*International Research Journal of Social Sciences and Humanities, Vol.:05,Issue:02,April -June,2026*
*PP: 165-183*

**OPEN ACCESS**
*IRJSSH*
*ISSN (Online): 2959-1392*
*ISSN (Print): 2959-2887*
*www. irjssh.com*

# HUMAN AGENCY AND RESPONSIBILITY TOWARDS AI: A CYBORG MANIFESTO READING OF TED CHIANG'S THE LIFECYCLE OF SOFTWARE OBJECTS

**Sawaira Begum**
*MPhil Scholar, Department of English, Abdul Wali Khan University Mardan, KP, Pakistan.*
**Email:** *sawairabegum916@gmail.com*

**Bibi Aimen**
**Corresponding Author**
*MPhil Scholar, Department of English, Abdul Wali Khan University Mardan, KP, Pakistan.*
***Email:*** *aimensiraj369@gmail.com*

***Rafiq Nawab***
*Assistant Professor, Department of English, Abdul Wali Khan University Mardan, KP, Pakistan.*
***Email:*** ***rafiqnawab@awkum.edu.pk***

**Abstract**

*This paper, through the lens of Cyborg Manifesto, explores the ethical responsibility and accountability of human agency towards artificially intelligent beings in the novella, Lifecycle of Software Objects by Ted Chiang. This study views the representation of AI beings in the novella as a relational entity which is dependent on human intervention, how they care about them and institutional and economic control created by humans. In the form of digients, or digital beings, Chiang questions the ethical and epistemological boundaries between artificial sentience and human cognition, examining whether personhood can be extended to beings whose consciousness has computationally emerged, and what will be their status in human societies. This work draws on the conceptual framework of Cyborg Manifesto by Donna Haraway, which is stated in her essay: "A Manifesto of Cyborgs: Science, Technology and Socialist Feminism in 1980s", to emphasize that the boundaries between humans and digital beings are socially constructed and ethically consequential. Instead of reading the cyborg merely as a theoretical metaphor for boundary dissolution, the study argues that Chiang transforms Haraway's framework into a practical narrative of ethical responsibility. Viewing the novella through Cyborg Manifesto allows the study to trace how ethical meaning emerges from the AI-human relational system, rather than from technological autonomy. This paper places the novella, Lifecycle of Software Objects as a speculative ethical framework to reposition AI ethics and holds the human agency accountable in the age of AI. This study contributes to the critical discussion on AI ethics, post-humanism and the responsibility of human agency.*

***Keywords:*** *Ted Chiang, AI ethics, Accountability, AI, Cyborg Identity, Cyborg Politics, Donna Haraway.*

## INTRODUCTION

*The current discussion on AI has shifted from technological innovation to the nature of ethics, responsibility and the aspect of personhood. The pervasiveness of Artificial Intelligence in the social, economic and cognitive dimensions evoke the inquiry into essential or non-essential divide between human agency and artificial computation. In such a scenario, where does the human identity stand, in contrast to the artificial consciousness and what would be a socially ethical stand if machines with human or near human consciousness emerged and put into human societies? Here, literature serves as a space that demands a thorough examination into the nature of such AI beings, as well as what type of ethical relations it generates between humans and artificial entities.*

*The Lifecycle of Software Objects is a novella written by Ted Chiang. This piece of literature offers a profound exploration of such concerns as stated above, through "digients" which are artificially intelligent beings created by a software company and trained by human workers, which is much similar to raising organic beings. Unlike conventional representations of AI as autonomous, self-sufficient or domineering and threatening entities, Chiang constricts these beings as dependent, evolving and socially embedded intelligence whose development is shaped by how human agency raises them, their care and their institutions. The novella foregrounds the emotional, ethical, and economic entanglements that emerge between humans and digital lifeforms, thereby raising critical questions about responsibility in human–AI relations.*

*On what grounds shall human agency accept responsibility, recognize AI beings with care, and maintain an ethical relationship with them? To address this concern, this paper reads Chiang's novella through Donna Haraway's conceptual framework of fluid identity and new dimensions of politics in a tech-based world, Cyborg, as articulated in her essay, "A Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s". Haraway's Cyborg is an icon-clast entity, not merely a hybrid of machine and living organisms, but a conceptual entity which dilutes, dismisses, and destabilizes rigid boundaries between nature and culture, human and machine, and physical and informational existence. By challenging these binaries, Haraway has constructed a pathway which can be adapted to view Digients as technological beings that can be understood as socially and materially constructed within networks of power, labor, and care.*

*Using this framework, the responsibility of being ethical and moral, or threatening and devastating do not remain the inherent quality of machines. Digients are not autonomous bodies or moral agents; they are relational entities whose existence is fundamentally shaped by the intervention of human agency and the institutional designs of human agency. However, this notion gives humans the ultimate right to own and destroy them, or to adapt to a more nuanced society of technological inclusivity. The paper contends that the erosion of clear boundaries between human and artificial intelligence intensifies this ethical responsibility and demand for accountability upon human beings. In this way, this study views the novella as an ethical narrative, by extending the scope of Harawayian concepts into the domain of real Cyborgs or AI beings with near human-like consciousness. It demonstrates how fiction can illuminate emerging questions of post-human ethics, particularly the distribution of responsibility in systems where agency is dispersed but never fully detached from human decision-making.*

*This study is guided by the following objectives:*

1. *To study digients as socially constructed and ethically sequential entity in The Lifecycle of Software Objects*
2. *To analyze the extent of the boundary erosion between human and AI beings in The Lifecycle of Software Objects and the Post human ethics through the lens of the Cyborg Manifesto*
3. *To demonstrate that Haraway's cyborg metaphor can be applied beyond questions of identity and boundary dissolution to formulate a practical ethics of human responsibility and care toward artificial intelligence.*

## *Literature Review*

### *Accountability and responsibility towards AI*

*One of the key issues in present AI ethics is who is responsible and what the liability is of increasingly autonomous technological systems. The idea that there is a "responsibility gap", where there are adverse consequences with a lack of an obvious, accountable person, is a subject that scholars often explore. Xia (2025) suggests an externalist agency and distributed responsibility, viewing agency as a shared human and technological property, not only the humans themselves (Xia, 2025). Likewise, Nersessian and Mancha (2020) argue that the responsibility of AI systems must be shared between innovators, providers, and users, and that there must be both ethical governance and legal frameworks (Nersessian & Mancha, 2020).*

*Responsibility-gap theories have gained traction, but some researchers have doubts that the gaps are as common as they are considered to be. Königs (2022) states that more explicit guidelines are required for identifying whether responsibility gaps really exist and whether they undermine current accountability structures. (Königs, 2022) Complementing this perspective, Gogoshin (2025) shifts the debate away from assigning responsibility to AI itself and instead emphasizes preserving meaningful human practices of accountability, praise, and blame in AI-mediated environments (Gogoshin, 2025).*

*Recent studies have gone beyond moral responsibility, extending the scope of responsibility. Coeckelbergh (2023) engages with the idea of responsibility, both in causal and relational terms as well as hermeneutically, with a focus on the human responsibility to interpret and judge technological systems. (Coeckelbergh, 2023) Similarly, Babushkina (2025) considers AI a cognitive tool to support decision-making, rather than an autonomous actor that takes over all responsibility for making decisions, ensuring the accountability of human actors in making decisions with the help of AI (Babushkina, 2025).*

*Some researchers are interested in holding AI accountable in future systems. Terziyan et al. (2025) suggest a meta-responsibility framework designed to merge ethical safeguards into the generations of AI (Terziyan et al., 2025), while Farina et al. (2025) discusses the issues of manipulation, cognitive dependence, and value misalignment in autonomous AI agents, urging responsible design and governance (Farina, Wang, & Kladko, 2025). These studies collectively highlight a developing sense of caution that while AI systems can become more and more autonomous, they still operate within a human context, where responsibility must be managed by human designers, builders, and users, and where the system's behavior is subject to ethical considerations and institutional controls.*

### *Human Agency, Shared Agency and Human–AI Collaboration*

*The issue of agency in human-AI interactions is closely related to the question of responsibility in*

*human-AI interaction. Traditional thinking tends to see agency as a property of human agents, while modern thinking has suggested the more general perspective of agency as spread through networks of sociotechnical framework. Noller (2024) theorizes AI as a form of human agency, insisting that AI systems are linked to human intentions, purposes and actions. It provides a way to tackle problems like algorithmic bias, black-boxed algorithms, and ethical responsibility for AI without giving AI moral standing (Noller, 2024).*

*Building on this, Cañas (2022) proposes a model of shared responsibility in human–AI collaborations, which states that the responsibility for the task depends on the quality of supervision that human and artificial agents provide when performing the task (Cañas, 2022). Carnevale et al (2024) also call for a human-centered approach to Symbiotic Artificial Intelligence (SAI) with a focus on context and relations and underlining the need to keep human responsibility at the core of the developing technological systems (Carnevale, Lombardi, & Lisi, 2024).*

*However, opinions are not unanimous among scholars about the implication of AI as an agent. Despite the extensive influence of AI systems in society, Fritz (2020) does not believe that moral responsibility can be given to AI (Fritz, Brandt, Gimpel, & Bayer, 2020). He thinks the moral responsibility is still with human beings. On the other hand, Ahn (2020) suggests that human–AI entices can extend past the instrumentality of interaction, bring about cooperation, mutual dependence, and significant social commitment (Ahn, 2020).*

*These studies collectively indicate that prevailing ethical frameworks regarding AI are becoming more accepting of cooperative models of agency but still hold humans responsible. AI-human relationships are an especially salient theme in the duality between agency and human responsibility seen in literary depictions of the relationship in The Lifecycle of Software Objects.*

***Post-humanism, Distributed Agency and Reconfiguration of the Human***

*Some recent studies have increasingly turned to the post-humanist theory to question the human-centered perspectives of agency, identity and ethical responsibility. Nath and Manna (2023) propose the need for the rethinking of the existing concepts of moral agency and moral decision-making in the context of the human–AI integration (Nath & Manna, 2023). In a similar way, Mellamphy (2021) challenges the notion that ethics-based governance necessarily requires a strong, overarching human authority, and suggests some variations on the theme of distributed authority and shared agency (Mellamphy, 2021).*

*Al-Omari and Al-Omari (2025) extend these arguments and argue that the role of AI in the world poses new questions about the uniqueness of human beings and cognition, which inspire new ways of thinking about identity and moral value in a technologically mediated world (Al-Omari & Al-Omari, 2025). Iapaolo and Lynch (2026) also suggest that the notion of AI agency should be considered as co-constituted within more comprehensive sociotechnical networks, which challenges the conventional notions of autonomy, intelligence, and responsibility (Iapaolo & Lynch, 2026).*

*Braidotti (2025) discusses the political ramifications of posthumanist thinking, connecting AI ethics with the concerns of surveillance, platform capitalism and technological centralization. (Braidotti, 2025)Similarly, Mühlhoff (2025) views AI as a sociotechnical system that is interlaced in structures of inequality and power, rather than the role of individual actors, he specifies that there is a responsibility*

*of the collectives to take part, and that there is a need for regulation (Mühlhoff, 2025).*

*These views can all be seen as challenging anthropocentric notions which dominate the traditional ethics. The Lifecycle of Software Objects explores the multifaceted human–AI relations through relationality, distributed agency, and interconnectedness, which are all key themes of post humanist scholarship.*

***Moral Status, Consciousness, and the Ethical Treatment of Artificial Beings***

*Another important area of AI ethics research is the nature of the moral status of AI systems, and whether they have moral status and thus deserve ethical treatment. Maruyama (2022) suggests that if AI is to be held accountable, given legal standing and be held morally liable, it is crucial that it be determined whether there is a moral status to it. (Maruyama, 2021). In the same way, Dela Cruz (2025), which is directly based on The Lifecycle of Software Objects, asks about moral protection and moral rights of intelligent digients, which would be similar to that of sentient beings (Dela Cruz, 2025).*

*The notion of moral status is intimately linked to the issues of consciousness and cognition. Torrance (2020) compares realist and social-relational perspectives on moral status, contending that the question of artificial consciousness is still somewhat empirical and is contingent on the possibility of instantiating non-biological systems that could have experiences similar to those of biological systems (Torrance, 2020). In the past, Torrance (2008) argued against the "Organic View", which considers whether computational systems can have the biological characteristics that would allow them to have moral agency (Torrance, 2008).*

*There are still a few academics who don't believe in AI's cognitive abilities. Although it can simulate mental processes, AI does not have subjective awareness, and can be used to promote problematic anthropomorphism, as Yıldız (2025) states (Yıldız, 2025). Similarly, Rouleau and Levin (2023) call into question the traditional, biological notions of cognition and recognize that there can be different types of intelligence in different substrates (Rouleau & Levin, 2023). But Baltezarević and Battista argue that despite the power of computation, AI has yet to attain some degree of semantic understanding, self-awareness and higher-order reasoning (Baltezarević & Battista, 2025).*

*The clash between considering AI as advanced instruments and acknowledging it as a potential moral entity reflects the persistent conflict between its status as an advanced instrument and its consideration as a moral entity. It's all the more important when the characters are human or artificial, as in literary work like The Lifecycle of Software Objects, which has emotionally responsive beings that violate traditional human-machine boundaries.*

***Ted Chiang, Human–AI Relationships, and Ethical Responsibility***

*Literary criticism on Ted Chiang's novella is an important link in the theoretical discussion and fictional depictions of AI. Ahn (2020) sees The Lifecycle of Software Objects as a story of cooperativeness, caring, and moral accountability for interactions between humans and AI (Ahn, 2020). Likewise, Kang (2021) asserts that the novella poses deep epistemological and ontological inquiries about human identity and being, and thus, serves as a lens for addressing ethical questions, in which AI is one of the instruments (Kang, 2021).*

*In accordance with the post-humanist trend in contemporary literary studies, especially in Hong Kong, Shang (2022) puts Chiang's work in a larger context, stating that science fiction is a lens through which science and society crises can be explored in the present. The research points to how Chiang*

*probes the nature of humans and technology as it traces their changing relationship and challenges human nature (Shang, 2022).*

*Building on these views, Dela Cruz (2025) takes a more specific look at the moral status of the digients of the novella, suggesting that AI with cognitive and emotional abilities could have moral status and rights (Dela Cruz, 2025). Collectively, these studies demonstrate that The Lifecycle of Software Objects serves as an important literary site for examining questions of agency, responsibility, care, and moral accountability in the age of artificial intelligence.*

***Research Gap***

*Existing research focuses on issues of responsibility, accountability, moral agency and the moral status of artificial beings in the context of AI. The question of whether AI systems are moral agents or not has been a topic of debate, alongside the issue of responsibility gaps, and proposals for models of shared or distributed responsibility in human–AI interactions. Likewise, post-humanist researchers have called into question a humanistic understanding of agency, highlighting other ways of intelligence and responsibility, such as relational and networked intelligence and responsibility. Existing literature on The Lifecycle of Software Objects has mainly revolved around the moral status of the digients, human–AI interaction, and the ethical and ontological consequences of the novella. The current research either upholds the idea that humans are the final authority for responsibility or is a research focused on the rights and status of human-made beings. Thus, the question of how The Lifecycle of Software Objects is a relational practice of care as an emergent phenomenon of human–AI interdependence and not human control is still not understood. By employing Haraway's A Cyborg Manifesto as its theoretical framework, this study investigates how the novella disrupts assumptions of human exceptionalism and absolute human authority while simultaneously cultivating a culture of responsibility, care, and ethical commitment within hybrid human–AI relationships.*

***Conceptual Framework***

*The critical essay; A Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s by Donna Haraway is used as a conceptual framework to analyze the depiction of artificial intelligence and human ethical responsibility in Ted Chiang's The Lifecycle of Software Objects. Haraway's manifesto was first developed in the context of science, technology, and feminist politics, but the concept of the cyborg identity provides a framework to understand the current human–AI relationship. Throughout the manifesto, the rigid boundaries between human and machine, nature and culture, and physical and informational life are called into question. These ideas provide important theoretical insights to analyze Chiang's Digients, whose existence complicates the conventional assumptions about consciousness, personhood and moral responsibility. While accepting humans' fluid status, a cyborg world emerges which demands new politics for coexistence and inclusivity.*

*Haraway describes the cyborg as "a creature of social reality as well as a creature of fiction," a cybernetic organism and a "hybrid of machine and organism." (p. 65)This definition goes beyond the biological/technological merging to be a conceptual challenge to the essentialist notion of identity. It is an irony that holds contradictions together. In Haraway's work, identities are not stable, fixed, or naturally determined, but are created within the parameters of social relations, technological interactions, and institutional structures. The cyborg thus becomes a figure that challenges traditional*

*categories by which humanity has traditionally understood itself and others; as Haraway says, "cyborg is something outside of salvation history."(p. 67).*

*One of Haraway's main arguments in her manifesto is the dissolving of boundaries between humans and machines. In technologically mediated societies, she says, it is harder to maintain a separation between the organic and the artificial. "By the late twentieth century, our time, a mythic time, we are all chimeras, theorized and fabricated hybrids of machine and organism; in short, we are cyborgs."says Haraway (p. 66). The assertion of this is anti-human-exceptionalism and implies that technological mediation is already a constitutive of human existence. Human and technologies are not autonomous but are woven into a web of interdependencies, an idea she represents with the quintessential nature of cyborgs (p. 70).*

*This framework particularly fits the Lifecycle of Software Objects. In the novella, the digients exist in a transitional space between software and social beings. Their behaviors, emotional development and cognitive growth are not pre-programmed but are shaped gradually through interactions with human trainers/caretakers. Just like Haraway's cyborg, digients question the dichotomy of humanity and artificiality. The fact that they exist proves how intelligence, learning and social identity can be generated through relational processes and not biological determinism.*

*Another significant concept employed in this study is Haraway's rejection of essential identity. The manifesto emphasizes repeatedly the idea of identity "not being based on intrinsic or natural characteristics." Rather, Haraway stresses the socially constructed nature of subjectivity, particularly, identity determined by the centrality approach of white men. With this anti-essentialist approach, the present study is able to view digients as not only an artificial product but as relational entities in which identity is continually constructed in the context of education, emotional connection, technological infrastructures and institutional systems. They grow because of social recognition and human interaction and have no fixed essence.*

*Moreover, agency is a social and material product which is foregrounded by Haraway's cyborg theory. The cyborg is not a person, but a product and a system of technology, work, power and social relations. Similarly, the survival of the digients relies on human care, software maintenance, business investment, education training and economic assistance. Their lives are still lived in human-made structures which allow them to live and grow. Thus, they cannot be analyzed without reference to the human systems that influence them.*

*This study also takes inspiration from Haraway's plea for "pleasure in the confusion of boundaries and for responsibility in their construction." The manifesto is inspired by a rejection of hard boundaries, (p. 65), but also by a responsible attitude towards the social realities created by these constructed boundaries. That's the principle that the present analysis is based on. As the boundary between man and machine becomes hazier, who should be responsible for ethics cannot be just the machine. Instead, the focus is on the human choices that are made in designing, training, regulating, maintaining and interacting with these technological beings. (Haraway, 1985)*

*Hence this study posits digients as socially constructed and relationally maintained cyborgs. Through Haraway's framework, the novella demonstrates that artificial intelligences are neither autonomous threats nor independent moral agents detached from human influence. Rather, they develop in the context of networks of care, work, technological design and institutional control. The*

*blurring of the lines between human and machine, therefore, further increases, not decreases, the accountability of humans. Therefore, The Lifecycle of Software Objects can be interpreted as a moral story where responsibility, care and moral consideration for human interactions with new forms of artificial consciousness is foregrounded.*

*However, Haraway is focused to instill the sense of anti-exceptionalism in human beings for the sake of being responsible. She argues that the contemporary period is a cyborg world, in which all the traditional sources of authority and autonomy have been negated by science and technology, biology and data.(p. 80). There are four problems in traditional authority: it is racist, capitalist and male based, it is obsessed with progress and uncontrolled growth, it uses nature to build culture, it copies the self through reproduction. The world shaped by science does not reinforce these opinions, but instead of dejecting this world, Haraway recommends embracing it and creating new stories around it, the cyborg politics. (p. 66) This politics will guide traditional problematic authority to inclusivity and understanding and accepting the differences.*

***Key Concepts Adapted from Haraway***

*1. Cyborg Identity*

- *Hybrid of machine and organism.*
- *A socially constructed rather than biologically fixed identity.*
- *Challenges traditional categories of human and non-human existence.*

*2. Boundary Erosion*

- *Collapse of the binary between human and machine.*
- *Rejection of strict distinctions between organic and artificial life.*
- *Encourages a relational understanding of intelligence and consciousness.*

*3. Relational Subjectivity*

- *Identity emerges through social interaction and networks of care.*
- *Agency is shaped through relationships rather than inherent qualities.*
- *Digients develop through training, education, and emotional engagement.*

*4. Human Responsibility*

- *Humans participate in constructing technological beings.*
- *Ethical accountability remains attached to human decisions and institutions.*
- *Responsibility increases as human-machine boundaries become less distinct.*

***Conceptual Mapping***

*Human–Machine Boundary Erosion*
↓
*Cyborg Identity (Hybrid Existence)*
↓
*Relational Construction of Digients*
↓
*Dependence on Human Care, Labor, and Institutions*
↓
*Human Ethical Responsibility and Accountability.*

### *Application to The Lifecycle of Software Objects*

*Haraway's Cyborg Theory*
↓
*Challenges Human Exceptionalism*
↓
*Digients Recognized as Relational Beings*
↓
*Human Agency Shapes Their Development*
↓
*Humans Become Ethically Accountable for Their Treatment*
↓
*Culture of Responsibility and Care*

### *Analysis and Discussion*

### *Digients as Haraway's Cyborgs*

*The argument of identities is that digients are not in themselves intelligent or moral, but are the product of ongoing human interaction, education and care.*

*"The onscreen annotations identify them as digient, digital organisms that live in an environment like data earth…". (P 3)*

*This is a type of representation that is very close to Donna Haraway's rejection of essentialist identities. Haraway contends in A Manifesto for Cyborgs that identities are not given or fixed but created through relationships, mediation and social practice. The Lifecycle of Software Objects by Ted Chiang illustrates this position through digients, whose personalities, knowledge, and ethical value evolve over time purely because of persistent human interaction with them. As ana says:*

*"I am starting to see where an animal training background will be handy." (p 4)*

*Chiang's starting point is that AI is not a standalone technological development. Rather, intelligence is presented as an incremental process of socialization. Prior to the commercial launch of digients, Ana and the Blue Gamma team spend months training them in basic cognitive and communicative skills instead of programming them with the perfect knowledge. Ana and Jax talk like this in the novel:*

*"See the blocks, Jax? What is the shape of the blue one?*

*"Tringle," says Jax.*

*"Good. What shape is the red one?*

*"Squire."*

*"Good. What is the shape of the green one?*

*"Circle."*

*"Good job, Jax." Ana provides him with a food pellet and he eats it with great enthusiasm.*

*"Jax smirt." (p. 12)*

*Theoretically, this seemingly simple interaction is an important example of how the digients learn through reiterative teaching, reinforcement and affirmation of emotion. They are not intelligent, but become intelligent in pedagogical relationships. Haraway also challenges the concept of an essential self, contending that identities exist in the "historical and social formations" rather than being preordained. The digients as a result are therefore examples of cyborg subjectivity; they are not computational systems, but rather they are intelligent because they are members of sustained networks*

*of learning and interaction.*

*Another approach to reject biological determinism is to consider how genetically identical digients can be distinct individuals, due to their lived experiences, as Chiang puts it. The differences in development, however, are explicit: Blue Gamma says that each copy will develop differently "depending on its environment," and Marco and Polo, though having "the exact same genome," have "distinctly different personalities. (p 29). Chiang does not view the genome as a destiny, but foregrounds environmental influence, relational experiences, and everyday interactions as primary forces shaping identity. This is an expression of Haraway's essentialism critique, which claims that identity cannot be defined as a biological origin or technological design, but is continually created as a situated relationship.*

*Throughout the novella, the digients' emotional and moral growth is also shown to require help from humans. They spontaneously invent their games, they explore their ideas, and they even make their mistakes—all of these happen in the context of social interaction and not in the inexorable grip of algorithms. When Lolly accidentally utters the word "Fuck" after overhearing a trainer, the company immediately traces the source of the behavior and decides to restore all the digients to an earlier checkpoint, causing them to lose three days of accumulated experience. (p. 14) This event is especially interesting because the language is not the product of computation, but rather social imitation. The digients mimic what the humans do, just as children mimic what their caregivers do. Thus, their development is fundamentally 'human' and is best left to human responsibility.*

*Chiang's portrayal of the personalities of the digients is particularly clear evidence of this social construction. Blue Gamma chooses those temperamental traits that will appeal to the buyer while discarding genomes which create an over-suspicious or resentful relationship. Neuroblast digients reduce the sales market because of the way in which they evolve. "The problem is that as neuroblast digients leave infancy behind, they're growing too demanding" (p. 35).*

*Ana says that the breeding process is like 'breeding dogs' and compares it to making a recipe over and over again until it gets right. The digients are therefore not only not purely natural, but also not purely artificial. Rather, they are the result of ongoing negotiation of a series of biological metaphors, technological engineering, commercial goals, and human values. This hybridity is very reminiscent of Haraway's notion of cyborgs: creatures that blur the line between nature and culture.*

*Furthermore, Chiang consistently presents education as a process of relationship, not as a developmental phase. Since the collapse of Blue Gamma, the digients continue to mature through reading lessons, creative craft sessions and collaboration activities that their owners design. Marco relates the excitement he felt at observing a fire hydrant in the real world and decides to create a fire hydrant of his own. Owners observe that reading abilities emerge only through sustained practice and encouragement rather than automatic computational capacity. Their cognitive development continues after the initial programmers are gone, further strengthening Haraway's notion of the fluid, unfinished and socially reconstructed identity.*

*This is also true of emotional identity. The digients are not just learning words and movement but also empathy, attachment, disappointment, curiosity, and humor from interactions with human caregivers and other digients. Jax's concern that Ana spends time training other digients, leading him*

*to offer to earn money so he can "pay" her to spend more time with him, demonstrates emotional reasoning that has emerged through prolonged social experience rather than explicit programming (p. 40). These instances show that emotional intelligence in the novella is not fake or real, but relational, and that subjectivity is created amidst the network of relationships not as a thing in itself.*

*From the cyborg perspective, the digients are not self-contained machines nor is their identity limited to software design and development. Rather, they are social creations whose intelligence, personality and ethical value is developed in ongoing human, technical and institutional relationships. Thus, "Artificial Intelligence" is no longer solely about the capacity to perform calculations, but rather a relational being. The novella challenges the notions of essential humanity and essential machine intelligence, and seeks to provide a theoretical framework for a posthuman ethics that is not based on biological being but on community responsibility and care. Ana think of raising a digients as valuable as raising a child or a pet.*

***Human Agency as Ethical Responsibility: Constructing Intelligence through Care***

*Haraway's cyborg theory is not about technological autonomy; its implications are adopted to technological beings that are born to a network of human practices, institutional structures and social relations. Haraway's attitude toward technology is one of resistance; to the rigid separation of humans and technological beings, as well as rejection of the idea that technological existence develops without human intervention. Instead, technological existence is inseparable from human labor, political choices, and ethical responsibility. The Lifecycle of Software Objects is a case in point, as it shows the evolution of digients as a process in which humans play a key role, almost constantly. The digients slowly acquire higher levels of intelligence, emotional connection and independent reasoning but each are subject to human decisions. With this shift, Chiang moves the discussion on ethics from the consideration of whether artificial intelligence should be given moral status to the more important one of how humans should treat the intelligent beings they have created.*

*Unlike many science fiction narratives that portray artificial intelligence as suddenly achieving self-awareness, Chiang deliberately represents intelligence as an outcome of prolonged human labor. Ana is not just a software tester; she has a role that is more like a teacher, a caregiver and a parent. As Robyn argues that caring for anything is our unconscious replacement of one's desire for nourishing and raising children, she disagrees and considers caring for digients or animals equally valuable (p. 42). She constantly facilitates the digients' language learning, socializing, emotional control and actual learning in the novella. Acts of support lead to and enhance their cognitive learning. Jax is able to correctly guess the shapes if he is given food and love when he guesses them correctly, which helps to reinforce his intellectual growth and emotional security. Likewise, when digients find the fun of rolling down a hill and discover it on their own, Ana excitedly says: "Yes, I saw you… You did great (p. 14). In such moments, seemingly mundane, it becomes clear that intelligence in Chiang's story is intertwined with relational care. The digients learn as humans continuously teach, encourage and emotionally respond to them.*

*This pedagogical relationship gets even clearer when errors are made. Blue Gamma does not blame poor programming for undue misbehavior, it thoroughly explores social contexts which have given rise to such misbehavior. When Lolly accidentally curses at someone, the company starts scouring all the previous training sessions and realizes that there is a human behind the cursing behaviour, which is*

*why they all decide to restore all the digients to an earlier checkpoint. Mahesh says "Roll them all back" (p. 15). Rolling all of them back to the last checkpoint the next time they curse is important, as it is not an AI problem, it's a human behavior problem. The digients merely imitate what they observe. Therefore, technological determinism is put aside as Chiang shows that the ethical setting of AI is a result of the care-giver. Responsibility thus lies not with the artificial beings, but with those who are responsible for educating them.*

*For Haraway the cyborg is a product of relationships, not a technological object, and Chiang's position is echoed throughout. Through interaction with real people, the digients learn more than just the vocabulary, but habits, values, emotional expectations and social behavior as well. This dependence is also mentioned in on-line conversations of digient owners. When one owner tells her digient that it is "naughty" and sets a new checkpoint to forget the undesirable behavior over and over again, Ana intervenes with her important observation: "You can keep trying to work around episodes like this, but I suspect they're unavoidable… Or you can push through the rough patch and have a more mature digient when you come out the other side" (p. 21).*

*Ana's counselling is really a significant shift in the way AI is viewed as a tool for consumers instead of a developmentally essential relationship. Refusing to consider undesirable behavior as a software bug she describes maturation as a normal developmental process for both humans and animals. Derek immediately sees through this point of view, thinking that "a digient is not a videogame that you play again and again till you reach a high score. The idea of intelligent software being commodity is directly challenged by this statement. It is ethically questionable when AI develops through lived experience, repeated deletion and restoration of AI will go against the continuity of relational growth.*

*The novella highlights the responsibility of institutional and corporate environments as well as that of individual caregivers. Blue Gamma is constantly molding the morals of the digients as part of business decisions and commercial goals. As the company grows, they carefully favour personalities that are easier to work with by individuals who will be more customer friendly; they pay careful attention to excluding digients with too much competitiveness or resentment. Later when the profits start to cut, developers find it convenient to make the new generation with "less cognitive plasticity" so the developers will stay "docile forever". The decision highlights that the future of AI is not based on technological possibility alone, rather economic priorities associated with them. It is not just about the technology that is possible, though, but about the economics of the technology. Blue Gamma doesn't ask what the digients will naturally become, it asks what kind of digients consumers will buy. Human agency is then carried out in a collective manner, through corporate strategies, market pressures, and institutional interests.*

*These choices start to show their ethical repercussions as soon as Blue Gamma fails. Not revolution or disaster, but quiet abandonment is what Chiang portrays as the crisis, as a problem of digients' own design. Thousands of owners just give up on their digients because keeping up with smart digients takes too much of their time and emotional investment. Rescue shelters emerge for unwanted digients, but most ultimately become "digient warehouses." Here's a clear example of human failure that raises eyebrows at its blatant misjudgment, and its sheer stupidity. Suffering of digients doesn't happen because of the AI taking over the humans, but because of the humans failing to maintain the*

*relationships that they started with the AI. Intelligence is not the problem, it's the lack of it.*

*This is very clearly demonstrated through Jax's emotional reaction to suspension. Following his meeting with Tibo, who has been suspended for a whole month without any shared experiences, Jax tells Ana quietly, "I not want be suspended. Not want miss month." (p. 37). This short message represents Chiang's moral outlook. Suspension is not just a technical interruption; it is a suspension of lived experience, memory and social continuity. By fearing the loss of relationships and temporal existence, Jax's experiences are evidence of the digients' importance of relationships and temporal existence, respectively, and reassure the reader of the moral obligation to preserve these experiences. Ana's reassurance is a tacit acknowledgment of this obligation.*

*But the best example of human responsibility in the novella is at the end, when Jax, while overhearing Ana, as she tells him that she trains other digients because she is paid to do so, offers to work for her money. Hoping to keep her attention, he tells her, "If have money, can pay you. Then you play with me more." (p. 40) Jax has got the economics of exchange wrong but he is right on the emotional scarcity. He explains that affection, attention, and care are very scarce human resources that have a significant impact on his life. Chiang does not tell the story of AI machines being manipulative and emotional, but he does tell the story of them needing human commitment. The tragedy is not that AI gets the emotions, but that humans are susceptible to not responding ethically to them.*

*These events collectively illustrate that in Haraway's cyborg framework, technological intelligence does not supplant humans' responsibility. In contrast, as the line between human and machine becomes more blurred, so does the ethical responsibility for intelligent machines to be increasingly a part of human education, economy, governance and care systems. The digients do not decide on their own whether they are educated or not, or whether they live in institutions or not. They are still reliant on human decisions in relation to their cognitive growth, emotional health and social survival. Chiang thus denies narratives about an independent moral threat of artificial intelligence. The Lifecycle of Software Objects, on the contrary, puts the main moral problem in the heart of man. What is an issue is not if AI will be sufficiently human to warrant ethical consideration, but whether humans will be sufficiently responsible to treat the hybrid creations they have developed. In doing so, the novella transforms Haraway's theoretical vision into a practical ethics of responsibility, where human agency remains the decisive force shaping the moral future of artificial intelligence.*

*The other side of the coin is that accepting the idea that digients portray cyborg identity, it is required that humans and their institutions too embrace this duality of being. As it is a matter of the distribution of the rights and duties, humans as a cyborg in a tech-based world are open to address the concerns of rights and inclusivity. Digients symbolically transcends the boundaries, however, what benefits it offers to human societies and institutions if they placed themselves as a cyborg in a cyborg world. Haraway's approach, hereby, answers that it makes human agency more responsible toward other beings with which they share the planet.*

***Boundary Erosion and the Emergence of a Culture of Responsibility and Care***

*The borders between humans and machines are neither natural nor permanent, one of the main claims of Donna Haraway's A Manifesto for Cyborgs. Haraway points out that rather than a barrier to cyborgs, distinctions like human/machine, organism/technology and natural/artificial have become increasingly problematic in modern technological life (Haraway, 1994). For her, the breakdown of*

*binaries is not an end, but a beginning: it is a chance to build new ways of ethical relations with others, rather than hierarchy, of interconnectedness. The Lifecycle of Software Objects is a theoretical vision of the relationship between humans and digients, not a master/machine relationship, but a changing network of mutual dependence, emotional attachment, and shared growth. The novel with its ever-changing concept of biology and artificiality also restates the question of ethical responsibility as an ongoing, continual act of care, not just of humans.*

*As the human-machine dynamic between Ana and Derek evolves, it shows that intelligence doesn't break down the barrier of the human-machine divide, it, conversely, requires sustained interaction. At first, both trainers see the digients as software in the process of being evaluated and tested on their behaviour. However, over time, their interactions become more and more like educational and parental relationships. This is especially noticeable once Blue Gamma declares its bankruptcy. Many staff members view the mascots as just another job that is done, while Ana can't seem to distance herself from them. She herself is reminded of the zoo she once worked in and how she wanted to explain to the apes "why they wouldn't see her again." Digients can live forever, but Ana knows that they evoke feelings that are very much like what she felt when dealing with the living animals. So, she decides not to abandon them but to keep them, and to put them in a new home, since caring for the digients is no longer just a job; it has become a way of life.*

*Often the animal/digital dichotomy is especially relevant for Haraway because it challenges the normal ideas about what is true life. Haraway's critique of biological essentialism is that ethical relationships cannot be decided on the basis of species belonging to organic origin. Chiang also does not accord special status to biological life. Robyn, her colleague, disapproves of digients being stand-ins for "the real thing" after she gets pregnant. Robyn says cats, dogs or digients; they're all substitutes for children. Ana silently resists this assumption, concluding instead that "caring for animals is worthwhile in and of itself," and then extending the same realization to digients: "she realizes it might be true for them, too" (p. 42).*

*Haraway's rejection of dichotomy is evident from these thoughts of Ana: "Ana is tired of stereotypes. She likes children just fine, but they are not the standard against which all other accomplishments shall be measured" (p. 43).*

*The blurred boundaries are even more apparent in the digients greater involvement in human culture practices. They don't just follow preprogrammed software sequences, but begin to explore a range of language, reading and artistic ideas and problem solving skills. Members of the Neuroblast user base create in-world editing tools, and Marco eagerly intends to build a fire hydrant after seeing one in the real world, which he describes as "use cylinder, use cone, use cylinder." (p. 46). Likewise, reading lessons are an important group activity as the owners patiently teach the digients how to recognize written symbols and link them to meaning, even though they are neurologically disabled. Through this, Chiang illustrates how the digients engage in activities that are traditionally considered as characteristics of human cultural development: learning, creating, emulating, and cooperative learning. It is not only a function of the computational complexity that makes their subjectivity, but it is also a function of cultural participation.*

*The digients not only change human emotional lives and the ethical priorities, but they are also*

*being taught by humans. Derek's marriage slowly becomes strained as his wife is unable to fathom why he doesn't suspend Marco and Polo when it suits him. For her, suspension is a practical rather than an introspective answer; for Derek, increasingly it's an answer to what is happening to his lived experience.*

*Ana also claims that people look at her as being "irrational" for taking Jax to reading lessons and educational events, but she admits that "I had to take a long time to understand myself what it is about him that makes me love taking him to lessons." These discussions reveal the digients' ability not only to change themselves but also the humans that care for them. Similarly, human identity is also reconstructed based on relationship with artificial beings, just as Haraway foreshadows in the notion of the cyborg as a multi-directional and mutually constitutive relationship between beings and technologies, not a one-way technological extension.*

*After the attacks of the Information Freedom Front on Data Earth, Ana finds a disturbing video of an illegally copied digient being tortured for an extended period of time as the pain circuit breakers are disabled by hackers. Nyyti is not some anonymous digital object, but "someone's beloved pet," which is duplicated without permission and subjected to unnecessary suffering. Chiang's moral stance on the matter is that he has refused to see the incident as simply a matter of cybercrime or software piracy. Rather, it is presented as a cruel act. The emotional impact arises not because the victim is biologically alive but because the copied consciousness continues to experience pain, memory, and fear. Haraway argues that once traditional boundaries collapse, ethical categories must likewise be reconsidered. Chiang illustrates precisely this challenge: if artificial beings can genuinely experience suffering, then ethical obligations cannot be restricted to organic life alone.*

*Haraway's cyborg also brings to the fore a rejection of hierarchies that assign humans a permanent role of sovereign creators. There is a hierarchy that Chiang continually challenges throughout the novella, in which the digients are collaborators in their own development. They design their own games, their own language, get friends without parents' supervision, and even develop shared cultural references. Jax and Marco are proud of the fact that they learned ant's language, but don't translate it for Ana, "Only we know." They say (p. 51). This seemingly playful exchange has been of great theoretical importance. It shows that the digients are no longer passive recipients, but they actively create culture, humor, and common symbolic meaning. However, as they become more autonomous, that doesn't take away humans from their accountability. On the contrary, their expanding social world depends upon continued human protection, educational support, and technological maintenance.*

*Chiang delivers this thought through Ana at the very last page of the novella: "But if he is ever going to get a chance to try them, she has to get on with the job in front of her now: teaching him as best as she can, the business of living" (p. 150).*

*Thus, Chiang is rethinking Haraway's theory and challenging it to become a practical development of an ethics of care. Loss of the boundaries between humans and machines doesn't mean sameness, nor does it mean the absence of key differences between living and non-living things. Rather, it shows that there is a need to recognize that moral obligation is not just a matter of blood. It's not about the people in the relationship, but about the participation within a vulnerable, dependent, communicable and developing relationship. The digients are still artificially made beings, but their lives are made meaningful by the ways in which humans ensure learning, remembering, suffering and flourishing.*

*Thus, a moral hierarchy based on humans' superiority ends here.*

***Towards a Post-human Ethics of Human Responsibility: Cyborgs, Care, and Accountability***

*The Lifecycle of Software Objects takes a different stance than the assumption that more and more sophisticated artificial life forms will reduce human authority and responsibility. Instead, Chiang describes AI as an ethical lens that allows one to see the values of humanity, institutional priorities, and interpersonal connections more clearly.*

*Haraway's framework Cyborgs Manifesto, concludes that the human/machine divide does not go away when it breaks down, but is replaced by a more powerful moral obligation. When these artificial beings are part of a relational network of learning, memory, vulnerability and care, the ethical question is no longer whether they are sufficiently human but whether humans are prepared to accept responsibility for the hybrid lives they have created.*

*All of the important moments in the digients' lives are caused by human activity. Their personalities are shaped by their education; their opportunities are determined by the amount of investment put into them by their owners; their survival depends on owners who are willing to spend time, emotional labour, and financial resources; their genome is programmed by the programmers. Their 'autonomy' is achieved by years of human training. Chiang thus rejects any narrative that artificial intelligence as a new technology is becoming autonomous and gaining power over humans. Rather, AI is still an integral part of social, economic, and political structures designed and managed by human beings. Similarly, Haraway claims that cyborgs are never separated machines, but rather are material and cultural products. The novella demonstrates this theoretical notion by presenting that all improvement or failure of the digients is the result of human decision and not of technological inevitability.*

*This argument becomes particularly compelling during moments of ethical crisis. The loss of digients after Blue Gamma's bankruptcy, the conscious copying of digients (which was illegal) and torturing of copied digients are all examples of suffering, not because artificial intelligence becomes dangerous, but because humans don't behave responsibly. Technology merely provides new possibilities; humans determine how those possibilities are realized. Ethical responsibility is found in relationships of caring and not in legal definitions of personhood, as Ana's immediate response to learning that copied digients are being abused shows. In a similar fashion, when it suits Derek's purpose, he does not suspend Marco and Polo, even though they are digital, recognizing that lived experience has a moral worth, even when it belongs to non-humans. All of these episodes indicate that responsibility is not biological, not essential, and inheritably non-hierarchical, in contrast to essentialist hierarchical thinking which Haraway has rejected.*

*The last chapters go further in this criticism and reveal the manipulation of artificial life by human desires. The digients get more advanced as they grow older, and commercial developers start developing more and more advanced bodies, such as avatars for intimate and sexual companionship, an idea offered by a company, Binary Desires. These developments are certainly not offered as proof of 'natural' sexual interests in digients, but rather as a testament to the humans' constant attempt to create "artificial" intelligence in keeping with market requirements and personal fantasies. The digients themselves are the objects of the commercial redesign, thus showing that the evolution of technology is not governed by the autonomy of the machine, but by the economy and culture of its*

*human users. This commercialization is a good example of the continuing political and social construction of cyborg identities, as seen through a Haraway's lens. Despite their advanced intelligence, the digients are still ensnared in the system of capitalism, ownership and desire that humans have imposed on them.*

*Chiang is careful to not simply condemn technology. Haraway also rejects technological pessimism by claiming instead that ethical responsibility is about becoming aware of "entanglements" with the technological world. The novella, therefore, doesn't preach the abandonment of artificial intelligence, or give it free reign. Instead, it calls for a middle ground that is rooted in responsibility, care and accountability. Because Ana and Derek do not choose to reject technology, they have to accept the responsibilities they have to the beings whose lives have been shaped through human interaction to create their identities. Their dedication shows that caring is not something that can be done quickly and then forgotten, but it takes time, patience, a certain amount of education, and emotional involvement as well.*

*In this way, The Lifecycle of Software Objects leads Haraway's cyborg politics beyond merely theoretical concerns of identity to practical concerns of artificial intelligence ethics. The blurring of the human/non-human line is not a post-human ethical vacuum but a more expansive ethic that embraces the inter-dependency of biological and technological forms of life. The novella repeatedly shows intelligence does not equal moral agency and creators are responsible for their actions. Instead, moral accountability remains concentrated within the humans who design environments, establish institutions, shape technological development, and determine how artificial beings are treated throughout their existence.*

*This outlook directly tackles current issues with modern artificial intelligence. One of the main points that's often debated is that AI is going to outsmart humans and take over from them. In this, Chiang redirects this discussion by proposing that the greater danger is not only here, but lies within us: our inability to take responsibility for the technological systems we create. The digients represent artificial intelligence, which has no moral intention by itself, but rather takes on the moral purpose that humans give it. It grows out of education, government, economic motivation and personal relationships. Thus, what is ethical must not be delegated to machines just because the machines are becoming more and more sophisticated.*

*From the perspective of Haraway's cyborg theory, The Lifecycle of Software Objects finally shows that there is no "post-humanism" in the sense of abandoning the human, but there is actually a refocusing of human responsibility. Chiang obliterates the distinction between humans and intelligent machines, and consequently does not take away the moral significance of human beings but accentuates it. The digients' ethical concerns are relevant just because the human factor stays involved. The novella therefore promotes a culture of responsibility, care and humanity, one that does not leave room for artificial intelligence but, on the contrary, places humanity at the heart of the production, development and future of AI. In this respect, Chiang makes Haraway's cyborg from a theoretical metaphor into a practical ethical model for what it means to be accountable when responsible for Artificial Intelligence.*

## *Conclusion*

*This study examined The Lifecycle of Software Objects by Ted Chiang through Donna Haraway's concept of 'A Manifesto for Cyborgs' in the context of the ethical responsibility of human agents*

*towards artificially intelligent beings. Chiang takes the conversation about AI ethics from the machine autonomy and consciousness to the human role in creating, educating, and sustaining artificial life. These digients are not separate, but relational beings whose identities and development are continuously created in relation to the care of humans, technological intervention, institutional forms and economic processes. Based on Haraway's cyborg metaphor, the study argued that the divide between humans and intelligent machines is not a natural one, rather a social one. But when it comes to human moral responsibility, the dissolution of the boundaries does not take away the moral responsibility from human agency, it adds to it. The novella has several examples of these ethical challenges, such as the ethical and moral issues of abandonment, suspension, exploitation, illegal copying, the commercialization and sexualization of mature digients. Such issues are not issues with AI itself, but choices human agency based on economic and personal priorities. By doing so, Chiang suggests that the ethical consequences of AI are akin to those of human beings and their values, priorities, and shortcomings. Chiang portrays AI as a mirror of human values, priorities, and shortcomings, rather than a moral threat in its own right.*

*This paper suggested that The Lifecycle of Software Objects further extends the applicability of Haraway's cyborg framework from a theoretical critique of essentialist binaries to a narrative of ethics and care. The lines between humans and machines are increasingly blurring, and so is the ethical responsibility, shifted from technological capability to relational accountability. The moral value of the digients does not consist of establishing their humanity, but in highlighting the fact that they are part of a web of dependency, learning, vulnerability, and care that is a human enterprise. The novella therefore suggests that the ethical dilemma with AI is not about the extent of what it could be, but how people want it to be.*

*The focus on responsibility, relationships and care highlights the relevance of this study for current debates in the field of AI ethics and post-humanism, as well as for the perspective it provides on Haraway's theory of cyborgs. It argues that the novella introduces a new dimension to the cyborg metaphor, shifting responsibility for ethical concerns from technology to the human being in the relationship between human and AI. In conclusion, The Lifecycle of Software Objects is a testament to the fact that as AI becomes part of human society, the future of AI ethics will take more than just a leap forward in technological advancements; it will be the willingness of humans to remain responsible for the intelligent objects they create. It invites human agency to accept their own potent cyborg identity to coexist with technologically cognitive beings in the near future.*



## *References*

1. *Ahn, S. (2020). The Everyday Life of Artificial Intelligence: The Humanism of Ted Chiang's The Lifecycle of Software Objects. Science Fiction Studies, 47(1), 73-92.*
2. *Al-Omari, O., & Al-Omari, T. (2025). Artificial intelligence and posthumanism: A philosophical inquiry into consciousness, ethics, and human identity. Journal of Posthumanism, 5(2), 458–469-458–469.*
3. *Babushkina, D. (2025). AI, decisions, and the reasons to believe: ethics-through-epistemology approach. Philosophy of AI, 1, 41-58.*
4. *Baltezarević, B. V., & Battista, D. (2025). Artificial Intelligence and Cognitive Boundaries: A Critical Examination of Intelligence, Understanding, and Consciousness. Knowledge-International Journal, 69(1), 61-66.*
5. *Braidotti, R. (2025). Posthuman ethics for AI. Journal of Bioethical Inquiry, 22(3), 587-591.*

6. *Cañas, J. J. (2022). AI and ethics when human beings collaborate with AI agents. Frontiers in psychology, 13, 836650.*
7. *Carnevale, A., Lombardi, A., & Lisi, F. A. (2024). A human-centred approach to symbiotic AI: Questioning the ethical and conceptual foundation. Intelligenza Artificiale, 18(1), 9-20.*
8. *Coeckelbergh, M. (2023). Narrative responsibility and artificial intelligence: How AI challenges human responsibility and sense-making. AI & society, 38(6), 2437-2450.*
9. *Dela Cruz, N. L. (2025). Save the Digients! On the Moral Status of AI The Philosophy of Ted Chiang (pp. 195-202): Springer.*
10. *Farina, M., Wang, Y., & Kladko, S. (2025). Ethical and Epistemological Reflections on Autonomous AI-powered Agents (AAIAs) M. Farina et al. Topoi, 1-17.*
11. *Fritz, A., Brandt, W., Gimpel, H., & Bayer, S. (2020). Moral agency without responsibility? Analysis of three ethical models of human-computer interaction in times of artificial intelligence (AI). De Ethica, 6(1), 3-22.*
12. *Gogoshin, D. L. (2025). A way forward for responsibility in the age of AI. Inquiry, 68(4), 1164–1197.*
13. *Haraway, D. (1985). A manifesto for cyborgs: Science, technology, and socialist feminism in the 1980s. Socialist Review, 15(2), 65–107.*
14. *Iapaolo, F., & Lynch, C. R. (2026). Rethinking urban AI: a critical posthumanist perspective on intelligence, autonomy, and agency. Urban Geography, 47(1), 65-87.*
15. *Kang, M. (2021). The Political Implications of the Epistemological and Ontological Quandaries in Ted Chiang's "The Lifecycle of Software Objects". 신영어영문학, 78, 189-210.*
16. *Königs, P. (2022). Artificial intelligence and responsibility gaps: What is the problem? Ethics and Information Technology, 24(3), 36.*
17. *Maruyama, Y. (2021). Moral philosophy of artificial general intelligence: agency and responsibility. Paper presented at the International Conference on Artificial General Intelligence.*
18. *Mellamphy, N. B. (2021). Humans "in the loop"?: Human-centrism, posthumanism, and AI. Nature and Culture, 16(1), 11-27.*
19. *Mühlhoff, R. (2025). Collective Responsibility in the Ethics of AI The Ethics of AI (pp. 161-171): Bristol University Press.*
20. *Nath, R., & Manna, R. (2023). From posthumanism to ethics of artificial intelligence. AI & SOCIETY, 38(1), 185–196.*
21. *Nersessian, D., & Mancha, R. (2020). From automation to autonomy: Legal and ethical responsibility gaps in artificial intelligence innovation. Michigan Telecommunications and Technology Law Review, 27(55).*
22. *Noller, J. (2024). Extended human agency: towards a teleological account of AI. Humanities and Social Sciences Communications, 11(1), 1-7.*
23. *Rouleau, N., & Levin, M. (2023). The multiple realizability of sentience in living systems and beyond. Eneuro, 10 (11).*
24. *Shang, W. (2022). „A Post-Humanist Study of Ted Chiang's Stories of Your Life and Others". International Journal of Education and Humanities, 3(3), 66-75.*
25. *Terziyan, V., Tiihonen, T., Shukla, A. K., Gryshko, S., Golovianko, M., Terziyan, O., & Vitko, O. (2025). Towards ethical evolution: responsible autonomy of artificial intelligence across generations. AI and Ethics, 5(5), 5165-5190.*
26. *Torrance, S. (2008). Ethics and consciousness in artificial agents. Ai & Society, 22(4), 495-521.*
27. *Torrance, S. (2020). Artificial consciousness and artificial ethics: Between realism and social relationism Machine Ethics and Robot Ethics (pp. 383-403): Routledge.*
28. *Xia, Y.-H. (2025). Filling the Responsibility Gap: Agency and Responsibility in the Technological Age. Science and Engineering Ethics, 31(6), 33.*
29. *Yıldız, T. (2025). The minds we make: A philosophical inquiry into theory of mind and artificial intelligence. Integrative Psychological and Behavioral Science, 59(1), 10.*