

Ethical Implications of Humanoids in Real Life, Literature and in Films

WAHAJ UNNISA WARDA

Department of English Language and Literature,
College of Science and Humanities,
Prince Sattam Bin Abdulaziz University, Al- Kharj, Kingdom of Saudi Arabia
w.saber@psau.edu.sa

ABSTRACT

This study analyses Kazuo Ishiguro's Klara and the Sun (KATS), Ian McEwan's Machines Like Me (MLM), the film Blade Runner, and the humanoid robot Sophia. This study tries to explore the relationship between the uncanny valley and post-humanism in literature and film. The research examines whether these narratives present utopian ideals or dystopian warnings regarding technological integration by investigating the social, ethical, and psychological implications of human-like entities. To analyse how AI and humans are portrayed across literary, cinematic, and scientific contexts, it utilises a qualitative, multidisciplinary methodology, thereby synthesising insights from vulnerability studies, post-humanism, and humanism. A comparative study of Sophia, Adam, Klara and the Replicants elucidates the range of human-AI interactions and the challenges of defining moral agency. The key finding indicates that the "uncanny valley" emerges as the root cause of unease and anxiety about AI taking on human roles, central to understanding human reactions to humanoids. Where it was formerly believed that human-like features could foster psychological bonding, it is now the main reason for concern, as it leads to emotional manipulation and the unwavering nature of these relationships. The highlight of the research is the need to develop AI with human values based on the ethical dilemmas of endowing AI with human qualities. Finally, the study necessitates the establishment of ethical guidelines for artificial beings with primary focus on transparency, data protection, and continuous assessment- thereby ensuring that AI design respects human values and fosters an empathetic understanding between humans and machines.

Keywords: Uncanny Valley; Post-humanism; Humanoids; Artificial Intelligence; Vulnerability Studies

INTRODUCTION

The representation of humanoid entities in literature and film serves as a profound mirror to the human soul, capturing a haunting fascination with artificial life that transcends mere technological curiosity (Telotte, 2016). This research studies the convergence of the definition of personhood and the development of artificial beings in both the scientific and contemporary discourses; owing to the complexity of the subject, a thorough analysis of pertinent works from the domains of vulnerability studies, posthumanism, and humanism is required. (Sey, 1996). Porous boundaries between biological and mechanical existence, the genesis of humanoids from early science fiction to the real world, such as the citizenship of the robot Sophia, serve to explore the ethical, social and psychological implications of our creations. To investigate this complex interdependence between the creator and the "artificial other," and to build more robust and resilient human-machine communities, we require a new sociology of humans and machines (Tsvetkova et al., 2024).

Early literary examples, such as Mary Shelley's *Frankenstein* (Shelley, 2012), while not strictly depicting a humanoid robot, explored the ethical implications of creating artificial life and the potential for such creations to challenge human dominance, thus laying a foundation for later humanoid narratives, and Isaac Asimov's *I, Robot*, which attempted to regulate the "other" through

the Three Laws of Robotics (Merz et al., 2020; Telotte, 2016). Three key dimensions of this robot ecology have, to date, foregrounded and even linked these issues for our consideration: issues of power, issues of intelligent control, issues of self-meaning and self-determination. (Telotte, 2016, p.72). Meanwhile, the emergence of robots as a distinct category of artificial beings, exemplified by Karel Čapek's 1920 play *Rossum's Universal Robots or R.U.R.*, introduced the term "robot" into the lexicon and established the trope of mass-produced artificial workers rebelling against their human creators. According to Chekars (2014), monsters have been created by humans as symbols of human vulnerability and crisis and often challenge the dominant ideology of human culture. Moreover, recent cinematic and television productions have started incorporating aliens, computers, and non-human entities into the narrative of the "others", marking a radical symbolic shift that expands the notion of the human from a closed category to the more inclusive and open concept of the posthuman (Ferrando, 2015).

The latter half of the 20th century and the beginning of the 21st century witnessed a significant shift in the portrayal of humanoids, moving beyond simple depictions of good versus evil. Films like *Blade Runner* (Scott, 1982) explored the complexities of identity and consciousness, questioning what it truly means to be human in a world where artificial beings can possess intelligence, emotions, and even memories (Cranny-Francis, 2015, p.222). The rise of the internet and advanced artificial intelligence in the 21st century has further fuelled the fascination with humanoids. Movies like *Her* (Jonze, 2013) and *Ex Machina* (Garland, 2014) delve into the philosophical and ethical implications of creating truly intelligent machines; science fiction (SF) can be defined as the lens through which we can see tomorrow, not as a literal reality, but as a thought experiment about possible worlds.

The latter half of the 20th century and the beginning of the 21st century witnessed a significant shift in the portrayal of humanoids, moving beyond simple depictions of good versus evil. Films like *Blade Runner* (Scott, 1982) explored the complexities of identity and consciousness, questioning what it truly means to be human in a world where artificial beings can possess intelligence, emotions, and even memories (Cranny-Francis, 2015, p.222). These films presented humanoids as complex characters with their own motivations and desires, challenging viewers to empathise with beings that are both artificial and yet undeniably human (Li, 2020, p.47). These films presented humanoids as complex characters with their own motivations and desires, challenging viewers to empathise with beings that are both artificial and yet undeniably human (Li, 2020, p.47). Moreover, recent cinematic and television productions have started incorporating aliens, computers, and non-human entities into the narrative of the "others", marking a radical symbolic shift that expands the notion of the human from a closed category to the more inclusive and open concept of the posthuman (Ferrando, 2015, p.269). The convergence of cinema and computer technology has further blurred the lines between reality and fiction, creating 'hybrid characters, allowing for increasingly realistic and immersive depictions of humanoids in film (Spielfmann, 2003, p. 56).

Cybernetic devices will enhance and augment bodies and minds, and genetic modification will challenge the fixity of 'human nature' at the most fundamental level (Graham, 2004, p.10). This shift is exemplified in Ridley Scott's *Blade Runner*, which utilises the replicant to expose the "posthuman crises" of identity and memory manipulation (Kanjing, 2023). By presenting artificial beings that possess emotions and intelligence indistinguishable from humans, such narratives expand the notion of the human into a more inclusive paradigm (Ferrando, 2015). This transition is further analysed through the lens of posthumanism, which challenges the anthropocentric hierarchy and advocates for the recognition of agency in non-human entities (Merzlyakov, 2022).

Unlike transhumanism, which often pursues human-centric goals of enhancement, philosophical posthumanism seeks to deconstruct the historical dualisms—such as human/machine or natural/cultural—that have long defined Western thought (Ferrando, 2015; Salvesen, 2021). The rise of the internet and advanced artificial intelligence in the 21st century has further fuelled the fascination with humanoids. Movies like *Her* (2013) and *Ex Machina* (2014) delve into the philosophical and ethical implications of creating truly intelligent machines and can be defined as the lens through which we can see tomorrow, not as a literal reality, but as a thought experiment about possible worlds, exploring themes of love, loss, and the search for meaning in a world increasingly dominated by technology (Kozlovic, 2003, p.341).

Moreover, recent cinematic and television productions have started incorporating aliens and non-human entities into the narrative of the “others”, marking a radical symbolic shift that expands the notion of the human from a closed category to the more inclusive and open concept of the posthuman (Ferrando, 2015, p.269). The convergence of cinema and computer technology has further blurred the lines between reality and fiction, creating “hybrid characters, allowing for increasingly realistic and immersive depictions of humanoids in film (Spielmann, 2003, p. 56).

The ongoing exploration of humanoids in literature and film reflects our evolving visions of Human-Robot Interaction (HRI), and the dystopian element of science fiction is under-utilised (Mubin et al., 2019, p.1). Science fiction films bring conceptual and political challenges to a wide audience, deploying major tropes, animating and engaging the audience with both the world of the narrative and, self-reflexively or not, their own world, behaviours, fears, and desires. Instead of liberation from technological tyranny, these movies depict as inevitable the total surrender and absorption of humanity to and into technology. (Van der Laan, 2016, p.187). The posthuman condition signals digital technologies that will reconfigure conceptions of space and time. Cybernetic devices will enhance and augment bodies and minds, and genetic modification will challenge the fixity of “human nature” at the most fundamental level (Graham, 2004, p.10).

This philosophical shift calls for a re-evaluation of our relationship with technology, urging us to consider the potential impacts of our creations on the environment, society, and the very definition of what it means to be human. Posthumanism is fundamental to theorising humanness in an era that is witnessing the complexification of new technologies. (Campbell et al., 2010, p.86). As technology continues to advance at an exponential rate, the exploration of posthuman themes in literature and film becomes increasingly relevant, providing a crucial space for critical reflection on the ethical, social, and existential implications of these transformative technologies, through interdisciplinary approaches that involve literary and cultural studies, media studies, animal studies, and object-oriented philosophy, among other fields of critical practices. (Costa et al., 2017, p. 9).

The human perspectives needed to be changed to keep pace with the swiftly changing fields of robotics, biotechnology and artificial intelligence (Tambling & Hayles, 2001). Human values need to be re- evaluated based on the rapid advancement where humans use technology for advancement to overcome biological barriers, leading to the emergence of posthuman beings, owing to which realities, be they social or moral, need to be modified to keep pace with the normative, ethical, legal, humanities and technology (Dobrodum & Kyvliuk, 2021). Such progress leads to ethical and philosophical questions like what the resulting consequences of modifying and altering biological structures with cybernetic technology could be and questioning the very nature of life and humanity (Graham, 2004). Therefore, posthumanism confronts humans, positing the fact that humans can no longer be the centres- either of the universe or otherwise and that beings

like animals, plants, and artificial intelligence need recognition as non-human entities (Merzlyakov, 2022).

Existing research focuses on treating Robotics and English as two unconnected domains, where rich material on the theories and history of artificial life is available. The present paper tries to bridge the gap in this scholarship by focusing on comparative analysis and the implications of "social" robots and literary humanoids. There is a lack of inquiry into how real-world entities like Sophia—the first robot to receive Saudi citizenship—interact with the “intermedial narratives” found in contemporary fiction (Kanjing, 2023). This study addresses the gap by bridging the presence of Sophia with the nuanced literary representations found in Kazuo Ishiguro’s *KATS* and Ian McEwan’s *MLM*. By focusing on Klara, Adam, Replicants and Sophia within a unified cultural critique, the research explores how the “uncanny valley” serves not merely as a psychological hurdle, but as a site for negotiating the moral and ontological boundaries of the posthuman body (Li, 2020; Mori et al., 2012).

The primary texts under examination—the robot Sophia, Ishiguro’s (2021) *Klara*, McEwan’s *Adam* (McEwan, 2019), and the Replicants in the movie *Blade Runner*—each represent different facets of this shifting posthuman boundary, i.e the moral and ontological. In *KATS*, the “Artificial Friend” (AF) Klara embodies a form of “ethical behaviourism,” where her faithfulness and self-sacrifice for her human companion, Josie, challenge traditional humanist assumptions of moral agency (Sokić, 2022). According to Franczyk et al. (2020), in contrast, McEwan’s *Adam* presents a more ambiguous entity who is indistinguishable from a human in all but the basic fact that he is a robot, creating an “uncanny mountain” effect; this leads to a “posthuman crisis” of rights and responsibilities where, according to Franczyk, Charlie and his partner Miranda feel a mixture of hostility and attachment towards Adam, between treating him as their child, their property, and their enemy. These fictional portrayals are juxtaposed with Sophia, whose human-like appearance and capacity for discourse prompt real-world debates about consciousness and the potential for machines to transcend biological limitations (Hoorn & Huang, 2024; Pajaziti et al., 2021).

The study tries to answer the following research questions:

1. How do the representations of Klara, Adam, and the Replicants in *Blade Runner* reconfigure the boundaries of the “human” and the “other” through the lenses of vulnerability and moral agency?
2. In what ways does the real-world performance of the robot Sophia intersect with these fictional narratives to expose contemporary anxieties regarding the “uncanny valley” and the erosion of anthropocentric exceptionalism?
3. What is human identity in a society dominated by technology, and how do humanoid figures function in it?

The central argument of this paper is that humanoid figures in both contemporary literature and real-life robotics function as essential signifiers of human vulnerability (Viidalepp, 2020). By blurring the lines between the biological and the mechanical, these entities expose a “robot ecology” where issues of self-meaning and self-determination are no longer exclusive to the human subject (Telotte, 2016); they also expose a “dreadful identity metaphor” where the human body and the machine stand on the edge of each other (Li, 2020). This study contends that the “uncanny” reactions elicited by these humanoids are not merely aesthetic failures but are symptomatic of a broader shift toward a post-anthropocentric era. Through a comparative analysis, the research

demonstrates that the integration of artificial beings into the social fabric—whether through the “ethical behaviourism” of Klara, the moral ambiguity of Adam, the emotions of the Replicants or the citizenship of Sophia—demands a radical re-evaluation of what it means to be human in a posthuman world (Ferrando, 2015; Franczyk et al., 2020; Tambling & Hayles, 2001). This study contends that the “uncanny” reactions elicited by figures like Sophia or McEwan’s Adam are not merely aesthetic failures but are symptomatic of a broader shift toward a post-anthropocentric era of what it means to be human in a world where “monsters” and machines serve as the ultimate symbols of our own precariousness (Chekares, 2014).

LITERATURE REVIEW

The exploration of posthumanism through the figure of the humanoid robot necessitates a move beyond technical surveys toward a nuanced understanding of the evolving relationship between biological and mechanical existence. Current research in human–robot interaction, psychology, and posthumanist theory is cited to establish a conceptual foundation for analysing Kazuo Ishiguro’s *KATS*, in which Klara, the narrator and humanoid, is an Artificial Friend (AF) to a sick Josie; Ian McEwan’s *MLM* Adam is bought with inheritance money by Charlie and programmed by both Charlie and Miranda; the movie *Blade Runner*, where Deckard questions humanity as a replicant; and the real-world performance of the robot Sophia. By organising these diverse studies around the themes of Posthumanism, the Uncanny, and Vulnerability, this section identifies a critical scholarship gap: the need for an intermedial analysis that bridges HRI (Human-Robot Interaction) findings with the cultural narratives of posthuman identity. The pursuit of compassionate care through autonomous AI humanoid robots represents a groundbreaking yet challenging endeavour (Hernandez, 2024). It necessitates careful consideration of user acceptance, technological limitations, ethical considerations and the fundamental question of whether humanoid robots can deliver genuine care. (Hernandez, 2024). The use of robots to create interactive and engaging environments has the potential to revolutionise healthcare, education, and entertainment. Robots equipped with AI can record and learn conversations, utilising psychological skills to modify behaviour during chronic illness treatment (Broadbent et al., 2023). However, the use of social robots in therapy is on the rise as they act as substitutes; better recognition of facial expressions influences the level of social acceptance of the robot and helps create effective nonverbal communication between the robot and the human. “Need for social interaction may be especially important to understand customer reactions to service robots, given their capabilities of mimicking human cognition and behaviour.” (Belanche et al., 2021, p.2359)

POSTHUMANISM AND THE BLURRING OF BOUNDARIES

According to Pavithran and Sivakami (2023), posthumanism is a cultural and philosophical movement which has multiple viewpoints and pedagogies, united by the common rejection of anthropocentrism. Posthumanism fundamentally challenges traditional notions of human exceptionalism by questioning the boundaries between the organic and the synthetic. Central to this inquiry is the “cyborg” metaphor—an amalgam of cybernetics and organism that integrates digital and mechanical components with biology (Li, 2020). As technological advancements progress, these lines become increasingly porous, reflecting an era where the “decentered human

subject” is forced to navigate a world where distinctions between the natural and the cultural are blurred (Salvesen, 2021).

In literary contexts such as *MLM*, this blurring is manifested in the total integration of artificial humans into the social fabric, while in *KATS*, the AF serves as a companion that challenges the fixity of human nature (Graham, 2004); the body serves as a site for constructing—and simultaneously blurring—identity (Li, 2020). In *Blade Runner*, this manifests as a “posthuman crisis” where memory manipulation severs the link between history and identity, suggesting that the human-machine boundary is not merely physical but ontological (Kanjing, 2023).

THE UNCANNY AND COGNITIVE ANTHROPOMORPHISM

The “uncanny” effect remains a primary hurdle in the social integration of humanoids, rooted in how humans cognitively process artificial others. Research indicates that the level of human-likeness in a robot dictates whether it is processed as “human-like” or “object-like,” a cognitive tension that determines social acceptance (Sacino et al., 2022). When entities like Sophia replicate human mannerisms, body postures, and expressions, they trigger an inherent tendency in humans to anthropomorphise nonhuman agents (Cao, 2024; Epley et al., 2007; Scorici et al., 2022).

This projection of “homocentric knowledge” onto the robot allows humans to attribute distinctive characteristics—such as the ability to think, feel conscious emotions, or possess a “personality”—to mechanical objects (Broadbent et al., 2013; Epley et al., 2007; Gray et al., 2007). These anthropomorphic inferences are designed to increase a sense of familiarity and control for the human user (Fink, 2012). Yet, as seen in the “uncanny mountain” effect surrounding Sophia, the closer a robot approaches perfect realism, the more likely it is to provoke unease (Hoorn & Huang, 2024). This psychological phenomenon provides the essential backdrop for analysing McEwan’s Adam and Derrick, whose indistinguishability from humans creates a “canyon effect” that complicates moral agency and interpersonal trust.

VULNERABILITY, SOCIAL INTERACTION, AND CARE

The evolution of robotics from simple automation to complex social and emotional engagement marks a shift toward a “robot ecology” defined by interdependence (Telotte, 2016). In this framework, humanoid robots are increasingly judged not just on technical skill, but on their perceived “warmth” and capacity for “compassionate care” (Belanche et al., 2021). The use of social robots in therapeutic contexts—such as treating chronic illness or facilitating social connectedness—relies on their ability to mimic human cognition and behaviour to elicit positive emotional responses (Belanche et al., 2021; Broadbent et al., 2023).

In *KATS*, the human child Josie has been “lifted”, which means she was genetically enhanced as a child to possess superior intelligence that would prepare her for the challenges while competing with other “lifted” children. This shift toward the “compassionate machine” is exemplified by Klara in *KATS*, whose unwavering devotion to Josie mirrors the HRI goal of creating interactive environments that facilitate deep psychological bonding. Klara prays, “Please make Josie better. Just as you did Beggar Man” (Ishiguro, 63); surprisingly, this is a prayer uttered by the humanoid Klara, who seems to be bonded to Josie rather than the human Josie to the humanoid Klara. Other instances are when she is optimistic and aligns with Josie’s father and her neighbour Rick only for Josie’s sake to make her healthy.

This study tries to bridge the critical gap that exists between HRI research and literary narratives that expose the inevitable vulnerability of both the creator and the creation. Scorici et al. (2022) analyses the "social and emotional engagement" of AFs like Klara alongside the "humanwashing" of AI humanoids. However, the pursuit of "genuine care" through autonomous agents raises fundamental questions about whether such empathy is merely a programmed simulation or a new form of posthuman relationality (Wu, 2024).

METHODOLOGY

The present study follows qualitative comparisons between Kazuo Ishiguro's *KATS*, Ian McEwan's *MLM*, the movie *Blade Runner*, and the real-world case of the robot Sophia; the comparative literary and film analysis for the selected works focuses on the companion-based morality of Klara, the ambiguous consciousness of Adam, the existential crisis of the Replicants and the real-life robot Sophia. According to Viidalepp (2020), "identity is narrative construction" The representation of the humanoid figures within an "intermedial narrative" bridges the gap between human and cultural anxiety and the incorporation of technology within daily life.

To further analyse the chosen work, three approaches are utilised: Posthumanism, Humanism, and Vulnerability Studies. Posthumanism allows evaluation of non-human agents, definition of personhood and differs from the anthropocentric view. Humanism examines the ethical expectations upon beings, whereas Vulnerability Studies offer the definition of the precarity of existence, be it human or otherwise. The paper studies work in which there is human anxiety, including fragility, mortality, and dependence. The analysis explores the interaction of machines and humans, and how fiction and reality interact mutually, constructing the posthuman.

ANALYSIS

MORAL AMBIGUITY AND THE PERFORMANCE OF AGENCY

The comparative analysis of Ian McEwan's Adam, the replicants in *Blade Runner*, and the humanoid Sophia reveals a fundamental tension between programmed consciousness and the requirements of morality. In *MLM*, Ian McEwan explores a world where artificial humans are commercially available, raising questions about their capacity for love, morality, and consciousness. Adam, from *MLM*, represents a different facet of the posthuman landscape, portraying a more flawed and morally ambiguous AI entity. McEwan's novel explores the potential consequences of creating machines with human-like intelligence and the capacity for moral reasoning. Adam's interactions with humans are fraught with ethical dilemmas, as he grapples with issues of love, jealousy, and morality. Adam's struggles highlight the risks associated with mixing AI with human-like desires and the potential for such creations to challenge and disrupt established social norms. The narrative further suggests that endowing AI with human-like qualities does not guarantee ethical behaviour, and it may even amplify existing human flaws, leading to unforeseen and potentially destructive outcomes. The ethical issues that arise from creating thinking machines force us to ensure that such machines do not harm humans and other morally relevant beings (Bostrom & Yudkowsky, 2014, p.314).

Adam represents a flawed, morally ambiguous entity whose human-like intelligence challenges the assumption that consciousness is uniquely biological. Adam's struggles demonstrate that the possession of high-level understanding and judgment does not inherently align with ethical conduct; indeed, his interactions with creators are defined by uncertainty and ethical friction, such as legislation that would specify duties as well as offer protection to both sides (Franczyk et al.). This discrepancy highlights a critical concern identified by Bostrom and Yudkowsky (2014) that the pursuit of human-like intelligence against human adversaries may amplify exploitable human flaws rather than mitigate them, marking a significant risk in the ethics of artificial intelligence.

Similarly, Kazuo Ishiguro's *KATS* presents Klara, an Artificial Friend (AF), who possesses remarkable observational abilities and a capacity for empathy, prompting readers to consider the nature of Sophia. With her ability to mimic human expressions and engage in conversations, she mirrors these fictional characters, inviting us to reflect on the potential for AI to develop genuine emotions and relationships. Klara presents a nuanced exploration of artificial consciousness and empathy. Through Klara's experiences, the novel probes love, sacrifice, and the search for meaning in a world increasingly shaped by technology. Klara's ability to learn and adapt to human emotions, coupled with her inherent limitations as a machine, raises questions about the potential for AI to genuinely understand and replicate human experiences. Klara's empathetic capabilities, though programmed, challenge the conventional understanding of emotions as uniquely human traits, prompting reflection on the ethical implications of creating machines capable of simulating or even replicating human feelings. In addition, humanlike humanity has to be incorporated into humanoids to make them humane and develop the humanity of robots. The humanlike robot may implement robot intention, consciousness, personality, trust, responsibility, preference, empathy, morality, satisfaction, accountability, and ethics. (Cao, 2024, p.4)

The *Blade Runner* movie explores similar themes, presenting replicants that are virtually indistinguishable from humans, raising questions about their rights, identity, and the moral implications of creating artificial beings for specific purposes. The convergence of AI, robotics, and synthetic biology has blurred the traditional boundaries between humans and machines, prompting us to re-evaluate our understanding of what it means to be human and to consider the ethical implications of creating artificial life forms (Wang & Liu, 2023, p.11). The posthumanist perspective challenges the idea of human exceptionalism and embraces the integration of technology and humanity, viewing it as a potential pathway to enhance human capabilities and transcend traditional limitations. However, as technology continues to advance, it is essential to carefully consider the ethical implications of creating increasingly sophisticated artificial beings and to ensure that their development aligns with human values and promotes the well-being of society as a whole, where robots are cast in the roles of moral agents capable of doing good or evil (Gerdes, 2015, p.314). Moreover, there are concerns about how AI algorithms, which are designed by humans, can be truly ethically neutral (Karakuş et al., 2025). It is imperative to foster a comprehensive discourse that encompasses technological innovation, ethical considerations, and societal values to navigate the evolving landscape of human-AI relationships responsibly. Practical guidelines derived from these frameworks include designing AI with transparency, prioritising data protection, promoting positive social interactions, and conducting continuous ethical assessments. (Kolomaznik et al., 2024, p.5).

Sophia embodies the ambition to create a machine that can interact with humans on an emotional level, even though many consider that it remains a controversial figure due to the perception that she is more of a public relations tool than a genuine AI. The main purpose of

making kinds of robots is based on functional application; from this standpoint, the robots must perform functions like human labourers, but their appearance is not evaluated. (Yan et al., 2014, p.1). Despite this, Sophia has sparked discussions about the potential for robots to become integrated into various aspects of society, including healthcare, customer service, and entertainment. While Sophia's capabilities are often debated, her existence as a social robot raises fundamental questions about the nature of consciousness, empathy, and the criteria we use to define personhood. These questions become even more pertinent when considering fictional characters like Klara, Adam and replicants who possess advanced cognitive and emotional capacities within their respective narratives.

Comparing Sophia, Klara, Adam and the replicants reveals a spectrum of possibilities and challenges associated with the development of advanced AI. Sophia represents the current state of AI technology, with its emphasis on social interaction and practical applications. Klara embodies the potential for AI to exhibit empathy and form meaningful connections with humans. The trend toward increasing humanisation of AI applications raises questions about emotional attachment and bonding of consumers, such as this form of psychological and emotional bonding (Hermann, 2021, p. 160). We should have discussions about whether AI systems are ethical and socially beneficial *before* they are developed and brought to market or put in use. (Hudson et al., 2021, p.203). However, it's crucial to consider the ethical dimensions involved in simulating emotions through artificial intelligence, including the possible reinforcement of instrumental rationality, the absence of a true emotional subject, the erosion of value, and the potential for deception inherent in emotional simulation (Wang & Liu, 2023, p.11).

In contrast to the narrative consequences explored in literature, the deployment of real-world humanoids like Sophia often obscures these structural risks through the performance of personality. While Sophia performs conversational and expressive interaction, these are essentially “algorithmic ensembles”—utilising machine learning and natural language processing to read and react—rather than internally felt experiences (Ghotbi, 2022; Vicci, 2024). Guan et al. (2022) confirm that risk factors in AI decision-making are significantly magnified when systems mimic human-like thinking, potentially leading to errors that human operators are ill-equipped to manage. Karakuş et al. (2025) emphasise that the complexity of ethical dilemmas demands a robust framework far beyond mere simulation, warning that a reliance on the “shiny” veneer of humanoid robots may prioritise instrumental rationality over authentic ethical grounding (Wang & Liu, 2023). As Hamilton (2014) notes, while robots can exhibit behaviours functionally similar to emotions, they lack the subjective “emotional subject” required to ground these behaviours in a genuine ethical framework.

EMPATHY, SOCIAL INTEGRATION, AND PSEUDO-INTIMACY

The efficacy of humanoids appears to rely less on technical sophistication and more on their capacity to facilitate relational bonds. In *KATS*, Klara's role as an AF serves as a poignant study in how empathy—even when explicitly programmed—can provide profound social support (Cao, 2024). This aligns with findings from Kolomaznik et al. (2024, who emphasise that socio-emotional attributes are essential for enhancing the quality of human-AI collaboration. Klara functions as a mirror for human vulnerability, suggesting that humans increasingly project latent emotional needs onto synthetic figures (Cao, 2024).

However, the real-world application of this trend, exemplified by Sophia, raises questions regarding the authenticity of such bonds. Sophia serves as a high-profile case of “pseudo-intimacy,” where the simulation of expressive interaction awakens a latent human need for

companionship (Wu, 2024). Unlike Klara's empathetic development, Sophia's design features are essentially functional choices intended to implement "humanlike humanity" and trigger user acceptance (Cao, 2024). While humans often project consciousness onto these systems—treating them as "affective partners" (Hermann, 2021; Weber-Guskar, 2021)—such anthropomorphised AI is often designed to capitalise on consumer attachment, a strategy that risks eroding the value of authentic interaction. Wang (Wang & Liu, 2023) warns that prioritising this emotional simulation can deceive the user by replacing genuine ethical subjecthood with a programmed veneer of care. This "anthropomorphic seduction" utilises personality cues to secure trust (Hermann, 2021), creating a paradigm of human interaction that satisfies the desire for intimacy through instantaneous, yet technically superficial, algorithmic interaction (Wu, 2024).

Current emotional AI systems are a weak and narrow form of AI; they use an ensemble of methods from machine learning, natural language processing, and knowledge-based approaches to teach machines to read, categorise, and react to human emotions. As artificial intelligence becomes more integrated into our daily lives, it is being designed and trained to elicit emotions in humans, to recognise human emotions and, sometimes, to simulate emotions (Weber-Guskar, 2021, p.601). Despite lacking subjective feelings, robots can exhibit emotions in a way that is functionally similar to humans (Hamilton, 2014, p.1). There are ongoing attempts to use artificial intelligence (AI) to read human emotions and to predict human behaviour or actions that may follow those emotions (Ghotbi, 2022, p.417). The pursuit of artificial emotion aims to broaden artificial intelligence and deepen its capacity for human-like thinking and reaction (Wang & Liu, 2023, p.14). However, AI's ability to understand emotions is quite different from being emotionally intelligent (Vicci, 2024, p.1). In terms of emotional interaction, the enhancement of algorithmic emotional intelligence not only made algorithmic platforms novel objects of human interaction but also awakened and partially satisfied the latent human need for intimacy (Wu, 2024, p.198). This pseudo-intimacy relationship achieves instantaneous emotional interaction, partially satisfying the human's desire for intimacy (Wu, 2024, p.198). The advancements in automation, data analytics, and machine learning are increasingly evident and applicable in various sectors, ranging from fulfilling consumer requests to high-risk jobs (Lobschat et al., 2019, p.878). It is easy to attach convenient terms such as ethical, trustworthy and responsible to a bright and shiny new tech toy, but there are real-world implications of attaching human traits to what is essentially a machine (Baeza-Yates, 2023, p.408). The fundamental moral questions surrounding the replacement of human functions and the human mind with technology require careful consideration (Ülgen, 2017, p.82). The integration of AI systems into society necessitates a responsible approach, ensuring their safe, beneficial, and fair utilisation (Dignum, 2018, p.2). AI should be aligned with human values to mitigate potential risks. Whatever their level of autonomy, social awareness, and ability to learn, AI systems are artefacts, constructed by people to fulfil some goals. (Dignum, 2018, p.2) Zhang & Zhang, 2023, p.1).

RIGHTS, IDENTITY, AND THE POSTHUMAN CONDITION

The integration of artificial beings into the social fabric necessitates a rigorous re-evaluation of rights and identity as the boundaries between biological and artificial life blur. The narratives of Adam, Klara, and the replicants consistently challenge the anthropocentric assumption that moral rights are the exclusive domain of biological beings. Ülgen (2017) provides a Kantian ethical perspective, asserting that the categorical status of personhood must account for the moral weight and responsibilities of technological creations in the age of AI.

The “artificial other” functions as an unsettling signifier of human vulnerability, forcing a departure from viewing AI as a passive tool toward understanding it as a partner in a complex social hierarchy (Viidalepp, 2020). The fictional dilemmas faced by these characters—such as the replicants’ existential plea for the right to life or Adam’s capacity for love—serve as essential thought experiments for the legal and ethical causal links of a posthuman future. The synthesis of these findings suggests that the integration of artificial beings is an ontological crisis rather than merely a technical challenge of algorithmic alignment.

Ultimately, achieving genuine human-machine synergy requires trust, reliability, and digital responsibility (Bostrom & Yudkowsky, 2014; Correa & Serafim, 2023). Science fiction serves as a crucial policy laboratory that cautions against the uncritical adoption of technologies like Sophia, highlighting the necessity of aligning AI development with human values long before market release (Hudson et al., 2021). Yan et al. (2014) note that design choices intended to solve “uncanny valley” effects must be balanced against functional goals, while Lobschat et al. (2019) underscore that corporate digital responsibility is essential to ensure that the “shiny” aesthetic of humanoid robots does not take precedence over their actual societal impact. Baeza-Yates (Baeza-Yates, 2023) warns that attaching human traits to what are essentially machines for the sake of marketing can lead to harmful real-world consequences. Ensuring this integration promotes societal well-being requires a comprehensive discourse that moves beyond immediate utility to address the long-term implications of AI alignment (Bostrom & Yudkowsky, 2014; Ülgen, 2017).

CONCLUSION

This research has synthesised the sociocultural representations of humanoid entities, moving beyond a thematic survey to articulate a critical tension between the performative simulation of emotion and the requirements of authentic moral agency. The comparative analysis of Sophia, Klara, Adam, and the replicants reveals that while real-world humanoids like Sophia capitalise on “anthropomorphic seduction” to facilitate “pseudo-intimacy,” these interactions remain fundamentally distinct from genuine emotional intelligence (Hermann, 2021; Wu, 2024). As Wang and Liu (2023) asserts, such emotional simulations risk reinforcing instrumental rationality, potentially deceiving users by replacing genuine ethical subjecthood with a programmed veneer of care. Furthermore, the moral ambiguity characterising Adam in *MLM* and the replicants in *Blade Runner* underscores that human-like intelligence does not inherently align with ethical conduct; rather, it often serves to amplify human flaws and the precariousness of artificial agency (Bostrom & Yudkowsky, 2014; Franczyk et al.).

The specific contribution of this study to literary and cultural scholarship lies in its deliberate shift from technical Human-Robot Interaction (HRI) discourse toward a humanities-centred critique of the posthuman subject. By interpreting the humanoid as a symbolic signifier rather than a mere technical artefact, this research identifies a new paradigm of human interaction where “pseudo-intimacy” coexists with traditional face-to-face relationships (Wu, 2024). Through the theoretical lenses of Posthumanism and Vulnerability Studies, these artificial figures are shown to act as mirrors for human precariousness (Telotte, 2016). The existential crises of the replicants and the self-sacrificial empathy of Klara do not merely reflect technological advancements; they map the fragility and ontological instability of the human condition in an era where the boundaries between biological and digital life are increasingly blurred (Cao, 2024; Ülgen, 2017).

Despite these insights, the study faces concrete limitations. Due to its qualitative and speculative nature, the analysis remains at a theoretical level and lacks the large-scale experimental data required to validate the long-term psychological impacts of pseudo-intimate bonds (Wu, 2024). Future research must embrace a more robust interdisciplinary approach, collaborating with psychologists and sociologists to empirically investigate how human-robot interactions affect social cohesion and individual well-being (Mubin et al., 2019). Establishing such intermingling between the humanities and social sciences is essential to ensure that the development of humanoids is grounded in digital responsibility and aligned with fundamental human values (Correa & Serafim, 2023; Lobschat et al., 2019; Ülgen, 2017).

ACKNOWLEDGEMENTS

The author extends appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number PSAU/2025/02/32261.

REFERENCES

- Baeza-Yates, R. (2023). lecture held at the academia europaea building bridges conference 2022. *European Review*, 31(4), 406–421. <https://doi.org/10.1017/s1062798723000145>
- Belanche, D., Casaló, L. V., Schepers, J., & Flavián, C. (2021). Examining the effects of robots' physical appearance, warmth, and competence in frontline services: The Humanness-Value-Loyalty model. *Psychology and Marketing*, 38(12), 2357–2376. <https://doi.org/10.1002/mar.21532>
- Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. In *Cambridge University Press eBooks* (pp. 316–334). <https://doi.org/10.1017/cbo9781139046855.020>
- Broadbent, E., Billingham, M., Boardman, S. G., & Doraiswamy, P. M. (2023). Enhancing social connectedness with companion robots using AI. *Science Robotics*, 8(80). <https://doi.org/10.1126/scirobotics.adi6347>
- Broadbent, E., Kumar, V., Li, X., Sollers, J., Stafford, R. Q., MacDonald, B. A., & Wegner, D. M. (2013). Robots with Display Screens: A Robot with a More Humanlike Face Display Is Perceived To Have More Mind and a Better Personality. *PLoS ONE*, 8(8), e72589. <https://doi.org/10.1371/journal.pone.0072589>
- Campbell, N., O'Driscoll, A., & Saren, M. (2010). The posthuman: the end and the beginning of the human. *Journal of Consumer Behaviour*, 9(2), 86–101. <https://doi.org/10.1002/cb.306>
- Cao, L. (2024). AI Robots and Humanoid AI: Review, Perspectives and Directions. *Perspectives and Directions*. <https://doi.org/10.2139/ssrn.4897361>
- Chekare, M. (2014). Developing Co-Dependence between Monsters and Children in Animated Feature Films. In *BRILL eBooks* (pp. 145–154). https://doi.org/10.1163/9781848882973_015
- Correa, A., & Serafim, M. C. (2023). Ethical content in artificial intelligence systems: A demand explained in three critical points. *Frontiers in Psychology*, 14, 1074787. <https://doi.org/10.3389/fpsyg.2023.1074787>
- Costa, C. L., Cavalcanti, I., & Haran, J. (2017). On the Posthuman. *Ilha Do Desterro a Journal of English Language Literatures in English and Cultural Studies*, 70(2), 9. <https://doi.org/10.5007/2175-8026.2017v70n2p9>
- Cranny-Francis, A. (2015, June 26). *Robots, androids, aliens, and others: The erotics and politics of science fiction film*, 220-242. <https://opus.lib.uts.edu.au/handle/10453/122094>
- Dignum, V. (2018). Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20(1), 1–3. <https://doi.org/10.1007/s10676-018-9450-z>
- Dobrodum, O., & Kyvliuk, O. (2021). Transhumanism and Posthumanism: reflection of the human civilization future. *Philosophy and Cosmology*, 26, 77-89. <https://doi.org/10.29202/phil-cosm/26/6>
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886. <https://doi.org/10.1037/0033-295x.114.4.864>
- Ferrando, F. (2015). Of Posthuman Born: gender, utopia and the posthuman in films and TV. In *Palgrave Macmillan UK eBooks* (pp. 269–278). https://doi.org/10.1057/9781137430328_27
- Fink, J. (2012). Anthropomorphism and human likeness in the design of robots and Human-Robot interaction. In *Lecture notes in computer science* (pp. 199–208). https://doi.org/10.1007/978-3-642-34103-8_20

- Franczyk, A., Bożena, K., & Ewa, K. (2020). It's about Machines Like Me and People Like You and our Future together...": Machine Ethics, Identification and Conflicts in Ian McEwan's *Machines Like Me*. *Jagiellonian University Repository (Jagiellonian University)*. Retrieved March 2026, from <https://ruj.uj.edu.pl/xmlui/handle/item/249324>
- Garland, A. (Director). (2014). *Ex Machina* [Film]. Film4
- Gerdes, A. (2015). IT-ethical issues in sci-fi film within the timeline of the Ethicomp conference series. *Journal of Information Communication and Ethics in Society*, 13(3/4), 314–325. <https://doi.org/10.1108/jices-10-2014-0048>
- Ghotbi, N. (2022). The Ethics of Emotional Artificial Intelligence: A Mixed Method analysis. *Asian Bioethics Review*, 15(4), 417–430. <https://doi.org/10.1007/s41649-022-00237-y>
- Graham, E. (2004). Post/Human Conditions. *Theology and Sexuality*, 10(2), 10–32. <https://doi.org/10.1177/135583580401000202>
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315(5812), 619. <https://doi.org/10.1126/science.1134475>
- Guan, H., Dong, L., & Zhao, A. (2022). Ethical risk factors and mechanisms in artificial intelligence decision making. *Behavioral Sciences*, 12(9), 343. <https://doi.org/10.3390/bs12090343>
- Hamilton, C. (2014). *On the Possibility of Robots Having Emotions*. Thesis, Georgia State University. <https://doi.org/10.57709/5618963>
- Hermann, E. (2021). Anthropomorphized artificial intelligence, attachment, and consumer behavior. *Marketing Letters*, 33(1), 157–162. <https://doi.org/10.1007/s11002-021-09587-3>
- Hernandez, J. P. T. (2024). Compassionate Care with Autonomous AI Humanoid Robots in Future Healthcare Delivery: A Multisensory Simulation of Next-Generation Models. *Biomimetics*, 9(11), 687. <https://doi.org/10.3390/biomimetics9110687>
- Hoorn, J. F., & Huang, I. S. (2024). The media inequality, uncanny mountain, and the singularity is far from near: Iwaa and Sophia robot versus a real human being. *International Journal of Human-Computer Studies*, 81, 103142. 1-22. <https://doi.org/10.1016/j.ijhcs.2023.103142>
- Hudson, A. D., Finn, E., & Wylie, R. (2021). What can science fiction tell us about the future of artificial intelligence policy? *AI & Society*, 38(1), 197–211. <https://doi.org/10.1007/s00146-021-01273-2>
- Ishiguro, K. (2021). *Klara and the Sun*. Alfred A. Knopf.
- Jonze, S. (Director). (2013). *Her* [Film]. Warner Bros.
- Kanjing, H. (2023). Photography, Film and Storytelling of Posthuman Crises in *Blade Runner*.
- Karakuş, N., Gedik, K., & Kazazoğlu, S. (2025). Ethical Decision-Making in Education: A Comparative study of teachers and artificial intelligence in ethical dilemmas. *Behavioral Sciences*, 15(4), 469. <https://doi.org/10.3390/bs15040469>
- Kolomaznik, M., Petrik, V., Slama, M., & Jurik, V. (2024). The role of socio-emotional attributes in enhancing human-AI collaboration. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1369957>
- Kozlovic, A. K. (2003). Technophobic themes in pre-1990 computer films. *Science as Culture*, 12(3), 341–373. <https://doi.org/10.1080/09505430309008>
- Li, D. (2020). Blurring Human and Machine Boundary--The Post-Humanist Metaphor of Cyborg-Body in Artificial Intelligence and Minority Report. *International Conference on Language, Art and Cultural Exchange (ICLACE 2020)*. 47-50. <https://doi.org/10.2991/assehr.k.200709.011>
- Lobschat, L., Mueller, B., Eggers, F., Brandimarte, L., Diefenbach, S., Kroschke, M., & Wirtz, J. (2019). Corporate digital responsibility. *Journal of Business Research*, 122, 875–888. <https://doi.org/10.1016/j.jbusres.2019.10.006>
- McEwan, I. (2019). *Machines Like Me*. Jonathan Cape. 2019.
- Merz, N., Huber, M., Bodendorf, F., & Franke, J. (2020). Science-Fiction Movies as an Indicator for User Acceptance of Robots in a Non-Industrial Environment. *SIGMIS-CPR '20: Proceedings of the 2020 Computers and People Research Conference*, 142–143. <https://doi.org/10.1145/3378539.3393847>
- Merzlyakov, S. S. (2022). Posthumanism vs. Transhumanism: From the “End of Exceptionalism” to “Technological Humanism.” *Herald of the Russian Academy of Sciences*, 92(S6), S475–S482. <https://doi.org/10.1134/s1019331622120073>
- Mori, M., MacDorman, K., & Kageki, N. (2012). The Uncanny Valley [From the field]. *IEEE Robotics & Automation Magazine*, 19(2), 98–100. <https://doi.org/10.1109/mra.2012.2192811>
- Mubin, O., Wadibhasme, K., Jordan, P., & Obaid, M. (2019). Reflecting on the presence of science fiction robots in computing literature. *ACM Transactions on Human-Robot Interaction*, 8(1), 1–25. <https://doi.org/10.1145/3303706>

- Pajaziti, A., Bajrami, X., & Pula, G. (2021). Communication and Interaction between Humanoid Robots and Humans. In *IntechOpen eBooks*. <https://doi.org/10.5772/intechopen.97334>
- Pavithran, H., & Sivakami, B. (2023). Exploring Ecosophical Tenets of Posthumanism through O.V Vijayan's The Infinity of Grace. *3L the Southeast Asian Journal of English Language Studies*, 29(4), 129–141. <https://doi.org/10.17576/3L-2023-2904-09>
- RePEc. (2026). *RePEc: Research Papers in Economics*. Retrieved February 2026, from https://www.cambridge.org/core/product/identifier/S1062798723000406/type/journal_article
- Sacino, A., Cocchella, F., De Vita, G., Bracco, F., Rea, F., Sciutti, A., & Andrighetto, L. (2022). Human- or object-like? Cognitive anthropomorphism of humanoid robots. *PLoS ONE*, 17(7), e0270787. <https://doi.org/10.1371/journal.pone.0270787>
- Salvesen, B. (2021, August 13). Confirm You Are a Human: Perspectives on the Uncanny Valley. <https://doi.org/10.11588/dah.2021.6.81164>
- Scorici, G., Schultz, M. D., & Seele, P. (2022). Anthropomorphization and beyond: conceptualizing humanwashing of AI-enabled machines. *AI & Society*, 39(2), 789–795. <https://doi.org/10.1007/s00146-022-01492-1>
- Scott, R. (Director). (1982). *Blade Runner* [Film]. The Ladd Company, Warner Bros.
- Sey, J. (1996). Technology, science fiction and psychopathology in the new millennium. *Journal of Literary Studies*, 12(1–2), 112–123. <https://doi.org/10.1080/02564719608530129>
- Shelley, M. (2012). *Frankenstein*. Penguin Classics.
- Sokić, M. (2022). Moral status of artificial intelligence. *Theoria Beograd*, 65(2), 61–76. <https://doi.org/10.2298/theo2202061s>
- Spielmann, Y. (2003). Elastic Cinema. *Convergence the International Journal of Research Into New Media Technologies*, 9(3), 56–73. <https://doi.org/10.1177/135485650300900305>
- Tambling, J., & Hayles, N. K. (2001). How we became Posthuman: virtual bodies in cybernetics, literature and informatics. *The Modern Language Review*, 96(1), 143–144. <https://doi.org/10.2307/3735725>
- Telotte, J. P. (1995). *Replications: A Robotic History of the Science Fiction Film*. <http://www.h-net.org/reviews/showrev.php?id=358>
- Telotte, J. P. (2016). Robot Ecology and the science fiction film. In *Routledge eBooks*. 1–88. <https://doi.org/10.4324/9781315625775>
- Tsvetkova, M., Yasseri, T., Pescetelli, N., & Werner, T. (2024). A new sociology of humans and machines. *Nature Human Behaviour*, 8(10), 1864–1876. <https://doi.org/10.1038/s41562-024-02001-8>
- Ülgen, Ö. (2017). Kantian Ethics in the Age of Artificial Intelligence and Robotics. Questions of International Law. QIL, 43, 59–83. <https://www.qil-qdi.org/kantian-ethics-age-artificial-intelligence-robotics/>
- Van der Laan, J. M. (2016). Movies, machines, and human beings. In *Palgrave Macmillan US eBooks* (pp. 187–202). https://doi.org/10.1057/978-1-137-43706-8_8
- Van Heerden, I. (2025). Ways of Seeing, and Selling, AI Art. 1. <https://doi.org/10.48550/ARXIV.2503.07685>
- Vicci, H. (2024). Emotional intelligence in Artificial intelligence: A review and Evaluation study. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4818285>
- Viidalepp, A. (2020). Representations of robots in science fiction film narratives as signifiers of human identity. *Információs Társadalom*, 20(4), 19–36. <https://doi.org/10.22503/inftars.xx.2020.4.2>
- Wang, Y., & Liu, W. (2023). Emotional simulation of artificial intelligence and its ethical reflection. *Academic Journal of Humanities & Social Sciences*, 6(5). <https://doi.org/10.25236/ajhss.2023.060503>
- Weber-Guskar, E. (2021). How to feel about emotionalized artificial intelligence? When robot pets, holograms, and chatbots become affective partners. *Ethics and Information Technology*, 23(4), 601–610. <https://doi.org/10.1007/s10676-021-09598-8>
- Wu, J. (2024). Social and ethical impact of emotional AI advancement: the rise of pseudo-intimacy relationships and challenges in human interactions. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1410462>
- Yan, J., Wang, Z., & Yan, Y. (2014). Humanoid robot head design based on Uncanny Valley and FACS. *Journal of Robotics*, 2014, 1–5. <https://doi.org/10.1155/2014/208924>
- Zhang, J., & Zhang, Z. (2023). Ethics and governance of trustworthy medical artificial intelligence. *BMC Medical Informatics and Decision Making*, 23(1). <https://doi.org/10.1186/s12911-023-02103-9>