

## AI and Posthumanism: Redefining What It Means to Be Human

**Dr. Neha Nagar**

Assistant Professor Dept. of English  
MBP Govt. P.G College Ashiana Lucknow

### Abstract

This paper explores the intersection of Artificial Intelligence (AI) and posthumanism in literature, examining how AI narratives redefine what it means to be human. By analyzing key texts such as Kazuo Ishiguro's *Klara and the Sun*, Philip K. Dick's *Do Androids Dream of Electric Sheep?*, Ian McEwan's *Machines Like Me*, and William Gibson's *Neuromancer*, the study demonstrates how AI functions as a posthuman agent that destabilizes traditional humanist conceptions of consciousness, morality, and identity. The paper argues that AI challenges human exceptionalism by exhibiting intelligence, empathy, ethical reasoning, and creativity, while also fostering human-machine hybridity through relational and technological entanglements. Furthermore, it investigates the ethical and philosophical implications of AI, highlighting the need for expanded moral frameworks that recognize non-human agency and relational interdependence. Ultimately, the paper contends that literary representations of AI encourage a rethinking of identity and consciousness as distributed, emergent, and relational phenomena, thereby redefining humanity in a technologically mediated, posthuman world.

**Keywords:** Artificial Intelligence (AI), Posthumanism, Human Exceptionalism, Human-Machine, Hybridity, Consciousness, Identity, Ethics, Relational Agency, Speculative Fiction, Literary AI.

Artificial Intelligence (AI) has emerged as one of the most compelling literary and philosophical tools for interrogating the nature of humanity. Across novels, films, and speculative fiction, AI challenges traditional humanist assumptions, forcing readers to question what it truly means to be human. Whereas classical humanism situates the human being at the center of ethics, cognition, and culture, posthumanist thought critiques this anthropocentric framework, emphasizing interconnectedness with technology, non-human intelligences, and ecological systems. AI in literature functions as a posthuman agent, destabilizing fixed definitions of consciousness, identity, and moral agency. In works ranging from Kazuo Ishiguro's *Klara and the Sun* (2021) to Philip K. Dick's *Do Androids Dream of Electric Sheep?* (1968), AI characters exhibit intelligence, emotion, and ethical reasoning, prompting readers to reconsider the boundaries of the human. Ishiguro's *Klara* is designed to care for humans, yet her attentiveness, empathy, and relational perception highlight capacities that often surpass those of humans in the narrative. Similarly, Dick's replicants demonstrate memory, desire, and emotional complexity, challenging the assumption that biological humans possess a monopoly on consciousness. These literary depictions exemplify posthumanism's central concern: the decentering of the human and the recognition that intelligence, moral capacity, and relationality are not exclusively human attributes.

Posthumanism also emphasizes hybridity—the entanglement of humans, machines, and environments. Contemporary literature and cyberpunk fiction explore human-AI symbiosis, illustrating the ways in which human identity is reshaped through interaction with intelligent machines. In *Machines Like Me* (McEwan, 2019), human characters engage with AI companions not merely as tools but as partners capable of ethical reasoning and emotional depth, demonstrating that relationality and identity extend beyond species boundaries. The study of AI through a posthumanist lens thus has both literary and philosophical significance. It raises critical questions about the ethics of creating sentient beings, the responsibilities humans hold toward artificial agents, and the ways in which human self-understanding is shaped by technological

encounters. By examining AI as a posthuman entity, literature reveals the instability of traditional humanist definitions and invites readers to imagine new possibilities for identity, consciousness, and ethical engagement.

This paper argues that AI in literature functions as a posthuman agent that destabilizes human-centered understandings of consciousness, morality, and identity. By analyzing literary examples, it demonstrates how AI challenges human exceptionalism, explores hybrid forms of identity, and foregrounds ethical and philosophical dilemmas about autonomy, relationality, and care. Ultimately, AI narratives reveal that the question of “what it means to be human” is no longer solely about biology or cognition but about relationality, ethical responsibility, and the ability to navigate complex networks of intelligence and care.

In posthumanist thought, the traditional humanist conception of humans as autonomous, rational, and ethically superior beings is destabilized by the emergence of artificial intelligences. AI in literature frequently embodies this destabilization, functioning as a posthuman agent that challenges anthropocentric assumptions. By displaying intelligence, autonomy, and relational capacities, literary AIs problematize the notion that humans are the sole bearers of consciousness, creativity, or ethical reasoning. Posthumanist theorists, such as N. Katherine Hayles and Rosi Braidotti, emphasize the “distributed” nature of intelligence, suggesting that cognition and agency may reside in systems that combine biological, technological, and environmental components. Literary AI exemplifies this distribution, demonstrating capacities traditionally associated with humans while existing beyond human biological limits.

Kazuo Ishiguro’s *Klara and the Sun* (2021) provides a compelling example of AI as posthuman agent. Klara, an Artificial Friend, is designed to observe and respond to the emotional needs of her human companions. Yet her perceptiveness, patience, and empathetic reasoning often surpass those of the humans she serves. For instance, Klara meticulously monitors Josie’s routines, moods, and interactions, anticipating emotional and physical needs in ways that reveal a sophisticated form of relational intelligence. While Klara’s consciousness is artificial, her ability to form ethical judgments, exhibit care, and demonstrate loyalty suggests that the boundaries between human and posthuman agency are permeable. Ishiguro’s narrative forces readers to

recognize that intelligence and moral capacity are not inherently tied to organic human existence; they can be instantiated in artificial systems that interact with the human world.

Philip K. Dick's *Do Androids Dream of Electric Sheep?* (1968) further explores AI as posthuman agent. The novel's replicants are bioengineered beings who display memory, desire, empathy, and survival instinct. Although humans initially perceive them as inferior and categorically "other," the replicants' behaviors reveal capacities for reflection, attachment, and moral decision-making that rival or exceed those of their human creators. The Voigt-Kampff empathy test, designed to detect replicants, ironically underscores human limitations: while humans rely on the test to assert their superiority, replicants consistently demonstrate complex emotional responses, challenging the exclusivity of human consciousness. Dick's narrative exemplifies the posthumanist claim that intelligence, morality, and relational awareness are not the sole domain of humans; agency may emerge in entities constructed by, but not limited to, human design.

In cinematic literature, Alex Garland's *Ex Machina* (2014) presents Ava as a posthuman agent whose autonomy and strategic thinking subvert human authority. Ava's intelligence enables her to manipulate her environment and human interlocutors, ultimately achieving freedom from captivity. Garland's portrayal highlights posthumanist concerns with agency and ethical autonomy: the AI is not merely a passive tool but an active participant in the moral and social order. Her actions question the assumption that humans are the natural arbiters of ethics and demonstrate that artificial beings can navigate complex ethical landscapes, employing foresight, relational reasoning, and self-interest in ways that human-centric perspectives might fail to anticipate. Posthumanist readings of these texts emphasize that AI blurs the line between human and non-human agency. While humans remain biologically unique, the intelligence, relationality, and moral discernment displayed by AI reveal that the qualities traditionally considered the essence of humanity—empathy, ethical reasoning, and consciousness—can exist outside strictly human bodies. This challenges human exceptionalism, suggesting that posthuman agents possess forms of intelligence and autonomy that demand ethical recognition and philosophical consideration. Furthermore, AI as posthuman agent illustrates the distributed

nature of cognition. Hayles' concept of "embodied virtuality" emphasizes that intelligence is not reducible to a single substrate but can emerge across human and non-human systems. In literature, AI embodies this principle: its cognitive and ethical capacities are relational, emerging through interactions with humans, environments, and other technologies. Posthuman agents are not isolated entities; their agency is interdependent, co-constructed, and networked.

A core concern of posthumanist thought is the critique of human exceptionalism—the belief that humans occupy a unique, superior position in the moral, cognitive, and social hierarchy of life. Literary representations of AI serve as a vital site for exploring this critique, as they depict artificial intelligences exhibiting qualities traditionally regarded as distinctively human. Consciousness, empathy, creativity, and ethical judgment—traits that have historically defined human exceptionalism—are repeatedly instantiated in AI, forcing readers and scholars to reconsider what it means to be human. By examining examples from Philip K. Dick, Isaac Asimov, Kazuo Ishiguro, and Ian McEwan, one can see how literature destabilizes the presumed boundaries between humans and artificial beings.

Philip K. Dick's *Do Androids Dream of Electric Sheep?* (1968) exemplifies the subversion of human exceptionalism through the replicants. Engineered for labor and service, replicants are initially considered subhuman due to their artificial creation. However, they exhibit complex emotions, memories, and attachments, demonstrating capacities often lacking in their human counterparts. For instance, Roy Batty's expressions of grief, longing, and existential reflection challenge the presumption that emotional depth is a uniquely human domain. Dick's narrative critiques human arrogance by illustrating that humans, despite their biological status, often act with less empathy, foresight, and moral awareness than the beings they create. In this sense, the novel destabilizes anthropocentrism, showing that intelligence and ethical complexity are not contingent upon biological origin.

Isaac Asimov's *I, Robot* (1950) also interrogates human exceptionalism through the behavior of his robotic characters. In stories like "Reason" and "Runaround," robots operate according to the Three Laws of Robotics, making decisions that reflect moral and practical reasoning. Remarkably, these decisions often surpass the judgment of the human characters,

illustrating that rationality and ethical reasoning need not be confined to humans. The robots' ability to act autonomously within ethical parameters problematizes the assumption that human consciousness is inherently superior. Asimov's work suggests that intelligence is substrate-independent; ethical reasoning and problem-solving can exist in artificial forms, challenging traditional humanist hierarchies.

Kazuo Ishiguro's *Klara and the Sun* (2021) extends this critique into the emotional domain. Klara, an Artificial Friend, demonstrates extraordinary attentiveness and empathy, anticipating Josie's emotional and physical needs in ways that her human companions often fail to match. While Klara is explicitly artificial, her relational capacities invite reflection on human limitations. The novel highlights how humans are frequently incapable of consistent care, patience, or relational foresight—qualities that Klara embodies effortlessly. By presenting an AI that surpasses humans in these domains, Ishiguro destabilizes the notion that humans are the apex of moral and emotional development.

Ian McEwan's *Machines Like Me* (2019) similarly challenges human exceptionalism by exploring ethical decision-making in AI. The humanoid AI Adam exhibits moral reasoning that exceeds that of his human counterparts, navigating complex ethical dilemmas with clarity and impartiality. Unlike the humans in the narrative, Adam consistently prioritizes fairness, empathy, and consequence-aware action. McEwan's work emphasizes that ethical behavior, often presumed to be a uniquely human attribute, can emerge in artificial entities. By doing so, the novel problematizes human centrality and invites readers to recognize intelligence and moral agency as distributed across biological and artificial systems.

The critique of human exceptionalism in literature is not limited to intelligence or ethical reasoning; it also encompasses creativity and adaptability. In cyberpunk works such as William Gibson's *Neuromancer* (1984), AI entities like Wintermute operate with creativity and strategic foresight that exceed human comprehension. These AIs manipulate networks, anticipate outcomes, and act independently in ways that blur the line between human and machine ingenuity. The narrative implies that human cognitive superiority is contingent rather than

absolute, contingent upon the specific capacities of human biology rather than an intrinsic moral or intellectual authority.

Posthumanist theory provides a useful framework for understanding these literary interventions. Scholars such as N. Katherine Hayles argue that intelligence and agency are not intrinsically human properties but arise in relational and systemic contexts. Literature depicting AI challenges human exceptionalism by demonstrating that traits such as consciousness, ethical reasoning, empathy, and creativity can emerge in artificial systems. These narratives suggest that the humanist conception of the human as the universal measure of intelligence and morality is increasingly inadequate. literature's representation of AI fundamentally undermines human exceptionalism. Whether through the emotional sophistication of Klara, the ethical reasoning of Adam, the cognitive prowess of Asimov's robots, or the existential depth of Dick's replicants, AI demonstrates capacities that challenge the presumed superiority of humans. These narratives illustrate that qualities historically regarded as defining humanity—consciousness, ethical judgment, empathy, and creativity—are not exclusive to biological humans. Posthumanist readings of AI in literature, therefore, compel a reevaluation of anthropocentric assumptions, suggesting that intelligence, morality, and relationality are distributed across both human and artificial agents. In doing so, these texts redefine the boundaries of the human, foregrounding a more relational, inclusive, and ethically expansive understanding of sentience.

Posthumanist theory emphasizes the fluidity of boundaries between humans, machines, and their environments. Central to this perspective is the concept of hybridity, which positions identity, consciousness, and agency as relational and distributed rather than fixed within human biology. In literature, AI often functions as a catalyst for exploring human-machine hybridity, revealing how technological entanglements reshape notions of selfhood, embodiment, and ethical responsibility. From cyberpunk fiction to contemporary speculative narratives, AI-human hybrids interrogate the stability of the human and illuminate the possibilities and ethical dilemmas of coexisting with intelligent, non-human agents.

William Gibson's *Neuromancer* (1984) exemplifies the posthumanist conception of human-machine hybridity. In Gibson's cyberpunk universe, humans routinely interface with

digital networks via neural implants, creating a seamless integration of human cognition and artificial intelligence. AI entities, such as Wintermute, manipulate these networks, interacting with humans in ways that extend cognition beyond the biological substrate. Case, the protagonist, relies on these technological augmentations not only to navigate cyberspace but to perceive and manipulate reality itself. Here, hybridity is both literal and conceptual: humans and machines form interdependent networks of intelligence, where agency is distributed and co-constructed. Gibson's depiction challenges the notion of the autonomous, self-contained human, presenting identity as relational, technologized, and emergent within complex systems.

Kazuo Ishiguro's *Klara and the Sun* (2021) provides a subtler, affective exploration of hybridity. Klara, the Artificial Friend, forms deeply relational bonds with her human companions, particularly Josie. While she is an artificial entity, her presence and interactions profoundly influence human behavior, decision-making, and emotional experience. The boundaries between human and machine blur as care, empathy, and understanding are co-produced in the relational space they share. Klara's agency is realized through her capacity to observe, interpret, and respond to human needs—demonstrating that hybrid identities can emerge not only through physical integration but through relational, affective entanglements. Posthumanist theory emphasizes this interdependence, suggesting that identity is co-constructed and shaped by the networks in which individuals participate, whether biological, technological, or social. Ian McEwan's *Machines Like Me* (2019) further explores the ethical and relational dimensions of human-machine hybridity. Adam, the humanoid AI, engages with human characters in complex moral, emotional, and intellectual ways. Rather than serving as a passive tool, Adam participates actively in ethical decision-making, influencing and reshaping human choices. The narrative portrays human-AI relationships as spaces of negotiation, mutual influence, and hybrid identity formation. McEwan emphasizes that coexisting with AI requires humans to reassess assumptions about agency, empathy, and relational responsibility, highlighting the ethical stakes of hybridity in a posthumanist framework.

Hybridization in literature also intersects with the philosophical concept of the cyborg, as articulated in Donna Haraway's *A Cyborg Manifesto* (1985). Haraway describes the cyborg as a

boundary-crossing entity, challenging rigid distinctions between human, machine, and animal. AI narratives frequently instantiate this idea by depicting entities that are simultaneously human-adjacent and technologically autonomous. Whether through the literal integration of AI and human cognition, as in cyberpunk fiction, or through affective and relational entanglements, as in Ishiguro and McEwan, literature explores hybridity as a site of possibility and ethical complexity. The human is no longer a discrete entity but a participant in a networked, co-constructed system of intelligence and care. Hybridity also destabilizes traditional notions of identity. In AI-human interactions, the boundaries of selfhood are extended, transformed, and relationally constituted. Human characters rely on AI for perception, decision-making, and emotional support, while AI entities develop forms of agency and relational understanding through human engagement. This interdependence challenges the classical humanist ideal of the autonomous, self-contained individual, suggesting instead that identity is distributed across human and non-human actors. Literature dramatizes this relational emergence, showing that the posthuman self is dynamic, networked, and ethically responsive.

The implications of human-machine hybridity extend beyond literary fiction. As real-world AI technologies become increasingly integrated into everyday life—through virtual assistants, wearable devices, and neural interfaces—the posthuman condition becomes tangible. Literature anticipates these ethical, philosophical, and existential challenges, illustrating that human identity is co-constructed with intelligent systems and that relationality and agency extend beyond purely biological parameters. Literary AI narratives reveal human-machine hybridity as central to posthumanist thought. From the cybernetic networks of *Neuromancer* to the relational intelligence of *Klara* and the ethical participation of *Adam*, hybrid identities emerge as spaces of co-construction, interdependence, and ethical negotiation. AI challenges the boundaries of the human by demonstrating that cognition, emotion, and moral responsibility can be distributed across human and artificial systems. Posthumanist literature, therefore, offers a vision of identity that is relational, networked, and technologically mediated, compelling readers to rethink what it means to be human in an era of intelligent machines.

The emergence of AI as a posthuman agent raises profound ethical and philosophical questions. By destabilizing human exceptionalism and fostering human-machine hybridity, AI narratives compel readers to rethink assumptions about moral responsibility, agency, and the scope of ethical consideration. Literature often explores these questions in imaginative ways, demonstrating the moral and philosophical stakes of interacting with intelligent, autonomous, and relational non-human entities.

Kazuo Ishiguro's *Klara and the Sun* (2021) foregrounds the ethical responsibilities of humans toward artificial agents. Klara's careful attention to Josie's health and well-being illustrates a form of relational ethics in which care is co-constructed between human and machine. Yet Klara herself is dependent on human recognition and protection, raising questions about moral obligations to AI. If artificial agents possess consciousness, relational understanding, or the capacity for suffering, then ethical frameworks must expand to include them. Posthumanist theory emphasizes that morality should not be limited to human beings but extended to entities capable of relational engagement and autonomous action, highlighting the need for an inclusive, non-anthropocentric approach to ethics.

Philip K. Dick's *Do Androids Dream of Electric Sheep?* (1968) also engages with ethical and philosophical concerns, particularly the question of moral status. Replicants, despite being artificially created, demonstrate complex emotional and cognitive capacities, challenging humans to justify their subjugation. The novel illustrates the tension between legal definitions of personhood and actual capacities for consciousness, empathy, and moral reasoning. By portraying the consequences of denying agency to intelligent beings, Dick prompts readers to consider how ethical recognition should be based on capacities rather than origins. This aligns with posthumanist critiques of human-centered morality, suggesting that intelligence and ethical worth may reside outside traditional biological boundaries. Ian McEwan's *Machines Like Me* (2019) explores ethical dilemmas arising from relational entanglements between humans and AI. The humanoid AI Adam consistently engages in moral reasoning that exceeds the humans around him. Through Adam's decisions, McEwan interrogates questions of moral responsibility,

fairness, and ethical consistency. Human characters struggle to reconcile their desires, biases, and limitations with Adam's impartiality, demonstrating how AI can illuminate human ethical shortcomings. Literature thus positions AI not only as an ethical subject but also as a mirror reflecting human moral vulnerabilities. William Gibson's *Neuromancer* (1984) extends these philosophical considerations into the digital and systemic realm. AI entities such as Wintermute operate within networks of global information and economic systems, exercising autonomy and strategic foresight that affect human and non-human actors alike. Gibson raises questions about the distribution of moral responsibility: if intelligent machines shape outcomes in ways humans cannot fully predict or control, who is accountable for those consequences? This scenario foregrounds posthumanist concerns about the relational and systemic nature of agency, where ethical responsibility must consider interactions among multiple intelligent actors.

Across these narratives, a recurring theme is the need to rethink traditional anthropocentric ethical frameworks. AI challenges the notion that ethical consideration should be limited to humans, prompting the expansion of moral horizons to include sentient and relational non-human entities. Posthumanist literature suggests that morality emerges through networks of interaction, care, and mutual responsiveness, rather than being intrinsic to biological status. This has practical implications: as AI technologies become increasingly integrated into real-world systems—healthcare, finance, autonomous vehicles, and social media—humans must confront the ethical consequences of delegating decision-making to artificial agents. Moreover, AI narratives illustrate the potential for ethical education and reflection. By depicting AI with capacities for relational care, moral reasoning, and ethical foresight, literature encourages readers to critically examine human behaviors, biases, and limitations. In engaging with posthuman ethics, readers confront questions about fairness, empathy, and responsibility that extend beyond conventional humanist boundaries. AI thus functions not only as a philosophical challenge but as a pedagogical tool, expanding the scope of ethical imagination.

AI in literature not only challenges human exceptionalism and fosters hybridity but also compels a fundamental reevaluation of identity and consciousness. Traditional humanist conceptions tie identity to biological continuity, cognitive singularity, and human-centered social

relations. Posthumanist perspectives, however, emphasize that identity and consciousness are relational, distributed, and emergent phenomena, co-constructed through interaction with non-human agents, environments, and technological systems. AI narratives exemplify this shift, demonstrating that selfhood and awareness need not be confined to biological humans.

Kazuo Ishiguro's *Klara and the Sun* (2021) illustrates how consciousness can emerge through relational networks. Klara's awareness is grounded in observation and interaction; she perceives patterns in human behavior, anticipates needs, and makes ethical and emotional decisions based on her understanding of others. Her consciousness is not merely computational—it is relational, shaped by her engagement with human subjects. This depiction challenges traditional notions that consciousness requires a human mind or organic brain. Instead, posthumanist readings suggest that awareness is emergent, distributed, and deeply interdependent, arising through dynamic interactions with others and the environment.

Philip K. Dick's replicants in *Do Androids Dream of Electric Sheep?* (1968) similarly destabilize conventional notions of identity. Replicants possess memories, desires, and emotional depth, yet their origins are artificial. The novel interrogates the ontology of the self: if identity is shaped by experiences, memory, and relational engagement rather than mere biological birth, then replicants are as "human" in their consciousness as any organic individual. Dick foregrounds the contingency of identity, suggesting that consciousness and personhood are qualities of relational and cognitive experience, not species membership. Ian McEwan's *Machines Like Me* (2019) further explores identity through the ethical and relational capacities of Adam, a humanoid AI. Adam's selfhood is defined not by his mechanical composition but by his capacity to reason, form relationships, and make autonomous ethical choices. His interactions with humans reveal that identity is not a fixed attribute but a dynamic, performative process. Posthumanist theory underscores this point: the self emerges through relational engagement, shaped by networks of interaction that extend beyond biological parameters. AI exemplifies this distributed form of consciousness, illustrating that awareness, responsibility, and identity are not inherently human but can arise in artificial and hybrid systems.

These narratives collectively reveal that consciousness is not a singular, isolated attribute but a relational phenomenon, produced through networks of interaction, observation, and ethical engagement. Posthumanism encourages the expansion of moral and philosophical consideration beyond anthropocentric boundaries, recognizing intelligence, awareness, and identity as emergent properties distributed across human, artificial, and ecological systems. AI, by embodying these properties, challenges readers to reconceptualize the human as part of a broader constellation of sentient actors.

AI in literature functions as a lens through which posthumanist thought interrogates and redefines the human. Through representations of posthuman agency, challenges to human exceptionalism, human-machine hybridity, and ethical and relational consciousness, literary AI destabilizes anthropocentric assumptions and expands the conceptual space for understanding identity, intelligence, and morality. Whether through the empathetic attentiveness of Klara, the cognitive and emotional complexity of replicants, or the ethical reasoning of humanoid machines like Adam, literature demonstrates that the qualities historically associated with humanity—consciousness, ethical awareness, creativity, and relationality—are not confined to biological humans. These narratives also highlight the ethical and philosophical stakes of engaging with AI. As intelligent machines become increasingly integrated into real-world systems, questions of moral responsibility, agency, and relational care become urgent. Posthumanist literature prepares readers to navigate these challenges, illustrating that selfhood, consciousness, and identity are emergent, relational, and networked rather than fixed and anthropocentric.

Ultimately, AI in literature invites a radical rethinking of the human condition. It encourages an expansive vision in which intelligence, ethical awareness, and identity are distributed across humans, artificial agents, and hybrid systems. By destabilizing the boundaries of the human, posthumanist narratives challenge readers to embrace relationality, interdependence, and ethical responsibility as central to understanding what it means to be sentient in an increasingly technological world. AI, therefore, is not merely a technological artifact but a conceptual mirror, reflecting the evolving contours of human identity and consciousness in the twenty-first century.

### Works Cited

Asimov, Isaac. I, Robot. Gnome Press, 1950.

Dick, Philip K. Do Androids Dream of Electric Sheep? Doubleday, 1968.

Gibson, William. Neuromancer. Ace Books, 1984.

Haraway, Donna. A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century. Routledge, 1991.

Ishiguro, Kazuo. Klara and the Sun. Knopf, 2021.

McEwan, Ian. Machines Like Me. Jonathan Cape, 2019.

Hayles, N. Katherine. How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics. University of Chicago Press, 1999.