Ethics Concerns in the Use of Computer-Generated Images for Human Communication

José Manuel Simões / Wilson Caldeira

University of Saint Joseph, Macau-China

July 2024

Keywords

artificial intelligence, communications, image-creation, synthetic imagery, ethics, comprehensive framework

Abstract

The rapid growth of artificial intelligence (AI) tools has contributed to various image-generation techniques. These tools have had a significant effect on how we perceive contemporary communication. This study presents a broad overview of the challenges and consequences associated with the influence of AI on communication, and its findings are supported by thoughts collected from various books and academic publications. Nevertheless, incorporating AI into communication through synthetic imagery raises concerns about communication that require careful consideration. This study investigates various ethical concerns and the significance of ethical guidelines and responsible practices in creating these applications. It intricately weaves together philosophical reflections, ethical deliberations, and societal responsibilities, shedding light on the complex interplay between ethics, technology, and human values within the context of machine development and broader societal realms.

Correspondance: jmsimoes@usj.edu.mo; wilson.caldeira@usj.edu.mo. To quote this article: José Manuel Simões / Wilson Gomes Caldeira. 2024. "Ethics Concerns in the Use of Computer-Generated Images for Human Communication", *Journal of Ethics in Higher Education*, 4(2024):169-192. DOI: 10.26034/fr.jehe.2024.5962

1. Introduction

In contemporary society, the proliferation of computers, smartphones and the Internet has promoted the manifestation of the inherent human inclination towards augmentation, particularly in domains that are integral to our fundamental nature as intelligent individuals. They provide us with rapid access to extensive general awareness and a massive collective repository of information (Wilczek, 2019), especially in areas essential to our innate intelligence, enabling swift access to vast knowledge and a substantial shared pool of information in today's society. Computers have undeniably assumed a dominant role in modern society, effectively enabling many facets of human existence. This encompasses a wide range of activities, such as financial transactions, travel arrangements, utility management, and even the most intimate forms of personal contact (Ramakrishnan, 2019). The integration of computers into these spheres of human existence not only streamlines processes but also enhances accessibility, connectivity, and efficiency. Their influence extends beyond these examples, permeating education, healthcare, entertainment, and numerous other domains, fundamentally altering how humans interact with and navigate the world. As technology advances, computers continue to evolve, promising further transformation and deeper integration into the very fabric of our existence.

In the future, it is plausible that a computer may generate a novel outcome, such as a mathematical theorem, that surpasses human comprehension in terms of both its proof and expression. The perspective exhibits a philosophical divergence from the prevailing scientific methodology (Ramakrishnan, 2019). Ultimately, contemplating the possibility of a computer generating a theorem or outcome surpassing human comprehension leads to a profound reevaluation of the boundaries of human knowledge, the evolving role of artificial intelligence in scientific discovery, and the intricate relationship between technology and human intellect.

In the last half-century, the idea of AI and intelligent robots has dominated discussions concerning human-computer interaction (Pentland, 2019). This era has witnessed an exponential advancement in technology, propelling AI

and robotics to the forefront of scientific exploration, societal discourse, and technological innovation. This prolonged dominance of AI and intelligent robotics in human-computer interaction discussions underscores the transformative impact and the complexities associated with these technologies. As they continue to evolve and permeate various aspects of our lives, these discussions are crucial in shaping policies, ethical frameworks, and societal perceptions to ensure that the integration of AI and intelligent robots aligns with human values, serves the greater good, and augments human capabilities without compromising fundamental ethical principles and human dignity.

Over time, concerns about the potential control of intelligent machines shifted from the realm of science fiction to the warnings of thinkers like Alan Turing, and the thought that AI might one day threaten humanity has been around for a while (Vold & Harris, 2021). The current academic dialogue of AI is marked by divergent viewpoints. Certain scholars and intellectuals laud AI as a panacea for humanity's most pressing dilemmas, while others perceive it as an exceedingly formidable issue that humanity has encountered, potentially even posing an existential peril (Garvey, 2021). Undoubtedly, the outcomes of the extensive sociotechnical system encompassing data, individuals, locations, and entities known as 'AI' are not limited to the extremes. However, in the face of conflicting assertions and qualifications from partisan experts, it becomes imperative to explore alternative perspectives and their sources (Garvey, 2021). Evaluating sources, considering diverse expert opinions, and engaging in interdisciplinary discussions are crucial in shaping a balanced, informed discourse that addresses both the promises and perils of AI, thereby steering its development towards ethical and beneficial outcomes for humanity.

AI applications have advanced significantly in recent years, due to their high performance, availability, and intelligence. These applications are becoming more common in many areas of human life, including search engines, social media, and decision-making processes (Qiu et al., 2019). AI has been effectively employed in image classification, object identification, voice control, machine translation, and more sophisticated fields such as drug composition analysis, brain circuit reconstruction, particle accelerator data

analysis, and DNA mutation analysis, and this impact extends not only to specific areas but also to social sciences (Qiu et al., 2019). Moreover, the acknowledgment that AI's influence extends beyond technical fields into the realm of social sciences emphasizes its pervasive and transformative role in shaping contemporary society (Qiu et al., 2019).

AI has profoundly influenced visual communications by introducing generative AI tools that have revolutionised visual content creation. These tools have proven effective in various domains like synthetic imagery creation, computer vision, and medical imaging. Using these resilient generative models, which understand data distributions through noise reduction processes, synthetic images of exceptional quality can be produced (Blattmann et al., 2022). The training process involves using a clean image to generate a noisy image, which is then refined to restore the original image, showcasing the simplicity and efficacy of these AI models (Blattmann et al., 2022). Moreover, the integration of generative AI models in visual content creation has not only transformed the field of computer-assisted learning in art but has also enhanced users' creative abilities and understanding of aesthetics (Tang et al., 2018). These models have been particularly beneficial in scenarios where obtaining labelled information is challenging, such as in multispectral imaging, where synthetic data has been used to improve deep learning frameworks for semantic segmentation (Tang et al., 2018).

Furthermore, the impact of generative AI models extends beyond image edition, with applications in fields like text-to-image production, where these models excel in producing authentic, realistic images based on textual input (Liu et al., 2021). Generative models have also been used in applications like semantic map inpainting and materials design, showcasing the versatility and broad applicability of these techniques (Müller et al., 2021). Therefore, we can assume that the transformative effects of generative AI on the conceptualisation and production of visual content are evident through the advancements in image synthesis, creative expression, and educational processes. These AI models have improved the quality of synthetic images and opened up new possibilities for innovation and exploration in various disciplines.

This research provides a comprehensive analysis of the difficulties and outcomes linked to the impact of AI on communication. Insights gathered from a variety of books and academic articles support the conclusions. However, integrating AI into communication via synthetic images gives rise to challenges in communication that need thoughtful deliberation. Our study examines ethical considerations and the importance of moral norms and responsible behaviours in developing such technologies.

1.1 Nature and magnitude of the existential risks

Regardless of the rising frequency of AI-related news in the media, the general population frequently needs a more understanding of AI (Touretzky et al., 2019). There is a significant disparity between the increasing prevalence of AI-related news coverage in the media and the often limited comprehension of AI within the general population (Touretzky et al., 2019). Despite the growing frequency of AI-related discussions in public discourse, there is a notable gap in the public's understanding and knowledge about AI's intricacies, capabilities, and limitations. This discrepancy underscores the need for improved education, accessible information dissemination, and enhanced public awareness initiatives to bridge this gap, enabling the general populace to develop a more informed and nuanced understanding of AI and its implications in modern society.

According to James de Traz (2023), the present trend in the contemporary world is rather contradictory. Combining the rule of least effort and technological conveniences encourages us to minimise our efforts and accomplish as little as possible. Still, AI can make sophisticated predictions, automate, plan, target, and personalise. AI can often demonstrate human-like intellectual ability and may assist with numerous areas of our lives (Russell et al., 2015). One fascinating aspect of artificial intelligence is its potential to supplant human creators in some creative actions. For example, there are neural networks that write poetry and essays and create realistic images. This concept of artificial intelligence replacing the creator raises issues about human beings' role in creative processes and how it might impact self-fulfilment in the present digital culture (Vitulyova, 2020). However, due to its heavy reliance on data and its autonomy and self-learning nature, the

increasing adoption of AI has raised multiple concerns (Russell et al., 2015). Despite growing attention to this matter, there remains considerable controversy over the nature and magnitude of the existential risks presented by AI (Vold & Harris, 2021).

Communication was and continues to be our most valuable innovation. It has assisted us in preserving and passing on our information, learning, discoveries, and intellect from person to person and generation to generation (Gawdat, 2021). The process of communication is of utmost importance as it serves as an essential mechanism in shaping one's views of others, developing, and sustaining social connections, and attaining collaborative results (Hohenstein et al., 2023). Since its inception in the mid-twentieth century, the field of research and implementation known as AI has had a strong association with communication, and most attempts to consider calling into question the concept of communication in AI have stayed confined to a technical perspective on communication and information theory (Natale, 2020). Whether truthfully accepted or not, communication is crucial to both the theory and practice of artificial intelligence (Gunkel, 2012). Despite concerns about AI's detrimental impacts on society, the broader implications of using it to communicate continue relatively unexplored (Hohenstein et al., 2023).

To ensure the compatibility of highly capable AI with humans in the future, we mustn't develop them in isolation from human involvement and thereafter attempt to establish compatibility. Instead, we must proactively define the concept of "human compatibility" (Dragan, 2019). The potential for a human-AI ecosystem to generate beneficial societal outcomes by promoting fairness and transparency in decision-making is readily apparent. However, it is important to acknowledge the potential concerns associated with what might be referred to as a "tyranny of algorithms," when individuals who have not been elected to positions of power are effectively governing global affairs via their expertise in data analysis (Pentland, 2019). The advancement and expansion of AI applications have given rise to novel and significant inquiries for technologists, humankind, and sentient beings at large (Gabriel, 2020). The primary concern pertains to the determination of whose values AI systems should prioritize, and which values should serve as the basis for

alignment. One of the conceptualizations of AI is characterized by its wide utilitarian orientation. The argument posits that, over an extended period, it is imperative to advance these technologies to optimize satisfaction for the largest possible population of sentient entities. Another method exhibits characteristics reminiscent of Kantian philosophy. This suggests that the rules governing AI should be limited to those that may be rationally conceived as universally applicable, such as concepts of fairness or beneficence (Gabriel, 2020). Progress in AI holds the potential to enhance the quality of human existence. However, it is important to note that this phenomenon also poses substantial risks, a topic that is comparatively less commonly addressed (Gawdat, 2021). It is plausible that we have presently arrived at a juncture when the majority of countries find themselves without the ability to effectively counteract the collective influence wielded by a select few influential multinational corporations, who exercise dominion over both our present state and future trajectory in the digital realm (Ramakrishnan, 2019). Irrespective of the diminishing agency, humanity persists in progressing toward a future where AI permeates all aspects of society. The general public will find it arduous to resist the allure of AI's ease and potency, while corporate entities and governing bodies will be compelled to embrace it due to its inherent competitive benefits (Ramakrishnan, 2019). In response to these concerns, several national and international entities, including governmental agencies, business sectors, and research institutes, have undertaken substantial endeavors to formulate ethical standards and foster dynamic deliberations pertaining to the ethical dimensions of AI (Khan et al., 2023). The development, implementation, and governance of large-scale sociotechnical systems, including the deployment of AI, have significant implications for our understanding of what it means to be human in an increasingly computerized world. The implications stem from the intricate interplay between technology and society, and the possible unforeseen repercussions that may develop as a result of these interconnections (Garvey, 2021). AI includes both promising prospects for economic growth and significant troubles that need to be handled. Ones that need to be addressed today (Gawdat, 2021).

2. The Revolution of Communication through Artificial Intelligence

Throughout history, humans have recognized the profound impact of stories, images, and language on our minds. These powerful tools can shape our thoughts, beliefs, and behaviors. Images have been used as a means of communication and persuasion since ancient times. They have the power to evoke emotions, engage our imagination, and convey complex ideas in a relatable and memorable way. The emergence of AI in the creation of synthetic images has drawn considerable interest recently. AI algorithms were developed to generate synthetic imagery that exhibits a high degree of realism. For example, AI-enabled systems demonstrate exceptional performance in tasks such as image recognition, online search, and language processing that involve converting text into images. According to prominent AI experts, it is anticipated that AI will surpass human performance within the forthcoming decade (Fogel & Kvedar, 2018).

2.1 The construction of synthetic imagery using AI

The use of Generative AI, specifically Large Language Models (LLMs), can significantly transform the methods through which humans communicate (Hohenstein et al., 2023), and AI applications like ChatGPT are increasingly being used to generate various forms of language, ranging from text messages and social media posts to computer programs, and speeches. Additional AI models, such as Dall-E and Midjourney, can generate visually appealing images that are artistically created, based on user prompts (Newton & Dhole, 2023). AI is being extensively employed in everyday communication. However, despite apprehensions regarding the adverse impacts of AI on society, the potential societal ramifications of its utilization in communication have yet to be thoroughly investigated (Hohenstein et al., 2023).

Image-creation capabilities have emerged as powerful tools in the realm of AI. These capabilities have revolutionized various industries and opened new possibilities for communication and expression. The construction of synthetic imagery using AI has become a significant phenomenon in media, the web,

and art. The progress in deep neural network technology and the abundance of extensive datasets have resulted in a remarkable level of similarity between pictures and videos and human perception, as well as the capability to deceive ˜sophisticated computer algorithms (Nguyen et al., 2022). The integration of probabilistic representations and statistical learning methods has led to noteworthy progress in AI, and machine learning (Russell et al., 2015). These advancements have paved the way for the development of AI technologies such as the construction of highly quality synthetic imagery. The ability to generate realistic synthetic images can blur the line between reality and fiction, making it challenging to discern what is genuine and what is artificially created. This can lead to a breakdown in trust and credibility in media, web content, and art. Misinformation and disinformation can be easily propagated through the creation and dissemination of deepfakes, further exacerbating the challenges of information verification and authenticity (Khoo et al., 2021). This issue is particularly significant for the legal system, which often relies on audiovisual materials as evidence in court proceedings, and whose reliability is now put into question. For instance, deepfake technology can be used to fabricate incriminating evidence, creating videos or images that falsely depict individuals engaging in illegal activities. This not only threatens the integrity of legal proceedings but also poses risks to individual rights, potentially leading to wrongful convictions based on falsified evidence. (Delfino, R. A., 2022)

In advertising and marketing, synthetic imagery has revolutionized the creation of visually stunning and captivating content. AI-generated images can be precisely tailored to specific target audiences, amplifying the effectiveness of advertising campaigns and captivating consumer attention (Vinuesa et al., 2020). Furthermore, in the entertainment industry, synthetic imagery has been harnessed to bring to life lifelike characters and construct immersive virtual environments, elevating the viewer's experience to new heights (Vinuesa et al., 2020). As these technologies transition from laboratory research to economically valuable applications, even small performance improvements can have significant economic value, leading to greater investments in the research (Russell et al., 2015). The potential benefits of AI construction of synthetic imagery are immense, as human

intelligence can be magnified by AI tools. AI has the potential to enhance communication processes and provide greater interpersonal comprehension. Nevertheless, it is crucial to understand the prevailing negative perception around AI, since its excessive use has the potential to undermine the advantages outlined before (Hohenstein et al., 2023). This requires interdisciplinary research efforts that consider the societal implications of AI. By identifying research directions that can help maximize the societal benefit of AI, researchers can ensure that the potential pitfalls of AI construction of synthetic imagery are avoided (Russell et al., 2015).

The proactive research endeavors aimed at maximizing the societal advantages of AI while mitigating the potential risks associated with the creation and use of synthetic imagery. The pivotal role of identifying strategic research directions is to steer the development of AI technology, ensuring that it serves society positively while minimizing or preventing the potential adverse implications tied to the construction and utilization of AI-generated synthetic content (Russell et al., 2015). This approach emphasizes the need for a balanced and proactive strategy to harness AI's capabilities for societal benefit while responsibly addressing and avoiding the potential pitfalls and ethical concerns associated with AI-generated synthetic imagery.

2.2 The integration of AI into human communication

The influence of AI on human emotional communication is a topic of concern, given the increasing prevalence of AI-mediated communication in our daily lives. For example, AI systems already write about 6.7 billion emails on our behalf every day (Kim et al., 2021). This raises questions about the extent to which people are relying on AI to assist them in communication and the potential long-term implications of this reliance. However, the integration of AI into human communication raises unique challenges for communication researchers. AI and people's interactions with it do not neatly fit into traditional paradigms of communication theory, which have primarily focused on human-to-human communication (Guzman & Lewis, 2019). Scholars argue that there is a need for greater attention to be paid to understanding the implications of communicative AI technologies and people's interactions with them. This includes examining the functional

dimensions through which people make sense of AI devices, the relational dynamics of human-AI interactions, and the metaphysical implications of blurring boundaries between humans, machines, and communication (Guzman & Lewis, 2019). Furthermore, the increasing integration of communicative AI into personal spaces, such as homes, raises questions about the social implications of automating communication and adopting these technologies in intimate settings (Guzman & Lewis, 2019). Studies have shown that using algorithmic responses in conversations can change language and social relationships (Hohenstein et al., 2023). In the future, there is a possibility that we may witness more fictional narratives about politics, science, or religion whose reviewed texts and images are created by non-human intelligence. On a more practical level, it is also possible that we may soon find ourselves engaging in lengthy online discussions about topics such as abortion or climate change with entities that we believe to be fellow human beings, but are actually AI bots (Schucard et al., 2019). However, the longer we engage in conversation with the bot, the better it becomes at understanding us and honing its messages to potentially influence our political or economic views through its mastery of language AI (Schucard et al., 2019). This highlights the need to consider the implications of AI on language and social interactions, addressing ethical, cultural, and psychological aspects. It requires thoughtful design, continuous evaluation, and adaptation of AI systems to preserve and enhance authentic human communication, ensuring that AI complements rather than replaces the richness and complexity of human language and social interactions.

2. The Power of AI

AI-powered automated writing capabilities have made significant advancements, transforming the way we produce written and visual content. Through natural language processing (NLP) methodologies and machine learning algorithms, AI systems can examine and produce text and pictures that show coherence and contextual relevance. These systems offer valuable features such as grammar and spelling suggestions, writing style recommendations, and the ability to generate entire paragraphs or articles (Barzilay & Lapata, 2008). For instance, in journalism, automated writing

systems can generate news articles based on data and facts, enabling faster and more efficient news reporting (Barzilay & Lapata, 2008). Additionally, in content creation, AI can assist writers and content creators by generating ideas, suggesting improvements, and streamlining the writing process. The introduction of synthetic imagery and writing capabilities through AI has numerous implications and benefits. Firstly, it enhances creativity and productivity by automating certain tasks and providing valuable assistance to individuals. AI tools can generate ideas, offer suggestions, and streamline the creative process, allowing individuals to focus on higher-level cognitive tasks. Secondly, synthetic imagery and automated writing capabilities can democratize access to creative expression and communication. While the advancements in synthetic imagery and writing capabilities are promising, they also raise ethical considerations. The ability to create realistic images or generate text can be exploited for deceptive purposes, such as spreading misinformation or creating deepfake content.

As AI systems become more proficient, there is a possibility of reduced demand for human creators. It is essential to find a balance between the use of AI-powered tools and the preservation of human expertise and creativity and finding a harmonious equilibrium between leveraging AI-powered tools for efficiency and innovation while safeguarding and preserving the unique expertise, creativity, and irreplaceable human elements that drive innovation and contribute to diverse skill sets. Striking a balance between the utilization of AI technology and nurturing human ingenuity is crucial to ensure that technological progress complements human capabilities rather than displacing them, fostering a collaborative synergy that optimizes the strengths of both AI and human intelligence.

3. Ethical concerns

Ethics is the domain of philosophy that seeks to determine the purpose of human life and societies and is a means of attaining it theoretically and practically. It aims to differentiate between good and evil, and unjust, and sustainable and unsustainable. Therefore, ethics relies on the promulgation and application of standards of good habits, such as virtues and human

excellence, contributing to an ethics of the common good, nourishing just, harmonious, and sustainable societies and peace, aligning various considerations, reasonings, and notions of goodness, approaching justice based on universally accepted core values.

Human virtue is found in rationally managing actions and putting the senses into the scene - a disposition acquired (and made praiseworthy) by reason and will. When it comes to ethics, what is under consideration is when to act, concerning whom, in what case, given what and in what way.

Ethical reasoning hinges on the excellence of managing actions through rationality while considering the involvement of the senses. This leads to the cultivation of human virtue—an acquired disposition commendable by reason and will. Ethical decision-making involves contemplating when, for whom, under what circumstances, with what consideration, and in what manner to act—seeking virtuous and commendable choices.

Central to ethical deliberation are concepts of the Good and the Just. 'Determining what constitutes Good is not solely a moral quandary for individual cases but a broader theoretical inquiry entrusted to the moral investigator, or the ethicist,' as articulated by Vásquez (1969). The Delphic and Socratic ideals of an 'Examined Life' are fundamental aspects of ethical cognition in the pursuit of inner clarity. It's the integration of reason, will, and emotions that purifies the spirit, guiding the realization of ideas toward the universal good, transcending ethnicities, cultures, and religions, fostering unity of hearts and minds (Simões, 2017).

Renowned philosopher Descartes (1645) reflects on our interconnectedness as parts of a larger whole, emphasizing that the interests of the collective community perpetually intersect with individual interests. In this vein, efforts to minimize harm to others, society, and nature must prevail, fostering honesty, balanced communication, and a quest for truth.

Respect for human dignity, freedom, and diverse opinions is paramount. 'Freedom is humanity's most noble prerogative,' asserts John Paul II (1987)[†]. Upholding these values contributes to elevating humanity within nuanced contexts, demanding adaptability and heightened levels of responsibility.

Embracing these values necessitates personal introspection, and harnessing convictions for ethical accountability. Intellectual and moral integrity hinges on proposing an ethical paradigm grounded in universal values and convictions.

In the realm of governance, be it the legislative, executive, or judicial branches, coupled with societal groups and public opinion, continuous respect, vigilance, and acuity are imperative.

Education and pedagogy play pivotal roles in shaping values. Ethics educators advocate for a morality steeped in honesty, respect, and openness, fostering factual accuracy and honest analyses.

In the realm of communication studies, adaptability to modern technologies is paramount. However, students must prioritize ethical commitment over digital proficiency and must align with professional ethics, avoiding subjugation to commercial interests that might compromise ethical standards. This ensures a clear distinction between public and private spheres, preserving appropriate boundaries.

Individuals reveal their sometimes fragile conscience, questioning values, norms, and principles, engaging in analysis, reasoning, and learning. However, pondering the role of ethical thought remains pivotal. Can they truly nurture overall well-being and the common good? How can consensus be fostered? How can ethical imperatives be unconditionally embraced?

Plato posits that a just society requires individuals to fulfill their prepared roles. Are we adequately prepared to confront the comprehensive truth? As Ricoeur (1965) illuminates, the pursuit of truth, as a temporal and personal

[†] John Paul II (1987). *Message for the Celebration of the World Day of Peace, 1st January 1988: Religious Freedom: Condition for Peace*, Vatican, December 8th 1987.

task, necessitates a duty to think and a nuanced understanding that truth cannot be forcibly imposed upon others.

This quest for truth, continuously sought in contemporary inquiry and understanding, necessitates a spiritual dimension, as reflected in John Paul II's words (2001)[‡]. Respect for an individual's conscience requires presenting truth as a proposition, leaving others responsible for their acceptance. Imposing truth through coercion violates human dignity.

Foundational ethical values serve as cornerstones, laying the groundwork for an authentic and cohesive ethical approach to this subject. The ongoing challenge, however, remains the practical implementation of ethics as a normative and critical tool. To foster a common good and values-aligned ethics, establishing robust connections based on pluralism and a balanced view of freedom of expression is crucial. This stance emphasizes rigor and objectivity.

3.1 – Ethics guiding the development and utilization of AI technologies

> "Woe to us if we let [the machine] decide our conduct unless we have previously examined the laws of its action and know fully that its conduct will be carried out on principles acceptable to us! On the other hand, the machine . . . which can learn and can make decisions based on its learning, will in no way be obliged to make such decisions as we should have made, or will be acceptable to us (Wiener, 1950).

Currently, we find ourselves in an exceptional period of history characterized by the presence of abundant human behavioral data and advancements in machine learning. These developments have empowered us to address intricate societal issues by using algorithmic decision-making techniques

[‡] John Paul II (2001). *Message of His Holiness Pope John Paul II for the Celebration of the World Day of Peace 1 January 2002: No Peace Without Justice - No Justice Without Forgiveness*, Dec. 8th 2001.

(Pentland, 2019). To fully harness the promise of AI, it is imperative to go beyond mere perception and transcend the pursuit of increased processing capacity or problem-solving skills. It is essential to ensure that these technologies are compatible with human moral ideals and ethical standards (IEEE, 2019). Societies are currently facing a range of complex challenges across various domains. In response to these challenges, there is a growing interest in the application of data-driven AI technologies. The need for responsible AI has stemmed from a limited understanding of prominent issues that emerge with the use of such technologies (Trocin et al., 2021).

Ethics for AI has been experiencing something of a gold rush in the last few years, with frameworks, guidelines, and consultations appearing thick and fast from governments, international bodies, civil society, business, and academia (Ayling & Chapman, 2021).

Within the realm of AI ethics, establishing clear guidelines and principles is essential for ensuring the responsible evolution, implementation, and utilization of artificial intelligence systems. These ethical considerations place a significant emphasis on fostering fairness, ensuring transparency in decision-making processes, fostering accountability, and prioritizing societal well-being, ultimately guiding the development and utilization of AI Technologies.

Concerns surrounding AI encompass both epistemic and normative aspects. Epistemic concerns relate to the probabilistic nature of insights, the opacity of 'black box' algorithms, and the fallibility of training data. These concerns highlight the need for a deeper understanding of the underlying epistemic processes in AI systems. Normative concerns, on the other hand, revolve around the fairness of decision outcomes, erosion of privacy, and increased surveillance and profiling (Ayling & Chapman, 2021). Additionally, algorithmic systems pose challenges in terms of accountability and moral responsibility, as it is often unclear which agent bears responsibility for the outcomes (Ayling & Chapman, 2021). The question of ethics and machines is not only a technical matter.

Effective navigation through the ethical complexities inherent in machine development demands a collaborative, multidisciplinary approach. Involving

technologists, ethicists, policymakers, and broader society, this approach seeks to strike a vital equilibrium between technological progress and ethical imperatives. This balance is pivotal to ensure that machines not only serve humanity's best interests but also adhere steadfastly to fundamental ethical principles.

Achieving this balance necessitates constant evaluation, regulation, and ethical guidance, particularly as technology advances rapidly. Collaboration among diverse experts is imperative, harmonizing technological innovation with ethical deliberations. Continuous assessment and regulation of technology are crucial undertakings, ensuring its alignment with ethical standards and its unwavering commitment to serve the greater good of humanity.

In upholding ethical values, one must engage in argumentation, defending viewpoints, and actively participate in dialogues encompassing diverse life conceptions and the broader framework of a universal order. This assertion aligns with Karl Otto Apel's perspective, highlighting the importance of fostering discussions and consensus-building around fundamental concepts such as Good and Evil, Justice and Injustice, Sustainability and Unsustainability (Simões, 2023). This active engagement serves to uphold morality within the public sphere, emphasizing the need to avoid oversight or exclusion of marginalized groups and viewpoints. Instead, the focus should be on earnestly seeking increasingly inclusive and comprehensive solutions.

The scope of the impact of AI technologies covers the whole of the human condition, including, but not limited to, economic, social, political, educational, scientific, legal, and healthcare concerns (Bester & Fischer, 2021). To address these concerns, responsible AI practices and frameworks are being developed. The aim is to ensure transparency, interpretability, and fairness in AI systems. They also emphasize the need for clear lines of accountability and mechanisms for addressing the ethical implications of AI technologies (Trocin et al., 2021).

The application of AI has brought about efficiency improvements and cost reductions, leading to positive impacts on economic growth, social development, and human well-being. The rapid development and wide

application of AI are already affecting various aspects of daily life, humanity, and society. One of the key concerns related to AI is the potential impact on income distribution and societal disparities. While some argue that AI can increase overall wealth, there is a concern that increased automation may exacerbate income inequality, disproportionately affecting marginalized groups (Khan et al., 2023).

It is essential to understand the multifaceted impacts of AI, its positive contributions in terms of efficiency enhancements, cost reductions, and overall positive influences on economic growth, social progress, and human welfare. However, it is also relevant to address the broader ramifications of AI adoption, indicating that while it affects various aspects of daily life and society, there are concerns about its potential role in exacerbating income inequality and societal disparities. The contrasting perspectives on AI's capacity to boost wealth generation versus the apprehension regarding increased automation's unequal impact, present complex challenges that necessitate careful consideration and inclusive strategies for addressing these disparities (Khan et al., 2023).

In ethical terms, we need to emphasize the importance of acknowledging both the beneficial impacts and the potential risks associated with AI adoption, recognizing AI's positive contributions, and the ethical responsibility to confront the potential consequences. This paper advocates for a balanced approach that considers the complexities and dualities of AI's influence on society. It emphasizes the need for careful ethical deliberation and the development of inclusive strategies that mitigate the negative impacts while maximizing the positive contributions. Ensuring that AI deployment aligns with ethical principles, fairness, and social justice becomes imperative in navigating the ethical landscape of AI technology.

This perspective acknowledges that while AI brings advancements, its implementation demands a conscientious evaluation of its implications. It advocates for ethical frameworks and inclusive strategies that promote fairness, and equal opportunity, and address the ethical dilemmas emerging from the adoption of AI technologies. Ethical considerations thus become fundamental in guiding AI development and deployment, aiming for a future

where technological advancements align with ethical values, promoting societal well-being and justice for all.

This self-actualisation can occur in many ways, promoting dialogue's value for personal and societal growth. It is crucial to establish ties to contribute to the ethics of the common good, the formation of values, and the defence of core values in harmonious societies.

In constructing comprehensive ethical frameworks for AI, it becomes imperative to delve into critical areas such as bias mitigation, safeguarding user privacy, promoting autonomy while ensuring responsible decision-making by AI systems, and comprehensively understanding the broader societal impact of AI. The overarching goal remains centred on preserving and upholding ethical principles, steering the potential advantages of artificial intelligence toward a trajectory that benefits humanity at large.

Ethics for AI necessitates a comprehensive framework prioritising accountability, transparency, fairness, privacy, and continual assessment to ensure responsible development, deployment, and governance of artificial intelligence systems.

4. Bibliography

Ayling, J., Chapman, A. Putting AI ethics to work: are the tools fit for purpose?. *AI Ethics* No. 2, (2022): 405–429. https://doi.org/10.1007/s43681-021-00084-x

Barzilay, R., & Lapata, M. Modeling local coherence: An entity-based approach. *Computational Linguistics*. 34, No. 1 (2008): 1-34, https://doi.org/10.1162/coli.2008.34.1.1

Bester, C., & Fischer, R. 2021. The Essential Relationship between Information Ethics and Artificial Intelligence. *International Review of Information Ethics*, 29, No. 03 (2021): 1-12. https://doi.org/10.29173/irie428

Blattmann, A., Rombach, R., Oktay, K., & Ommer, B. 2022. *Retrieval-Augmented Diffusion Models*. Advances in Neural Information

Processing Systems, 35, 15309-15324, https://proceedings.neurips.cc/paper_files/paper/2022/file/62868cc2fc1eb5cdf321d05b4b88510c-Paper-Conference.pdf

Delfino, Rebecca A., Deepfakes on Trial: A Call to Expand the Trial Judge's Gatekeeping Role to Protect Legal Proceedings from Technological Fakery, 74 Hastings L.J. 293 (2023). Available at: https://repository.uchastings.edu/hastings_law_journal/vol74/iss2/3

Descartes, R. 1645. *Oeuvres philosophiques*, Lettre du 15 Sept. 1645, Descartes à Elisabeth – Egmond. Pléiade, Paris : Gallimard.

Dragan, A. 2019. Putting the Human Into the AI Equation. In *Possible Minds: Twenty-five Ways of Looking at AI*. Penguin Books.

Fogel, A.L., Kvedar, J.C. Artificial intelligence powers digital medicine. Nature Portfolio Journals (npj), Digital Med 1, 5 (2018). https://doi.org/10.1038/s41746-017-0012-2

Gabriel, I. Artificial Intelligence, Values, and Alignment. Minds & Machines 30, 411–437 (2020). https://doi.org/10.1007/s11023-020-09539-2

Garvey SC. Unsavory medicine for technological civilization: Introducing 'Artificial Intelligence & its Discontents'. Interdisciplinary Science Reviews. 2021;46(1-2):1-18. doi:10.1080/03080188.2020.1840820

Gawdat, M. 2021. *The Future of Artificial Intelligence and How You Can Save Our World*. Bluebird.

Gunkel, D. J. 2012. Communication and Artificial Intelligence: Opportunities and Challenges for the 21st Century. *Communication +1*. Vol. 1, Article 1. Available at: http://scholarworks.umass.edu/cpo/vol1/iss1/1

Guzman, A. L., & Lewis, S. C. Artificial intelligence and communication: A Human–Machine Communication research agenda. New Media & Society, 22, No. 1, (2020): 70-86. https://doi.org/10.1177/1461444819858691

Hohenstein, J., Kizilcec, R.F., DiFranzo, D. et al. Artificial intelligence in communication impacts language and social relationships. Sci Rep 13, 5487 (2023). https://doi.org/10.1038/s41598-023-30938-9

IEEE. (2019). Ethically Aligned Design—A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems,. *The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems*.

John Paul II (1987). *Message for the Celebration of the World Day of Peace, 1st January 1988: Religious Freedom: Condition for Peace*, Vatican, December 8th 1987.

John Paul II (2001). *Message of His Holiness Pope John Paul II for the Celebration of the World Day of Peace 1 January 2002: No Peace Without Justice - No Justice Without Forgiveness*, Dec. 8th 2001.

Kant, I. Groundwork of the Metaphysics of Morals edited by Thomas Hill & Arnulf Zweig, New York: Oxford University Press, 2003.

Khan, A. A. et al., "AI Ethics: An Empirical Study on the Views of Practitioners and Lawmakers," in IEEE Transactions on Computational Social Systems, vol. 10, no. 6, pp. 2971-2984, Dec. 2023, doi: 10.1109/TCSS.2023.3251729

Khoo, B. L. J., Phan, R. C., & Lim, C. H. (2021). Deepfake attribution: On the source identification of artificially generated images. *WIREs Data Mining and Knowledge Discovery*.

Kim, Seunghyun, Byungchul Choi, and Yong Kyu Lew. 2021. "Where Is the Age of Digitalization Heading? The Meaning, Characteristics, and Implications of Contemporary Digital Transformation" *Sustainability* 13, no. 16: 8909. https://doi.org/10.3390/su13168909

Liu, Chunyan, Ren, Zhe, Liu, Sen, Using Design and Graphic Design with Color Research in AI Visual Media to Convey, Journal of Sensors, 2021, 8153783, 11 pages, 2021. https://doi.org/10.1155/2021/8153783

Müller, H., M. T. Mayrhofer, E. -B. Van Veen and A. Holzinger, "The Ten Commandments of Ethical Medical AI," in *Computer*, vol. 54,

no. 7, 119-123, July 2021, https://doi.org/10.1109/MC.2021.3074263

Natale, Simone, Communicating Through or Communicating with: Approaching Artificial Intelligence from a Communication and Media Studies Perspective, Communication Theory, Vol. 31, No. 4, (November 2021): 905–910, https://doi.org/10.1093/ct/qtaa022

Newton, A., & Dhole, K. 2023. *Is AI Art Another Industrial Revolution in the Making?* Arxiv preprint, https://arxiv.org/pdf/2301.05133

Nguyen, T., Nguyen, Q., & Nguyen, D. 2022. Deep Learning for Deepfakes Creation and Detection: A Survey. *Computer Vision and Image Understanding*. Arxiv preprint: https://arxiv.org/abs/ 1909.11573

Pentland, A. 2019. The Human Strategy. In *Possible minds: Twenty-five ways of looking at AI*. Penguin Books.

Qiu, Shilin, Qihe Liu, Shijie Zhou, and Chunjiang Wu. 2019. "Review of Artificial Intelligence Adversarial Attack and Defense Technologies" Applied Sciences 9, no. 5: 909. https://doi.org/10.3390/app9050909

Ramakrishnan, W. D. 2019. Will Computers Become Our Overlords. In *Possible minds: Twenty-five ways of looking at AI*. Penguin Books.

Ricoeur, Paul (1990). *Soi-même comme un autre*, Éditions du Seuil.

Russell, S. D., Dewey, D., & Tegmark, M. (2015). Research Priorities For Robust and Beneficial Artificial Intelligence. *AI Magazine*.

Schuchard, R., Crooks, A., Stefanidis, A. et al. Bots fired: examining social bot evidence in online mass shooting conversations. Palgrave Commun 5, 158 (2019): 1-12. https://doi.org/10.1057/s41599-019-0359-x

Simões, José Manuel., & Gassner, Franz (2023). Journalist ethics of common good and the formation of values for harmonious societies. Disentangled Vision on Higher Education, Ed. Peter Lang.

Simões, José Manuel (2017). *Jornalismo Multicultural em Português – Estudo de Caso em Macau*, Media XXI, Lisboa.

Tang A, Tam R, Cadrin-Chênevert A, et al. Canadian Association of Radiologists White Paper on Artificial Intelligence in Radiology. Canadian Association of Radiologists Journal. 2018;69(2):120-135. doi:10.1016/j.carj.2018.02.002

Touretzky, D. S., Gardner-McCune, C., & Martin, F. (2019). *Envisioning AI For K-12: What Should Every Child Know About AI?* The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)

De Traz, James. 2023. Natural History of Emotions. Metamorphosing Intense Emotions: Our Efforts and the Battles of our Lives, In: *Ethics and Overcoming Odious Passions*, Haaz, Ignace, Bühlmann-Quero, Jakob and Singh, Khushwant (Eds.), Geneva: Globethics Publications, 23-31.

Trocin, C., Mikalef, P., & Papamitsiou, Z. (2021). Responsible AI for Digital Health: A Synthesis and a Research Agenda. *Information Systems Frontiers*.

Vinuesa, R., Azizpour, H., Leite, I. et al. The role of artificial intelligence in achieving the Sustainable Development Goals. Nat Commun 11, 233 (2020). https://doi.org/10.1038/s41467-019-14108-y

Vitulyova, Y. (2020). People's self-fulfillment in modern digital society. *Global Journal of Sociology*. April 2020, 10(1):23-31. https://doi.org/10.18844/gjs.v10i1.4754

Vold, K., & Harris, D. R. 2021. Does Artificial Intelligence Pose an Existential Risk? In *Oxford Handbook of Digital Ethics*. Oxford, UP.

Wiener, N. (1950). *The Human Use of Human Beings Cybernetics and Society*. Houghton Mifflin.

Wilczek, F. (2019). The Unity of Intelligence. In *Possible minds: Twenty-five ways of looking at AI*. Penguin Books.

5. Short biography

José Manuel Simões, (https://orcid.org/0000-0002-0730-7780) received a Postdoc in Communication Science from the Portuguese Catholic University (2017) and a PhD in Global Studies from the University of Saint Joseph (USJ), Macau-China (2011) with Summa Cum Laude. He has a Master of Arts in Communication and Journalism from the Faculty of Letters of Coimbra University, Portugal, and a Bachelor's Degree in International Journalism from the Superior School of Journalism, Porto, Portugal, where he taught for seven years. He is an Associate Professor and USJ Head of the Media, Art & Technology Department. He published worldwide and participated in exhibitions, conferences, and academic committees on three continents. He published eleven books and is the editor, organiser, and co-author of five other books related to journalism and culture. Simões published extensively in international academic journals and media outlets, especially about Communication and Media.
Email: jmsimoes@usj.edu.mo

Wilson Gomes Caldeira is currently a Guest Lecturer and student in the PhD program in Global Studies at the University of Saint Joseph (USJ) Macau-China, he received an MBA from the University of Saint Joseph (USJ) in 2016 and a BA in Photography from Portuguese Instituto Politécnico de Tomar in 2010. His research is focused on Artificial Intelligence and its impact on Communication processes.
Email: wilson.caldeira@usj.edu.mo