nature, within reason, and a universal moral order (Finnis, 2011, pp. 34–36). According to this view, AI can simulate moral decision-making but is incapable of moral reasoning. For this reason, the moral use of AI needs to be informed by something outside of the machines themselves.

Additionally, the implementation of AI into fields such as criminal justice, warfare, medicine, and surveillance risks algorithmic bias, discrimination, and dehumanization. Scholars have shown how

structural inequalities continue to be perpetuated through AI systems that encode data and operationalize flawed assumptions about human behavior (O'Neil, 2016, pp. 84–86). This raises an immediate ethical challenge: how to match the burgeoning power of A.I. to longstanding principles of justice and human dignity. A natural law-based moral frame can be applied to these technology modalities and offers guidance, focusing on what is in the common good, awareness of the sanctity of life, and an understanding of persons and personhood as irreducible. Whereas utilitarianism and technocracy are more predictable explanatory types, all of which produce empirical data, natural law provides a normative vision based on moral realism.

Another key component of this investigation entails responsibility and accountability. As these AI systems assume the roles of those traditionally played by moral agents, it becomes more and more ambiguous who should be held responsible when harm occurs: the designer, the user, or the system? Natural law philosophy holds that moral responsibility can be based on free will, self-awareness, and intentionality — all attributes not present in machines (Aquinas, 1947, 93–95). Delegation of life-and-death decisions to autonomous systems, particularly in digital warfare or surveillance, threatens moral and legal accountability. Hence, without a philosophical framework that serve as template for human agency, AI development risks taking society down ethically treacherous routes.

While AI can perform some impressive cognitive feats, this paper argues that it lacks the moral autonomy, metaphysical depth, and ethical orientation that is required by natural law traditions to undertake in full moral or legal agency. It thus cannot be treated as a moral agent in any sense deserving of that term. The significance of this study is in providing an essential philosophical framework — that of natural law — to judge and govern the ethical implications

of A.I. This research informs the expanding discussion around responsible governance of AI by interrogating machine morality limits. It calls on developers and policymakers to acknowledge that the value of any technological future will depend on the connections that can be made to human moral reasoning (MacIntyre, 1984, pp. 245–248; Finnis, 2011, pp. 102–104).

1.1. Conceptual Framework

The importance of clear and meaningful language has been a recurring theme in philosophical discussions (Mondin, 2007: 136). Confucius, for instance, emphasized the need for precise language, warning that unclear language can lead to misunderstandings and negative consequences (Confucius quoted in Borruso, 2009: 63). To ensure clarity, we begin by conceptualizing the following terms:

1.1.1 Artificial Intelligence

Artificial Intelligence (AI) is a term that originates from the Latin word *artificialis* meaning “made with art” and *intelligentia* meaning “understanding” or “knowledge”. It is defined as the field of computer science that deals with the development of machines that can perform a variety of tasks usually requiring human intelligence, including learning from experience, reasoning, solving problems, and understanding language. AI is based on algorithms and data-driven models that aim to imitate or replicate human cognitive capabilities. Some categories of AI today include: facial recognition systems, autonomous (self-driving) car technology, and natural language processing tools such as ChatGPT. Even though AI can manipulate and decide on data, it remains without consciousness, intentionality, and moral awareness (Coeckelbergh, 2020, p. 24). This is because AI operates through programmed algorithms and statistical patterns rather than subjective experience, it cannot possess consciousness or intentionality. Its decisions emerge from

coded responses to data, not from any inner awareness or moral judgment. As John Searle explains, a computer can simulate understanding without truly understanding anything—it lacks the “aboutness” that characterizes human thought (Searle 1980, p. 420). This is why, despite its impressive capabilities, AI remains fundamentally different from human intelligence.

1.1.2 Natural Law

Natural Law springs from the Latin term *lex naturalis*, which means “law of nature,” and is often considered a system of moral principles that is part of human nature and can be discovered by reason. In his explication of natural law, Aquinas posits that moral order is embedded in the natural order, and can be arrived at through the human intellect, the moral order being a reflection of God’s divine design (Aquinas, *Summa Theologica*, I-II, Q. 94, Art. 2). For Aquinas, natural law is what the rational creature knows about the eternal law in accordance with the participation of the creature with the eternal law, directing him or her to common good (Aquinas, *Summa Theologica*, I-II, Q. 94, Art. 2). The principles of natural law — preservation of life, pursuit of truth, etc. — have universal application and are the common sense (Aquinas, *Summa Theologica*, I-II, Q. 94, Art. 2). The first principle of natural law is that “good is to be done and pursued, and evil is to be avoided,” from which flow particular moral directives (Aquinas, *Summa Theologica*, I-II, Q. 94, Art. 2). Natural law (like all good things) points each person toward that good, that end, that happiness, which (ideally subserviently) brings about divine law fulfillment and promotes virtue that leads all to flourish in society (Cf. Aquinas, *Summa Theologica*, I-II, Q. 94, Art. 2).

Therefore, deeply embedded in classical philosophy and Christian theology, natural law theorists claim that some rights and ethical norms are inherent to humanity and not dependent on human

legislation. Natural law affirms moral realism — the idea that moral truths exist outside of human opinion — and addresses ideas like justice, human dignity and the common good. Unlike legal positivism, which is built upon written laws, natural law derives its power from reason and the human nature itself. For this reason, it serves as a preliminary structure for addressing ethical considerations within technology and governance (Finnis, 2011, p. 35).

1.1.3. *Morality*

Morality is derived from the Latin *moralitas*, meaning “manner, character, proper behavior,” and refers to principles related to the distinction between right and wrong or good and bad behavior. It is the application of such ethical reasoning to actions of humans, commonly regulated by traditions in culture, religion, or philosophy. It is concerned with autonomous human actions which have both legal and ethical implications. AI, unlike the latter which relies on free will, moral conscience and intrinsic worth, is expected to follow the instructions programmed into it. Thus morality is essential to an understanding of human action and responsibility in private and public life (MacIntyre, 1984, p. 246).

1.1.4. *Moral and Legal Autonomy*

Moral autonomy is the ability of an agent to make moral decisions and acting on them have instead an internalized ethical guide. By the same token, legal autonomy refers to the capacity to hold oneself accountable for one’s actions before the law. These concepts depend upon the presence of rational deliberation, self-awareness, and intentionality — attributes which are fundamentally human. Unlike human beings, machines, including AI systems, do not have these features, and must therefore never be held morally or legally responsible in the same sense. This difference is important

in establishing liability and responsibility in cases with autonomous technologies (Boddington, 2017, p. 13).

1.2. Artificial intelligence and the Question of Morality

The question of morality automatically raises the question of what is the nature of morality? What are those acts that constitute a moral act? How do we determine them so as to recognize them in artificial intelligence? Do we categorize all issues as moral? To begin with, morality deals with the notion of right and wrong, good and bad. Gert and Gert (2023) have it that the term “morality” can be used either (i) descriptively to refer to certain codes of conduct put forward by a society or a group (such as a religion), or accepted by an individual for her own behavior, or (ii) normatively to refer to a code of conduct that, given specified conditions, would be put forward by all rational persons. In the two senses, we will accept the second definition by which morality is used normatively because that captures the universal nature of morality.

Approaching morality from its etymology, Long and Sedley agree that morality is taken from Latin : *moralis* , lit. ‘manner, character, proper behavior’, thus meaning the differentiation of intentions, decisions and actions between those that are distinguished as proper and those that are improper. Morality can be a body of standards or principles derived from a code of conduct from a particular philosophy, religion or culture, or it can derive from a standard that a person believes should be universal. Morality may also be specifically synonymous with ‘goodness’ or ‘rightness’ (Aton and Dona, 1989: 366-367).

In other words, the question of morality deals with the issue of what kind of beings are said to be moral agents. What qualities do moral beings possess which differentiate them from non-moral or amoral beings. Morality has to do with moral norms, the idea of what is

good and bad, right and wrong, painful and pleasurable. Morality is the reflection and state of moral acceptability. Concerning the question of morality in this context, one expects and seeks to know whether artificial creatures like robots enjoy moral status.

A being is said to be moral if such exercises both rationality and free will. Man is a moral being because both his passions (will) and intellect (reason) are perfectly combined in making and executing decisions without external control. It is in the above sense that animals, infants and trees are excluded from the moral realm. Man acts from rational choices and desires. Although some of his actions cannot be said to be totally free. However, there is freedom in his moral choices. His desires may be controlled by emotions but those emotions are all rational emotions, they are different from animal instincts which lack reason whatsoever (Glen, 2007: 66-69). From the above, the human person is different from artificial intelligence, since man carries out both physical and mental events with moral implications, whereas artificial intelligence carries out only physical events, though it may look as if it were mental events. It is not known to common sense neither have AI experts, through logic and experience, established beyond doubt that machines do have free will. Animals may even be closer candidates for morality than machines and robots because whereas ordinary brutes manifest instincts and sentience, we don't think machines do.

In addition, even though AI does not enjoy moral status there is moral implication of equating machine with man. If we begin to equate artificial intelligence to human intelligence then we are ipso-facto assigning moral status to AI. But, can AI ever be morally responsible for its actions. In other words, can a machine be taken to court for appropriate judgment or be jailed should it cut off someone's hand? Could machines and robots begin to acquire fundamental moral rights whose violation would require legal action against the culprit?

Obviously, that would be irrational and unrealistic. In the same vein, to equate man with machine would imply machine having freedom of choice and of decision making. However, from our knowledge, it is obvious that machines are determined on what to do and what not to do. By implication, machines are not free to make choice of their own or to decide on their own whereas man has and enjoys freewill. Also, man is dynamic, flexible and creative, whereas if a machine is invented for a particular work, it cannot do otherwise, because to do so signifies fault in the machine.

1.3. Natural Law and Ethical Issues in Artificial intelligence Development

Artificial intelligence poses some ethical and Natural Law issues which we would consider briefly in this section. Issues bothering on biases, bioterrorism, violation of the principle of teleology, of human privacy and autonomy etc are all ethical and Natural Law issues confronting man in the use of artificial intelligence.

1.3.1 Artificial Intelligence and the Problems of Autonomy, Justice and Agency

One of the many problems plaguing AI automation is autonomy. This problem bothers on how machines emulate decision-making that would normally belong to moral agents. "True autonomy", as natural law theorists understand it, requires rational and moral deliberation, based on human nature (Finnis2011, 101). But it is obvious that AI systems are driven by algorithms that lack the moral reasoning of the real person. This has triggered fears over the use of AI systems in environments where human autonomy must be considered, like health care and criminal justice. This disconnection gives rise to a risk of eroding the individual's moral agency by delegating choices to systems that cannot appreciate human dignity or purpose (George, 2001, p. 134).

Furthermore, justice and agency make the ethical application of AI under natural law even more complex. According to the natural law tradition, justice demands individuals receive what they are owed according to an objective moral order, not just results foretold in statistical inferences (Finnis, 2011, p. 213). AI, on the other hand, frequently reproduces the biases inherent to data, with disproportionately unjust outcomes in domains such as predictive policing and hiring algorithms (O'Neil, 2016, p. 85). Furthermore, granting agency to the machines confuses accountability, flouting natural law's premise that only rational humans can be moral agents responsible for their actions (George, 2001, p. 99).

1.3.2 Artificial Intelligence and The Principle of Teleology

Natural law is assumed to operate on the principle of teleology. The principle of teleology in natural law holds that everything in nature has an inherent purpose or end (telos) toward which it naturally tends. Human reason, by reflecting on nature and the structure of human life, can discover these purposes and align actions accordingly. Thus, moral goodness is found in fulfilling the natural ends embedded in our nature, such as the pursuit of truth, preservation of life, and flourishing in community. This built-in purposes may not be easily susceptible to empirical verification. One natural law principle that AI challenges is the principle of "Telos" or "Purpose" (Aristotle, 1999, p. 104-105). Telos is a concept which describes the ultimate goal or direction of something, which is often seen as a fundamental aspect of natural law. AI systems, particularly those using machine learning, can be seen as violating this principle in the following way: firstly, AI systems are programmed to optimize specific objectives, but they do not have an inherent telos or purpose. Instead, their purpose is assigned by humans. Secondly, AI systems can adapt and change their behaviour based on data, but this adaptability can lead to unintended consequences, making it difficult to determine

their telos. These examples are a direct challenge to the teleology principle identifiable in Natural Law.

1.3.3. Artificial Intelligence and Biases

The notion that artificial intelligence (AI) operates without biases and prejudices has been widely touted, suggesting its potential to revolutionize fields like law and morality. However, a closer examination reveals that AI is not entirely bias-free. A UNESCO report, «Artificial Intelligence: Examples of Ethical Dilemmas» (2023), highlights this issue. For instance, a search engine query for «greatest leaders of all time» predominantly yields male personalities, perpetuating a gender bias rooted in societal stereotypes.

This bias is further exemplified in image searches, where queries for “school girl” often produce results featuring women and girls in sexualized costumes, whereas searches for “school boy” yield more innocuous results. These examples demonstrate how AI systems can perpetuate and amplify existing biases, rather than providing objective results.

The underlying issue lies in the fact that AI systems process vast amounts of data, prioritizing results based on user preferences and location. This creates an “echo chamber” effect, where biases and prejudices are reinforced and further entrenched online (UNESCO, 2023). As AI becomes increasingly integrated into various aspects of our lives, it is essential to acknowledge and address these biases to ensure that AI systems promote fairness, equality, and justice. (see UNESCO (21 April, 2023).”Artificial intelligence: Examples of Ethical Dilemmas.” <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics/cases> Accessed September 30 2023).

1.3.4. Artificial Intelligence and the moral issue of copyright

Artificial intelligence (AI) has sparked intense debate regarding copyright issues. Copyright refers to the legal right to control the publication, copying, and distribution of original works, such as literature, music, and art (Cornish & Llewelyn, 2013, p. 12). The integration of AI in creative industries raises significant ethical concerns. For instance, in 2019, Huawei announced that an AI algorithm had completed the unfinished Symphony No. 8 by Franz Schubert (Huawei, 2019). This development prompts questions about authorship and copyright in the context of AI-generated creative works.

If AI algorithms can produce original works of art, do they qualify as authors, and should they be granted the same rights as human artists? The answer to this question requires a reevaluation of the concept of authorship and the role of creativity in society (Bostrom & Yudkowsky, 2014, p. 15). Creativity is a vital component of open, inclusive, and pluralistic societies, and the impact of AI on human creativity warrants careful consideration (Florida, 2014, p. 10).

1.3.5. Cyberterrorism and Digital Warfare

Artificial intelligence raises ethical issue which borders on the use of AI tools and applications by both state and non-state actors to harm common people. According to C. Kenneth, AI provides a number of tools that are particularly useful for authoritarian governments: smart spyware, face recognition and voice recognition allow widespread surveillance; such surveillance allows machine learning to classify potential enemies of the state and can prevent them from hiding; recommendation systems can precisely target propaganda and misinformation for maximum effect; deepfakes aid in producing misinformation; advanced AI

can make centralized decision making more competitive with liberal and decentralized systems such as markets (Cukier, July/August 2019 : 192–98).

Terrorists, criminals and rogue states may use other forms of weaponized AI such as advanced digital warfare and lethal autonomous weapons. By 2015, over fifty countries were reported to be researching battlefield robots. Machine-learning AI is also able to design tens of thousands of toxic molecules in a matter of hours which are used for warfare of higher dimensions (see UNESCO (21 April, 2023).» Artificial intelligence: Examples of Ethical Dilemmas.»<https://www.unesco.org/en/artificial-intelligence/recommendation-ethics/cases> Accessed September 30 2023).

1.3.6. Artificial Intelligence and Bioterrorism

The convergence of artificial intelligence (AI) and biotechnology poses a profound moral threat to human well-being and the natural environment. By manipulating living organisms through AI-enabled biotechnology, individuals can cause widespread harm, disrupting human health and the environment (Richard, 2015, p. 31). This can lead to immense suffering and diminished happiness among humans. Furthermore, the intentional release of harmful biological agents, known as bioterrorism, is a growing concern. As defined by the U.S. Centers for Disease Control and Prevention (CDC), bioterrorism involves the deliberate release of viruses, bacteria, toxins, or other harmful agents to cause illness or death in people, animals, or plants (Richard, 2015, p. 31).

Bioterrorism agents, such as smallpox and anthrax, can have far-reaching consequences, including widespread fear and panic (Bill, 2007, p. 96). The ease of obtaining and disseminating these agents makes them an attractive option for terrorists. However,

military leaders have acknowledged the limitations of bioterrorism as a military asset, including the difficulty of targeting only enemy forces. The potential consequences of AI-driven biotechnology are alarming, and experts like Bill Joy have warned of the dangers of genetic engineering in the hands of bio-terrorists (Bill, 2007, p. 96). The moral implications of AI-enabled biotechnology are significant, and it is essential to address these concerns to prevent devastating consequences.

Therefore, the integration of AI and biotechnology poses a substantial moral threat to humanity and the environment. Bioterrorism is a growing concern, and the potential consequences of AI-driven biotechnology are alarming. It is crucial to address these concerns to prevent devastating consequences.

1.3.7. Acceleration of Crimes:

Artificial intelligence accelerates the rate of crimes in human society. If artificial intelligence makes anything faster at all, it is the perpetration of heinous crimes that it makes even the more faster. Crimes we never heard of in the 18th century have been committed with ease in our time due to the facilitation given to them by artificial intelligent agents. Let us take two internet crimes committed much easier with the aid of AI to illustrate our point.

Phishing: this occurs in the form of theft. It is the malicious act of keeping a false website or sending a false e-mail with the intent of masquerading as a trustworthy entity in order to acquire sensitive information, such as usernames, passwords, and credit card details. It is also a high-tech identity theft that steals personal information and identity of unsuspecting users as well as commits acts of fraud against the legitimate individual or organization that have been victimized (Patrazena, 2006: 88). The perpetrators send false emails fraudulently designed to mimic legitimate businesses,

requesting for personal information which if released can allow them access one's bank accounts, including details of debit and credit cards.

Cyber Terrorism - cyber-terrorism is defined as the malicious act of using computer network to disrupt the normal processes of critical national infrastructures (such as energy, transportation, government operations), including the coercion or intimidation of public and private citizens. Cyberterrorism is also cyber extortion as a sort of cyber terrorism whereby the website, e-mail server, i or computer system leveraging on ransom ware is put under attack by hackers for denial of services and demand for ransom (Ibid, p.90). Before the emergence of Artificial Intelligence these heinous crimes committed on daily basis now were never heard of. However, with the aid of AI, cybercrimes of higher magnitude have come to be part of our daily lives, terrorizing and upsetting human trust and relationships. This reminds an egregious elements in Artificial Intelligence.

1.3.8. Assault on Human dignity:

Artificial intelligence poses a great threat to human dignity in some respects. Philosophically, artificial intelligence seems to make human dignity felt only where and when we are useful. In other words, it seems we can be replaced with artificial beings that can work faster and more efficient than we do. Therefore, only those who can do what artificial intelligence cannot currently do or can work efficiently with artificial intelligence, are perceived as relevant and respected whereas those with little or no technical skills are made to appear redundant, weak, useless less creative and even undesirable. This has in many ways degraded the intrinsic and inviolable worth and dignity of human beings. Thus Artificial Intelligence is succeeding in reducing human beings to mere functionaries.

1.3.9. Artificial Intelligence and the Crisis of Employment

No one can understand the place of artificial intelligence in the employment sector if he has not first understood the sacredness, value and morality of labour or work. In other words, no one can evaluate the moral harms and benefits of artificial intelligence in the labour market who has no sense of the dignity of labour. When we talk about employment we are talking about engaging the totality of the human person to manifest his skills and realize his essence. He never can learn who never acquired new experiences from the work environment. It is through employment to work in certain areas that one manipulates the brute, natural environment for his survival. Employment enables the human person to work, it engages the human person in the kind of job for which he trained. Employment, job and work each has a significant status in philosophical discourse. The way the philosopher sees working is different from the way the sociologist, the lawyer, the labourer himself sees it. From a philosophical perspective, work and working conditions acquire a significantly moral dimension. Little wonder the ancient sage, Plato, defined justice in terms of division of labour. According to him, it would be an ideal state that “every member of the community must be assigned to the class for which he finds himself best fitted.” Justice is, for Plato, at once a part of human virtue and the bond, which joins man together in society. It is the identical quality that makes good and social. Justice is an order and duty of the parts of the soul, it is to the soul as health is to the body. It consists in giving each what is his due. And this essentially manifests in society when each citizen is seen doing the kind of job that feats his nature according to the three classes of the soul (Plato, 1951: 117).

The idea is that work, doing one's duty, being placed on the right job through employment enables man to realize his essence.

That being noted, artificial intelligence (AI) has been increasingly integrated into various industries, including banking, law, and medicine. According to Russell and Norvig(2021: 18), AI applications are being used to perform tasks such as data analysis, pattern recognition, and decision-making. This has led to concerns about job displacement and the potential for AI to automate jobs previously performed by humans. Some experts estimate that up to 47% of US jobs are at «high risk» of potential automation (Osborne & Frey, 2013). As a result, there are growing concerns about the impact of AI on employment opportunities.

The International Labour Organization (ILO) has published a report on the potential impacts of AI on employment. According to the report, while AI may displace some jobs, it could also create new opportunities and augment existing ones (Lanz, 2023). The report notes that office administration work is particularly vulnerable to automation, with women disproportionately affected due to their overrepresentation in clerical and service roles. This highlights the need for policymakers to consider the potential impacts of AI on employment and to develop strategies to mitigate any negative effects. By doing so, they can help to ensure that the benefits of AI are shared by all.

The ILO report emphasizes that the impact of AI on employment will depend on how the technology is managed and regulated. As Lanz (2023) notes, “generative AI is neither inherently good nor bad... its impacts will largely depend on how the technology is managed and regulated.” This highlights the need for policymakers to take a proactive approach to regulating AI and ensuring that its benefits are shared by all. By doing so, they can help to mitigate the negative impacts of AI on employment and ensure that its benefits are shared by all. This will require a coordinated effort from governments, industry leaders, and civil society organizations.

The integration of AI into the workforce also raises concerns about job quality and the potential for AI to exacerbate existing inequalities. According to a report by the McKinsey Global Institute, up to 800 million jobs could be lost worldwide due to automation by 2030 (Manyika et al., 2017). This highlights the need for policymakers to consider the potential impacts of AI on job quality and to develop strategies to mitigate any negative effects. By doing so, they can help to ensure that the benefits of AI are shared by all and that its negative impacts are minimized. This will require a coordinated effort from governments, industry leaders, and civil society organizations.

The integration of AI into the workforce raises important concerns about job displacement, job quality, and the potential for AI to exacerbate existing inequalities. While AI has the potential to bring significant benefits, including increased productivity and efficiency, it is essential that policymakers take a proactive approach to regulating AI and ensuring that its benefits are shared by all. By doing so, they can help to mitigate the negative impacts of AI on employment and ensure that its benefits are shared by all. This will require a coordinated effort from governments, industry leaders, and civil society organizations. By and large, it is up to policymakers to ensure that the benefits of AI are shared by all.

1.4. Proposed Solutions to Natural Law and Ethical Dilemmas in AI Development

1.4.1 Integrating Human Oversight in Improving AI Moral Reasoning

One of the most critical solutions to many ethical dilemmas related to AI is the integration of human oversight in the decision-making process of AI (such as detecting bias during prediction etc.). Although algorithms can mimic moral reasoning, no AI

system possesses the inherent morality present in human beings. Human oversight can be an effective way to keep AI in check. This oversight would make sure that humans who could make moral judgments review AI decisions, especially those with significant moral implications. One of the reasons is that in critical areas such as autonomous vehicles or medical diagnosis, human intervention can help reduce the unexpected consequences of AI actions. This approach protects the moral agency of artificial systems and ensures that they operate within the moral landscape determined by nascent natural law.

1.4.2 Integrating Algorithms that Resist Bias and Ensure Accountability

In responding to the many fundamental natural law and ethical issues in AI involves building algorithms that can robustly resist bias so that they not reinforce biased results. AI systems can be designed in such a fashion that they ensure bias minimization by promoting fairness and transparency during design and training. One way this could be accomplished would be through training on diverse data sets that represent a range of social and cultural perspectives, decreasing the likelihood of discriminatory behaviour. Equally important is accountability in AI, where those creating and deploying AI that affects the life of people should be held responsible. Also, creating mechanisms for auditing AI systems and keeping the decision-making process transparent can be helpful in identifying where biases exist and how they can be corrected. These design changes will move AI technology closer to the ideals of natural law which are anchored in justice and equity.

1.4. 3 International Legal Frameworks for AI Governance

We are aware that many developed nations of the world, have their AI regulative frameworks. For example, the EU AI Act (Europe), U.S. Executive Order on AI (2023) (America),

and China's Administrative Measures for Generative AI (2023) (Asia): are legal frameworks that reflect each region's values—Europe's focus on rights and safety, America's innovation-led and security-driven governance, and China's emphasis on state control and social order. However, the world is yet to develop AI regulatory frameworks that can apply across countries. Therefore, a sound international legal architecture for AI is utterly important to meeting the ethical challenges presented by AI and to regulate its development and use. This framework would set forth universal guidelines and rules that guarantee AI to conform to standards of ethics agreed upon on a global level. It would establish means for stewardship of A.I. applications, placing restrictions on their employment in sensitive areas — warfare, surveillance and criminal justice, to name a few. Nations working together and within the context of transnational law can prevent AI from violating basic rights and ensure it remains in line with natural law principles such as the dignity of the human person and the common good. Such global cooperation would also give us the legal tools needed to hold people and institutions accountable for any AI harm they cause. Regulatory measures like this would help to ensure that AI will ultimately be developed in keeping with the general ethical and moral aims of the world.

1.4 4 Incentivizing AI developers to embed principles of natural law into AI development.

An essential remedy to the ethical dilemmas AI presents is the overt infusion of natural law fundamentals into the design and development of AI's archetypal propositions. Natural law, which prioritizes human dignity, justice and the common good, offers a moral framework that can help steer the direction AI will take. These principles should be taught to all developers, and they should be incentivized to design AI systems that respect fundamental human rights, preserve justice, and promote human

flourishing. In this way, developers can help to avoid possible dangers of these technologies by making sure that the values of AI mirror those of the natural law tradition (such as respect for life, truth, and the autonomy of humans). Embedding these principles in the very beginning would enable AIs to perform not only under technical constraints but also under moral and ethical conditions that further human welfare. This would ensure that the momentum of AI development is guided by time-tested ethical principles, anchored on human-centered values essential to protect the moral fabric of society.

1.5 Summary of Findings

This study dealt with the examination of some fundamental natural law and ethical issues raised by the development and rapid deployment of Artificial Intelligence. Such issues such as moral bias, job displacement, bioterrorism, acceleration of crimes, legal and moral agency, the problem of teleology, problem of autonomy and agency, etc were thoroughly examined to determine the moral status of AI systems. The paper argued the thesis that, though AI performs cognitive functions very impressively, it cannot exhibit moral autonomy or ethical reasoning that human beings obtain from, for example, natural law traditions. In response to the above problems, the study proposed strategic solutions to control AI based on human choices, and to create algorithms that are explainable and display high-level of bias avoidance based on natural law principles. Adopting the method of qualitative analysis, the study finds that progress in AI technology must be aligned with sound ethical frameworks to prevent the erosion of moral and natural law principles in AI development and deployment.

Furthermore, the study connects the wisdom of classical philosophical traditions with the challenges emerging from AI technologies — marshaling natural law concepts to the ethical

dilemmas that are presented by real-world artificial intelligence practices, such as bias, surveillance and autonomous decision-making. The study advances such a potential framework in which developers and policymakers are morally responsible for the outputs of AI. The paper highlights the shortcomings of AI's moral capabilities, presenting a counter-argument against granting AI full agency, motivating readers to recognize the irreplaceable value of human moral judgment. Such a critical perspective supplements the latest discourse on responsible AI with principles engaged with lasting human values. The study is a clarion call for leveraging the strengths of AI in ways that promote genuine human flourishing in the society.

1.6 Conclusion and Recommendations

It is the submission of the study that the evolution of AI must be guided by commitments to moral accountability, human dignity, and ethical stewardship. The study reinforces the thesis that although AI can display complex cognitive capabilities, however, it can never replace human moral agency and spiritual depth as enshrined in natural law principles. Making authentic progress in AI ecosystem, ethical frameworks should be created in conjunction with AI technologies to ensure that the fundamentals of humanity are preserved. The future of AI is not about replacing human judgment, it is about augmenting it under appropriate oversight, regulation, and moral alignment. The study submits, among other things, that AI developers should incorporate natural law principles as a compass in building AI algorithms so as to promote human dignity rather than being a threat to it.

On the strength of the above conclusion, the study offers the following recommendations:

i. *Create a certification for ethical AI systems (For Government agencies and standardization bodies)*

Mandatory ethics certification for AI systems before their public deployment, which can be assigned by governments and standardization bodies like International Organization for Standardization (ISO) and Institute of Electrical and Electronics Engineers IEEE), would be necessary in producing AI professionals that adhere to best ethical practices in their daily operations. This would build public confidence, that AI technologies can serve society responsibly.

ii. *Implement Interdisciplinary AI Curriculum (For Higher Education Institutions and Related Bodies)*

University program courses in AI and computer science should integrate ethics, philosophy, and law. This will ensure that developers possess not only technical prowess, but also ethical character. Promotion of joint activity across engineering and humanities faculties would ensure the cross-pollination of dynamic ideas and frameworks that enhance the production of responsible AI technologies.

iii. *Create Explicit Public Ethical Review Panels (For Civil Society and National Parliaments)*

Independent panels of civil society and government should be formed to review the societal impact of AI. These panels should consist of ethicists, legal scholars, tech experts, and representatives from different professional communities. They would flag, stop or modify ethically dubious A.I. rollouts. Doing this would enhance transparency, accountability, and public participation in AI governance.

iv. Create Natural Law Design Frameworks (For Software Houses and Open Source Communities)

Human-centered design methodologies for AI must be implemented by tech firms and open-source communities using shared ethical design templates based on natural law. Such frameworks would inform developers about ways to embed fairness, justice, and human dignity in AI systems. (Organizations such as Mozilla Foundation and universities can work on this) It would clarify ethical compliance and institutionalize responsible innovation.

v. Launch AI Ethical Impact Bonds — For International Financial Institutions and Private Investors

Investors and development banks must create “AI Ethical Impact Bonds” that tie funding to positive ethical outcomes. Such bonds could fund efforts to reduce algorithmic bias or increase access to justice and education. These instruments could be propelled by organizations such as the World Bank or by philanthropists. That would marry ethical obligation and financial gain.

Once these recommendations are adopted, AI will not only get smarter but also more humane. It would embody our deepest values, not merely our technical ambitions. In that future, AI will be able to deal with a lot of fundamental natural law and ethical issues that currently confront it.

References

- Adams, J. (1979). *A Defence of the Constitutions of Government of the United States of America*. 3rd edition. Philadelphia; repr. Darmstadt, Germany: Scientia VerlagAalen.
- Albertus Magnus. (1951). On the nature of truth. In *Anselm of Havelberg & Albertus Magnus* (Eds.), *Opera omnia* (Vol. 1, pp. 1-25). Paris: Vrin.
- Andrew, G. (2008). "Algorithm". In Fuller, Matthew (ed.). *Software studies: a lexicon*. Cambridge, Mass.: MIT Press.
- Aquinas, T. (1948). *Summa theologiae*. Translated by Fathers of the English Dominican Province. New York: BenzigerBrothers.
- Aristotle. (1999). *Metaphysics*. Translated by J. Sachs, Green Lion Press.
- Aton, L. & Dona, S. (1989). *The Hellenistic Philosophers: Traditions of the Principal Sources with Philosophical Commentary*. Cambridge: Cambridge University Press.
- Bill, J. (2007)., *Why the Future Does Not Need U: How 21st Century Theology Threat make Humans an Endangered Species*. London: Random House.
- Boddington, P. (2017). *Towards a Code of Ethics for Artificial Intelligence*. Springer.
- Borruso, S. (2009). *The Art of Thinking*. Nairobi, Pauline Publications.
- Bostrom, N. (2014a). *Superintelligence: Paths, dangers*, Oxford: Oxford University Press.
- Bostrom, N., & Yudkowsky, E. (2014b). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Cath, C. (2020). "Natural Law and the Development of Artificial Intelligence". *Journal of Moral Philosophy*, 17(1), 1-18.
- Cicero. (1998). *On the republic*. Translated by C. W. Keyes. Cambridge, MA: Harvard University Press.
- Coeckelbergh, M. (2020). *AI Ethics*. MIT Press.
- Cornish, W. R., & Llewelyn, D. (2013). *Intellectual property: Patents, copyrights, trademarks & allied rights*. Sweet & Maxwell.
- Cukier, K. (July/August 2019). "Ready for Robots? How to Think about

- the Future of AI”, *Ethical Dilemmas*. vol. 98, no. 4: pp. 192–98.
- Finnis, J. (2011). *Natural Law and Natural Rights* (2nd ed.). Oxford University Press.
- Florida, R. (2014). *The rise of the creative class: Revisited*. Basic Books.
- George, R. P. (2001). *The Clash of Orthodoxies: Law, Religion, and Morality in Crisis*. Wilmington, DE: ISI Books.
- Glenn, J. P. (2007). *A Tour of the Summa of St Thomas Aquinas*. New Delhi, Bekha Printers Ltd.
- Gert, B. & Gert, J. (2017). “The Definition of Morality” in *Stanford Encyclopedia of Philosophy*. Available at: <https://plato.stanford.edu/archives/fall2017/entries/morality-definition/> Accessed on: 26th July, 2024.
- Gunkel, D. J. (2012). *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*. MIT Press.
- Kirk, G. S., Raven, J. E., & Schofield, M. (1983). *The Presocratic philosophers: A critical history with a selection of texts*. Cambridge, UK: Cambridge University Press.
- ILO (August 21, 2023) “Generative AI more likely to Create jobs rather than Destroy Jobs. Available at <https://www.ilo.org/resource/news/generative-ai-likely-augment-rather-destroy-jobs> Accessed 26th July, 2024
- Lanz, J. A.. (21 August, 2023). “ AI Won’t Lead to an Unemployment Crisis, but Jobs Are at Risk: Report” International Labour Organization <https://decrypt.co/153209/ai-wont-bring-an-unemployment-crisis-but-jobs-are-at-risk-ilo-report> Accessed September 29 2023
- Manyika, J., Chui, M., Bisson, P., Bughin, J., Woetzel, J., & Stolyar, K. (2017). *A future that works: Automation, employment, and productivity*. McKinsey Global Institute.
- MacIntyre, A. (1984). *After Virtue: A Study in Moral Theory* (2nd ed.). University of Notre Dame Press.
- McCarthy, J. and Hayes, P. (1969). ‘Some Philosophical Problems from the Standpoint of Artificial Intelligence’. Meltzer, B. and Michie, D.(eds.) *Machine Intelligence*. Edinburgh: Edinburgh University Press.
- McGaughey, E (2022). ‘Will Robots Automate Your Job Away? Full Employment, Basic Income, and Economic Democracy’. Industrial

Law Journal. 51(3), 511–517

- Mondin, B. (2007). *Philosophical Anthropology*. Trans. Myroslaw A. Cizdyn, Rome: Urban University Press.
- O’Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishing Group.
- Osborne, M. A., & Frey, C. B. (2013). The future of employment: How susceptible are jobs to computerization? *Technological Forecasting and Social Change*, 114, 254-280.
- Patrazena, N. (2006). « Internet Crimes and Crime-related cases». *Italian Journal of Cybercrimes*, 2(3): pp.88-107
- Plato, (1951). *The Republic*, trans. W.D. Ross. Oxford: Clarendon House.
- Richard, P. (2005). *The Demon in the Freezer*. New York: Ballantine Books.
- Richard, J. T. (2015). *Bioterrorism and Biosecurity: A New Era of Concern*. Nova Science Publishers.
- Russell, S. J.; Norvig, P. (2021). *Artificial Intelligence: A Modern Approach*. (4th ed.). Hoboken: Pearson, 2021.
- Searle, J. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences*, 3(3), 417-424.
- Searle, J., (2014). “What Your Computer Can’t Know,” *New York Review of Books*, October 9.
- Schmidhuber, J. (2022). *Annotated History of Modern AI and Deep Learning*. London: Everyman’s Library.
- Sternberg, R. (2003). “A broad view of intelligence: The theory of successful intelligence”. *Consulting Psychology Journal: Practice and Research*. 55 (3): 139–154.
- Tarnoff, B. (August 2023), “Lessons from Eliza”. *The Guardian Weekly*. 4. pp. 34-3
- UNESCO (21 April, 2023). “Artificial intelligence: Examples of Ethical Dilemmas.” <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics/cases> Accessed June 30 2024
- Wallach, W. (2015). “A quest for human dignity in the face of AI”. *Ethics and Information Technology*, 17(2), 135-144