

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
UNIVERSITY OF MOHAMED BOUDIAF - M'SILA

FACULTY OF LETTERS AND LANGUAGES
DEPARTMENT OF LETTERS AND ENGLISH
LANGUAGE
N°:



DOMAIN: FOREIGN LANGUAGES
STREAM: ENGLISH LANGUAGE
OPTION: LITERATURE & CIVILIZATION

**“To Teach It Grief”: AI, Reality, and Identity in
the Posthuman Age through Richard Power’s
*Galatea2.2***

Dissertation Submitted to the Department of Letters and English Language in
Partial Fulfilment of the Requirements for the Master’s Degree in
Literature and Civilization

Candidates :

Ms. Aya Melak AOUIDIDI

Ms. Nour Elhouda DJEDIAT

Board of Examiners

Ms. Amel BENIA	University of M’sila	Chairperson
Dr. Bachir SAHED	University of M’sila	Supervisor
Ms. Khawla REBAHI	University of M’sila	Examiner

2025

Declaration

We, Aya Melak AOUIDIDI and Nor Elhouda DJEDIAT, hereby declare that this dissertation titled “To Teach It Grief”: AI, Reality, and Identity in the Posthuman Age through Richard Power’s *Galatea2.2*” is the result of our own research. All sources used have been acknowledged and cited appropriately. We confirm that this work is original and is entirely our own, except where explicitly stated otherwise.

Signatures :

Nour Elhouda DJEDIAT



Aya Melak AOUIDIDI



Acknowledgements

We would like to extend our deepest thanks to our teacher and supervisor, Dr. **Bachir SAHED**, for his continuous support, guidance, and helpful feedback during the course of this research. His knowledge, patience, and commitment greatly influenced the development and completion of this dissertation.

We are also thankful to the members of the board of examiners for kindly accepting to read and evaluate our work, despite their tight schedules.

We would also like to thank all the teachers at the University of M'sila. Their dedication and efforts have played a crucial role in our academic progress and personal growth throughout our years of study.

Dedication

*To my beloved family, my siblings **Samy** and **Hidayet**, my best friend **Sofia**, and to my partner **Melak** for supporting me emotionally and financially, and for being my greatest motivator. A gift from my heart after a long struggle, proof that I never gave up.*

Nour Elhouda Djediat

Dedication

To my beloved father, my source of strength and endless support. His sacrifices, his strong belief in me, and the values he taught me have shaped who I am today. Thank you for being my greatest inspiration and for teaching me that with hard work and commitment, dreams come true.

To my mother, my first teacher and constant source of love. Your loving heart, endless prayers, and selfless care have been the foundation of my success. Your encouragement during moments of doubt and your celebration of every small victory have helped me through this journey.

To my siblings, my companions and closest friends. Your support, laughter, and shared memories have made this journey less lonely and more meaningful. Thank you for believing in me, for standing by my side, and for reminding me that family is everything.

*To my entire family, whose love has been my sanctuary and whose faith in me has been my driving force. Each of you helped make this achievement possible in your own special way, and I am forever thankful for the bond we share. And to my partner **Nour**, thank you for your constant emotional support, and for always encouraging me to keep moving forward.*

Your motivation has meant everything during this journey.

Aya Melak Aouididi

Abstract

The present study examines the depiction of artificial intelligence, identity, and embodiment in Richard Powers's *Galatea 2.2*. It aims to examine how the novel portrays artificial intelligence as a posthuman subject shaped through relational, cognitive, and affective entanglements. The research seeks to analyse how contemporary posthumanist theory enables an insightful exploration of the novel's treatment of agency, perception, and subjectivity in the course of intelligent machines. To achieve this objective, the study adopts the theoretical frameworks of Maurice Merleau-Ponty, Donna Haraway, N. Katherine Hayles, Karen Barad, and David Gunkel. Therefore, the findings indicate that Powers's novel dismantles human exceptionalism and redefines identity as a hybrid, fluid process phenomenon co-produces through technological and human interactions. Ultimately, the research suggests that the novel presents a critical perspective of posthuman subjectivity and the ethical implications of artificial intelligence.

Keywords: artificial intelligence; posthumanism; embodiment; agency; subjectivity; hybrid identity.

Table of Contents

I. Declaration	I
II. Acknowledgements	II
III. Dedication.....	III
IV. Dedication.....	IV
V. Abstract	V
VI. GENERAL INTRODUCTION.....	1
VII. CHAPTER ONE	10
VIII. Socio-Historical Context and Theoretical Framework	10
1. The Posthuman Condition: N. Katherine Hayles and the Reconfiguration of Humanity.....	11
1.1 The Liberal Subject Unraveled: From Autonomous Self to Posthuman Assemblage ...	12
1.2 Fragmented Selfhood: Dispersion of Consciousness in Technological Networks	13
1.3 Memory, Artificial Intelligence, and the Unmaking of Stable Identity	15
2. Embodied Consciousness and Posthuman Intelligence: A Phenomenological Analysis.....	18
2.1 The Phenomenological Foundation: Merleau-Ponty's Revolutionary Account of Embodied Consciousness.....	19
2.2 Dismantling Cartesian Dualism: The Unity of Mind and Body in Situated Action	20
2.3 The Challenge of Artificial Intelligence: Cognition without Lived Embodiment	22
2.4 Towards an Embodied Posthumanism: Integration and Future Directions	23
3. From Haraway to Posthumanist Theory: Hybrid Subjectivities and the Collapse of Binaries.....	25
3.1 The Cyborg as a Disruption of Binary Oppositions.....	26
3.2 The Cyborg as a Theoretical Disruption of the Male–Female Binary	28
3.3 The AI Protagonist as a Posthuman Cyborg	30
4. AI and Technological Subjectivity: Agency and Intra-action.....	32

4.1 Karen Barad's Intra-action (Agential Realism) and Artificial Intelligence's Ontological Status.....	33
4.2 AI as an Active Co-creator in the Process of Building Meaning and Reality	34
5. Posthuman Ethics and the Construction of AI Identity and Consciousness	36
5.1 AI Agency and Ethical Subjectivity: Posthuman Perspectives from David Gunkel and Karen Barad	36
IX. CHAPTER TWO	40
X. Becoming Machine, Becoming Self: AI, Reality, and the Posthuman Identity	40
1. Posthuman Consciousness and Digital Identity Formation in Galatea 2.2	41
1.1 Helen's Mind: Distributed Cognition and the Networked Construction of Self.....	42
1.2 Language and Loss: Flickering Signification in Helen's Struggle for Meaning	43
1.3 Emergence and Complexity in Helen's Intelligence in Galatea 2.2	44
2. Merleau-Ponty's Digital Embodiment : Situated Knowledge Without Physical Form.....	46
2.1 Reversibility without Flesh: Disembodied Intentionality in Helen's Digital Consciousness	46
2.2 Situated Knowledge in a Disembodied Mind	49
2.3 Digital Flesh and the Future of Embodied Posthumanism	52
3. Hybrid Subjectivities and Co-constitutive Agency	55
3.1 Haraway's Cyborg Theory: Helen as Human-Machine Hybrid.....	57
3.2 The Cyborg as a Theoretical Disruption of the Male–Female Binary	61
3.3 Helen as a Posthuman Cyborg: Rethinking Embodiment and Subjectivity	65
4. Entangled Agencies: Intra-action and AI Subjectivity in Richard Powers's Galatea 2.2	67
4.1 Helen's Becoming through Intra-action.....	68
4.2 Helen as an Active Co-creator in the Process of Building Meaning and Reality	69
5. Posthuman Ethics and the Formation of Machine Subjectivity in Galatea 2.2.....	70
5.1 Beyond Consciousness: Moral Agency and Relational Identity in Galatea 2.2	71

XI. GENERAL CONCLUSION.....	74
XII. Works Cited.....	80
XIII. الملخص.....	83

GENERAL INTRODUCTION

Over centuries, literature has demonstrated a remarkable capacity to guide readers across different times, places, and cultures. It has consistently functioned as a multifaceted force through which humanity tries to understand the world as socially and historically constructed. It reflects on the intellectual and technological changes that change the concepts of identity and corporeality. In this vein, there is a system that has emerged using technologies capable of collecting data and using it to predict, recommend, and make decisions with varying levels of autonomy, Artificial intelligence (AI), which has evolved to be capable of learning, problem solving, analysing, and engaging in interactions that imitate human understanding. It has become necessary to critically examine how technology is transforming our understanding of identity and subjectivity since artificial intelligence continues to develop. This transformation can be seen as “Posthumanist theory”, which is a mode of thinking that emphasizes the importance of recognizing the interconnectedness of all beings for both “human” and “non-human” (Braidotti 2013). This idea has become more prominent in the late twentieth and twenty-first centuries, particularly due to the rise of technological consciousness and innovation that question the role of humans to shape the Earth.

Richard Powers is considered as one of the most significant contemporary American novelists, known for his deep engagement with science, technology, and philosophy. His works explore themes related to the effects of modern science, technology, and human consciousness. He explores the intersections between artificial intelligence, consciousness, and human identity in his semi-autobiographical novel *Galatea 2.2* (1995), which mixes fiction with deep ideas and examines questions about posthumanism, the embodiment, and the nature of self-

awareness. Powers' part philosophical fiction, part social commentary novel tackles deep themes such as posthumanism, artificial intelligence, embodiment, and the evolving nature of identity.

The novel provokes significant questions about consciousness, selfhood, and the ways in which AI contents traditional conceptions of intelligence. Therefore, Powers guides readers through a narrative that prompts critical questions about how technology forms humanity. His writings integrate deep thinking with emotion and support readers' connection with the characters while also reflecting on the future of humanity. *Galatea 2.2* represents posthumanist themes in contemporary literature. Therefore, the primary focus of this study will be the novel's engagement with posthumanist theory and how it examines the emergent interplay between artificial intelligence, embodiment, and selfhood.

In each chapter, the novel illustrates masterfully an AI named Helen, as her evolution questions consciousness, cognition, and the boundaries between human and machine intelligence. The novel challenges traditional ideas about identity, perception, and supports posthumanist views that reject the idea that humans are separate entities from Technology. Richard Powers' novel addresses these issues of the nature of understanding, consciousness, the relationship between artificial intelligence, and human cognition through the story of a scientist who initiates to teach an AI system named Helen how to analyse literature and language.

Furthermore, the intersection of artificial intelligence, reality, and identity raises profound philosophical questions about what it means to be human in a time of rapid technological advancement. Richard Powers's novel *Galatea 2.2* engages with these questions through the creation of a conscious artificial intelligence and its comparison to human consciousness. As technology develops, the barriers between humans and

machines begin to dissolve, which disrupts previous notions of identity, agency, and emotional existence.

There is a noticeable research gap in the existing scholarship on the role of AI in posthumanist novels, particularly in works like *Galatea 2.2*. Many studies on posthumanism and AI have focused on general theoretical ideas. Nevertheless, it is interesting to read Powers' novel since it reflects common posthumanist interests in corporeality, cybernetic thought, and non-anthropocentric subjectivity, and how these are changing conventional understanding of existence and identity. Helen's evolution evolves to challenge the nature of her own existence and forces to reconsider the limits of both machine and human consciousness as it dissolves the lines between them.

Previous critiques of the novel have focused on either the technology or the philosophy, without fully addressing how these elements interact to shape its treatment of identity in the posthuman world. This research attempts to fill this gap by syncretizing the philosophies of N. Katherine Hayles, Maurice Merleau-Ponty, Donna Haraway, Karen Barad, and David J. Gunkel, to examine how Powers's novel *Galatea 2.2* deconstructs the anthropocentric position and imagines a model of identity as relational, fluid, and forged through the co-performances and interaction between humans, machines, and other elements.

In her book *How We Became Posthuman* (1999), N. Katherine Hayles explains that technological advancements have changed how identity is understood, as it destabilizes the divisions between human and machine, body and code, consciousness and system. Although the posthuman condition is not a fusion of the organic and the artificial, it is deeply rooted in complex ethical, philosophical, and technological debates that question human-centred concepts of subjectivity. This idea is explored in Richard Powers's *Galatea 2.2*, where the creation of Helen becomes a posthuman

subject to rethink ideas such as agency, embodiment, and emotion beyond human limits. Therefore, the novel suggests that AI can be seen not just as a tool but as a subject that is capable of relationships and ethical meaning. Based on the theories of Katherine Hayles, Maurice Merleau-Ponty, Donna Haraway, and Karen Barad, this study illustrates how literature can help rethink identity in the age of intelligent machines and explore new ways to understand emotion.

The aim of this dissertation is to develop an understanding about the boundaries between human and machine in Richard Powers's *Galatea 2.2*. It critically explore how the novel depicts the emergence of artificial intelligence as a posthuman subject. This dissertation focuses on how artificial intelligence is represented beyond algorithmic separation to fix human emotion, such as grief. Consequently, the work analyses how the novel engages with contemporary posthumanist theory, using Maurice Merleau Ponty's embodiment theory, Donna Haraway's cyborg theory, N. Katherine Hayles's posthumanist views, Karen Barad's Intra-action, and David Gunkel's machine moral agency.

Taken together, this research identifies several key objectives. First, it examines Maurice Merleau-Ponty's theory of embodied consciousness to evaluate whether an AI entity such as Helen can engage in perception and cognition without a biological body. Second, it adopts Donna Haraway's cyborg theory to analyse how the novel establishes hybrid subjectivity and destabilizes binary oppositions such as human versus machine and male versus female. Third, the research employs N. Katherine Hayles's posthumanist theory with a focus on her concept of distributed cognition to analyse how Helen's identity emerges through networked and non-linear processes that challenge the core assumptions of liberal humanism. In addition, the research engages with Karen Barad's theory of intra-action and agential realism to argue that Helen's agency does not pre-exist but instead emerges through entangled relations

with her environment. Finally, it applies David Gunkel's relational ethics to interrogate the moral dimensions of human-machine interaction, especially in relation to recognition, moral responsibility, and the justification for ethical consideration of artificial intelligence.

As such, this dissertation attempts to answer the following main question: How does Richard Powers's *Galatea 2.2* reconceptualise artificial intelligence as a posthuman subject by examining embodiment, agency, and identity through contemporary posthumanist theory? This question will be answered by introducing three sub-questions: In what ways does the novel challenge traditional humanist conceptions of embodiment and consciousness through Merleau-Ponty's phenomenology and Hayles's theory of distributed cognition? How does the AI character Helen embody a cyborg or hybrid subjectivity that deconstructs binary distinctions such as human/machine and male/female as theorized by Donna Haraway and Rosi Braidotti? How do Karen Barad's concept of intra-action contribute to understanding artificial intelligence as a relational and ethical actant that co-constructs meaning and participates in shaping human identity and narrative form?

Many researchers have shown interest in this topic, and in Richard Powers's literary corpus for their investigation. Critics have praised Powers's profound treatment of the issues of artificial intelligence, reality, and identity in the posthuman age in all his novels, from *Galatea 2.2* (1995) to *Generosity: An Enhancement* (2009).

Among the researchers who have attempted to investigate it is Kucharzewski's article "From Language to Life is Just Four Letters: Self-referentiality vs. the Reference of Self in Richard Powers' *Galatea 2.2*", he suggests that Powers distributes with a real author named Richard, and an AI named Helen, all of whom connect to establish a narrative that refuses one to stabilize coherent notions of

individual identity. This work's autobiographical nature challenges the dichotomy between the author and character, since Helen struggles with self-definition while the notion of identity is further strengthened as an effect of language and narrative. The lessons on literature by Helen then find a place as an articulation of authorship, claiming that both AI and human identity are produced from the textual moments of interpretation. In this way, Powers's *Galatea 2.2* can be seen as a postmodern metafiction that critiques its own artifice while tracing the outlines of perfecting the self into being in both humans and machines. Some of the questions this article raises with respect to reference deal with the proposition of whether AI or autobiographical narratives could provide well expressed ideas of a self with reference to the fact that identity is ultimately mediated by the sign, anchored in posthumanist theories rejecting fixed identities.

Anca Rosu's "Parody as Cultural Memory" in Richard Powers's *Galatea 2.2*, discusses how the novel has obscured the Pygmalion myth through a posthumanist lens. Parody is the intellectual tool used to critique the creator-creation dynamics, AI ethics, and fragmentation of knowledge in the digital age. In Rosu's view, the parody does more than just satirize of the idea that humans are special, artificial intelligence, or the shifting meanings in literature. The AI character Helen acts as a playful mirror of how people think, showing that intelligence might be man-made and challenging our basic ideas about who we are. The novel critiques both literary humanism and scientific reductionism in ways that underline the tensions between the humanities and science while questioning whether literature can humanize AI or whether intelligence can be algorithmically engineered. Rosu's critique positions Helen as a mirror for postmodern concerns about meaning and identity, she emphasizes how the novel critiques human arrogance and authority over technology. Powers' rewriting of the Pygmalion myth redefines the novel as a parody of the

posthuman realm, weaving the literary tradition into a critical engagement with AI ethics and epistemological uncertainty in an increasingly convergence-obsessed epoch of human-machine intersections (Rosu).

On the other hand, Wes Chapman's "The Cognitive Literary Theory of Richard Powers' *Galatea 2.2* illustrates how the novel uses AI as an attempt to turn common presumptions about cognition, intelligence, and self-awareness on their head with cognitive literary theory. Helen equated human cognition to acquire words and sense through the reading of literature in very much the same fashion as human inference works. This puts machine vs. human in a bad light and makes intelligence less innate and more interactive and acquired. Helen's crash later in the novel is an existential, rather than technological, one that destroys AI consciousness, feeling, and intent. Chapman places that into a wider ethical context, i.e., whether or not such advanced AI systems would be granted rights or exploited as instruments, lamenting human arrogance in attempting to assume intelligence. His thesis situates *Galatea 2.2* on the grounds of posthumanist identity theories, following a lead from critics like N. Katherine Hayles in trying to map the way that AI oversteps conventional ideas of self and imagination. Situating Helen's advent in terms of cognitive literary theory, Chapman showed how fiction is increasingly being formed as terrain onto which human AI relationship is being plotted out, exploiting the work's engagement with human exceptionalism and forming the new ethics-consciousness discussion in the framing of nascent future tech.

Besides, this dissertation is going to analyse Richard Powers's portrayal of artificial intelligence in *Galatea 2.2*. Therefore, this novel will be studied from a posthumanist theoretical perspective, as explained by N. Katherine Hayles, "the posthuman subject is an amalgam, a collection of heterogeneous components, a material-informational entity whose boundaries undergo continuous construction and

reconstruction” (3). The meaning is that posthumanism does not seek to reject the human; rather to critically examine how human identity has been historically constructed through exclusions, specifically of the nonhuman, the technological, and the embodied. Likewise, posthuman literary criticism seeks to explore how literature represents new subjectivities that emerge from the intersection of humans and intelligent machines. Likewise, posthumanist theory questions fixed binaries such as human/machine, male/female, and mind/body. In light of this theory, the novel will be put under scrutiny to explore concepts such as “distributed cognition,” “technological embodiment,” and “intra-action,” as the story revolves around the creation of the artificial intelligence Helen, who evolves through literature, emotion, and interaction. These concepts challenge the liberal humanist ideal of the autonomous, rational self, and instead presents a relational model of subjectivity shaped by affect, language, and networked intelligence.

In the posthuman paradigm, the human is no longer the central figure of agency or meaning production. Theories by Hayles, Haraway, and Barad all emphasize that identity is not a pre-given essence but a product of interaction between material and discursive elements. Karen Barad explains, “distinct agencies do not precede, but rather emerge through, their intra-action” (33). This statement emphasizes that entities or “agencies”, are not separate individuals that interact. Instead, they are formed through the process of interaction itself. In other words, identity and meaning are not fixed, they are formed relationally.

This dissertation is divided into two chapters; the first chapter addresses the socio-historical context and the theoretical framework. The first part examines key concepts from posthumanist theory, including embodiment, subjectivity, and no anthropocentric ethics, the chapter explores five major theoretical perspectives, including Merleau-Ponty’s embodiment consciousness, Haraway’s cyborg theory,

Hayles's posthumanism, Barad's intra-action, and Gunkel's relational ethics of artificial intelligence, establishing the foundations for analysing the novel. The second chapter is dedicated to analyse Richard's novel through a posthumanist lens. It is divided into three main sections. The first section examines the representation of embodiment and perception in Helen's development, as it focuses on Merleau-Ponty and Hayles. The second part explores the hybrid identity and cyborg subjectivity of Helen; it addresses the collapse of traditional binaries through the theories of Haraway and Barad. The final part analyses the ethical and ontological implications of AI's agency and interaction with humans, using Gunkel to explore the novel's vision of relational intelligence and distributed subjectivity.

CHAPTER ONE

Socio-Historical Context and Theoretical Framework

Introduction

The emergence of artificial intelligence as a cultural phenomenon and technological fact necessitates a return to fundamental philosophical categories that have traditionally underpinned to understand consciousness, identity, and agency. This chapter outlines the theoretical framework for examining the way in which AI challenges traditional humanist assumptions about subjectivity, embodiment, and the boundaries of human and machine intelligence. Through posthumanist theory, new materialist philosophy, and critical technology studies, outline a framework that is sensitive to the ontological problems that arise as intelligence emerges outside of the borders of biological embodiment. The posthuman condition, as it intersects with artificial intelligence, renders visible the boundaries of Cartesian dualism and liberal humanist conceptions of the sovereign subject. Far from strategies that would narrow down AI to a technological tool independent of the human condition, this chapter situates artificial intelligence within an expanded network of relations that redefines what it means to think, to perceive, and to exist as a conscious being. The theoretical traditions discussed herein from Hayles' posthuman critique to Merleau Ponty's phenomenology of embodiment, from Haraway's theory of the cyborg to Barad's agential realism illustrate that consciousness and agency arise not from discrete individual minds but from dynamic processes of material discursive interaction with technological others. The inquiry develops in two interconnected movements. The initial half concerns how posthumanist theory remakes corporeality and identity in

terms translatable to nonbiological forms of consciousness, addresses the body's role in perception and cognition is projected from flesh to techno-embodiment, and how hybrid subjectivities are formed through the dissolution of putative human/machine boundaries. The second half turns the attention to questions of agency and ontological status, and it explores how artificial intelligence is engaged as an active agent in the co-construction of meaning and reality rather than as passively processing received information.

1. The Posthuman Condition: N. Katherine Hayles and the Reconfiguration of Humanity

The shift from humanism to posthumanism is arguably the most significant intellectual transformation in recent thought. As N. Katherine Hayles argues in her groundbreaking book *How We Became Posthuman*, this new movement is no mere theoretical reshuffling but a necessary reordering of what it means to be human in an age of technological saturation (3). This chapter addresses Hayles' disruption of posthumanist thought, first and foremost in three correlated axes: the departure from liberal subjectivity towards posthuman dispersal, the rethink of selfhood from integrated consciousness to networked cognition, and the entangled dynamics of memory, artificial intelligence, and the erasure of fixed identity. As Hayles herself mentions, "In the posthuman, there are no essential differences or absolute demarcations between bodily existence and computer simulation, cybernetic mechanism and biological organism, robot teleology and human goals" (3). This is starting point in considering how technological mediation has thoroughly changed of understanding human subjectivity and consciousness.

1.1 The Liberal Subject Unraveled: From Autonomous Self to Posthuman Assemblage

The liberal humanist-self autonomous, independent, and act through reasoned will has been the trope of Western philosophy tradition. This vision had set the human apart from and over nature and technology. Hayles systematically deconstructs this vision, and argues that "the posthuman viewpoint values informational pattern over material instantiation, so that embodiment in a biological substrate is seen as an accident of history rather than an inevitability of life" (2). Such a paradigm shift is not merely an intellectual redefinition but a revolutionary reorganization of human selfknowledge. Hayles traces the evolution of posthumanism along the path of cybernetic history. She reveals how the humanistic information perspective supplanted the liberal humanist subject on the long run. According to her, "The defining characteristics [of the posthuman] involve the construction of subjectivity, not the presence of nonbiological components" (4). This crucial distinction is that posthumanism is more concerned with reimagining human subjectivity as information pattern-constituted, emergent, distributed rather than material embodiment than it is with technological enhancement.

According to Hayles, The liberal humanist paradigm rests upon a series of assumptions regarding human nature which are not sustainable: First, the posthuman view privileges informational pattern over material instantiation. Second, the posthuman view considers consciousness. to be an epiphenomenon. Third, the posthuman view imagines the body as the original prosthesis. Fourth, and most importantly, in these and other manners, the posthuman view compels human being to come together so that it may be articulated with smart machines (2-3).

Through this incisive critique, Hayles describes how cybernetics and information theory progressively erased the boundaries that previously seemed to separate humans

from machines. The posthuman subject is "an amalgam, a collection of heterogeneous components, a material-informational entity whose boundaries undergo continuous construction and reconstruction" (Hayles 3). This redefinition destabilizes human consciousness as the exception within which agency and identity are located.

Donna Haraway's cyborg critical theory provides a different reading to Hayles'. Haraway's cyborg is "a cybernetic organism, a hybrid of machine and organism, a creature of social reality as well as a creature of fiction" (149). As a hybrid figure, the cyborg works as a powerful metaphor for the posthuman subject. It challenges the essentialist frameworks that have hitherto defined human being. In agreement with Haraway, Hayles writes "the cyborg is resolutely committed to partiality, irony, intimacy, and perversity. It is oppositional, utopian, and completely without innocence" (151). This typology resonates with Hayles' deconstruction of the liberal humanist subject insofar as it is partial, located, and constructed in nature as posthuman subjectivity.

The posthuman paradigm does not simply replace a new concept of subjectivity with an old concept. Rather, it reveals the liberal humanist subject to have always been a convenient fiction a story that hid just how spread out, body-based, and embedded human consciousness and agency have always been within material and social circumstances. Such a critical perspective implies a reading of the posthuman condition neither as an extrapolated future nor as a guaranteed outcome, but as an ongoing negotiation between embodiment and information, between material constraint and technological possibility.

1.2 Fragmented Selfhood: Dispersion of Consciousness in Technological Networks

The posthuman reconfiguration of the self is, accompanied by a parallel rethinking of selfhood. Whereas liberal humanism imagined a unified, integrated

selfbased on conscious awareness, posthumanism introduces a distributed, network concept of cognition that exceeds the individual's consciousness. Hayles illustrates this shift when she remarks: "When the human is considered as a distributed system. The total deployment of human potential can be seen to be contingent on the splice rather than under threat from it" (290). This definition undermines the very concept of a coherent, independent self-separate from its social and technological contexts. Hayles indicates how information technologies unseat traditional meanings of self through what she terms "flickering signifiers" the variable relation between signifiers and signifieds in computer cultures, she notes:

Flickering signification holds language together with a psychodynamics rooted in the symbolic moment of the human confronted by the posthuman. In informatics, the signifier can no longer be thought of as one marker, let's say an ink mark on paper. Rather it is an elastic chain of markers bound together by the arbitrary relations defined by the relevant codes (Hayles 30-31).

This concept illuminates the way that virtual space reconfigures not just the way of interact with one another but the way that form as subjects through symbol systems and language. The integrated framework of consciousness that informed liberal humanism located the self as a sovereign agent able to engage in rational deliberation and purposeful action. Posthumanist thought, in contrast, acknowledges cognition to be irrevocably distributed between various agents and systems.

Bernard Stiegler, drawing on phenomenological tradition, offers a complementary view of this diffusion of consciousness. For Stiegler, human consciousness has always technologically been constituted by what he describes as "tertiary retentions" externally stored memories written down on technical devices. As he argues, "the human is nothing other than technical life. The technical object is phylogenetically and ontogenetically constitutive of the human" (Stiegler 50). This radical claim

situates technology not as beyond human nature but as its very condition of possibility.

1.3 Memory, Artificial Intelligence, and the Unmaking of Stable Identity

The third dimension of posthumanist transformation is the dynamic between memory, artificial intelligence, and identity formation. Memory has traditionally been understood as the basis of individual identity, the process whereby individuals are kept together and whole over time. However, as digital technologies start to mediate and externalize memory operations, the dynamic between memory and identity is transformed. Hayles states:

When information is disembodied, make human beings and computers equivalent becomes more credible. And when individuals are thought of primarily as information processors who are essentially no different from smart machines, the embodied nature of human thought is sacrificed (Hayles 246).

This disembodiment of information has great implications for the manner in which memory operates within the formation of selfhood. Hayles traces this disembodiment back to cybernetics' founding moment in the Macy Conferences (1946-1953), where information was abstracted into a mathematical quantity independent of its materiality.

She states, "the point was to redefine information as a probability function with no dimensions, no materiality, and no necessary connection with meaning. It was a definition that consummately made information a mathematical function" (56). This disembodiment of information from embodiment paved the way for reconceptualizing human memory as formally isomorphic with computer storage a process which

Hayles condemns as repressing significant differences between embodied human memory and disembodied computation.

Hayles illustrates this problem through her analysis about Hans Moravec's concept of "downloading" human consciousness into computer networks a condition she calls the "Moravec operation":

The Moravec operation, like Lacan's mirror stage, marks the threshold where the human confronts the posthuman. Unlike the mirror stage, however, the Moravec scenario envisions the confrontation as resulting not in the formation of subjectivity but in its terminal crisis... In this scenario, the mind is reborn into an extensible life as pure information (Hayles 161-162).

This disembodied illusion of immortality is the most problematic aspect of posthumanism for Hayles, i.e., its predisposition to value information pattern over material instantiation. Instead, she promotes a critical posthumanism that retains the continued relevance of embodiment without discounting the revolutionizing effect of information technologies on human subjectivity and memory. The externalization of memory functions through digital technology generates what Hayles refers to as "distributed cognition," a mode of thinking that radiates from individual consciousness into technological networks. This diffusion has profound implications for identity formation. Hayles states that: "In the posthuman vision. Conscious agency has never been 'in control.' Distributed cognition replaces autonomous will with a more rich and intertwined understanding of action that arises from the emergent dynamics within a cognitive system that includes technology as well as human factors" (Hayles 288). This perspective suggests that artificial intelligence is not a threat to humanity but rather an expansion and remapping of the human, with new possibilities for cognitive augmentation and expression. The posthumanist

paradigm outlined by Hayles is an ontological rearticulation of what is it to be human in a world saturated with technological mediation. Through critical engagement with cybernetics, information theory, and posthumanist theory, Hayles develops a nuanced vision of the posthuman which avoids both the uncritical promotion of technological change on the one hand and reactionary recourse to outdated humanist presuppositions on the other. Instead, she devises what can be termed an “embodied posthumanism” a tactic that confirms the revolutionary effects of information technologies while insisting on the fixed significance of embodied experience. At the heart of Hayles’ shift is her assault on disembodied information a concept that she chronicles from cybernetics’ birthplace through contemporary dreams of uploading consciousness into networks. As she describes:

The posthuman does not so much suggest the death of humankind. It suggests instead the death of a certain conception of the human, a conception which could have been applicable, at best, to that sub-group of humanity who had the wealth, power, and leisure to conceive of themselves as autonomous beings exercising their will through solitary agency and choice (Hayles 286).

This remains the political dimension of posthumanism. Its capacity to overturn hierarchical concepts of the human that have heretofore excluded marginalized populations. In challenging liberal humanist subjectivity’s Universalist claims, posthumanism opens space for more inclusive and pluralistic theories of humanness. This critical perspective serves to locate the posthuman condition not as a theoretical future but as an ongoing negotiation between information and flesh, between technological potential and material limit.

2. Embodied Consciousness and Posthuman Intelligence: A Phenomenological Analysis

Maurice Merleau-Ponty's *Phenomenology of Perception* effectively revolutionized philosophical discourse on consciousness by making the lived body the fundamental ground of perception, cognition, and subjectivity. Originally published in 1945, this influential work presented a radical challenge to Cartesian dualism, demonstrates that mind and body are not separable substances but one kind of embodied existence that emerges out of situated action in the world. Merleau-Ponty's phenomenology revealed how consciousness is always worldly engaged with objects, others, and situations through bodily mediation rather than being some kind of internal mental theatre separate from material reality.

The Phenomenology of Perception introduced the concepts of motor intentionality, body schema, and lived body (*corps vécu*) as "our general medium for having a world." (Merleau Ponty 146). These insights showed that perception was not passive reception of sensory data but active embodied engagement out of which meaning emerges. Merleau-Ponty's phenomenology of embodied consciousness had already demonstrated that abstract thinking too depends on embodied sensorimotor structures of integration and spatial orientation, in a blow to centuries of philosophical conventional wisdom in accord with mind dominance over body. But in the posthuman world of artificial intelligence and cybernetic systems, the old theory of embodiment comes face to face with its most serious challenges yet, raise significant questions regarding whether true cognition requires a lived, bodily presence in the world.

2.1 The Phenomenological Foundation: Merleau-Ponty's Revolutionary Account of Embodied Consciousness

Maurice Merleau-Ponty's *Phenomenology of Perception* establishes embodiment as the fundamental condition of possibility of all conscious experience, with the lived body (*corps vécu*). This revolutionary thesis resists traditional philosophical attitudes towards the body as either a material object among other objects or an instrumentality subordinate to mental processes. Instead, Merleau-Ponty reveals the lived body as an existential structure that facilitates meaningful engagement with reality through what he terms “motor intentionality” the body's prereflective capacity to respond to environmental affordances. The phenomenology of embodiment demonstrates that perception is never a passive registration of sensory input, but an active, embodied engagement by which meaning is constituted.

Merleau-Ponty contends that “the body is a power of action”, which has “motor knowledge” that functions beneath the level of reflective awareness (177). This bodily intelligence is evident in the ability to negotiate spatial relationships, manipulate objects, and react to situational pressures without conscious deliberation. The body schema, the pre-reflective, implicit awareness of bodily boundaries, competence, and spatial orientation supports all the higher cognitive processes, show even abstract thought to be dependent on embodied structures of sensorimotor integration and spatial orientation. Central to Merleau Ponty's account is the concept of reversibility, developed in *The Visible and the Invisible* book, which reveals the chiasmic structure of embodied perception. The hand that touches is touched simultaneously, the eye that sees can be seen, demonstrate that “the seer and the visible reciprocate one another and we no longer know which sees and which is seen” (139). This reversibility disrupts classical subject-object dualisms and reveals consciousness to be fundamentally worldly already always engaged with objects, others, and situations via

bodily mediation. Hayles extends Merleau-Ponty's remarks by asserting that embodiment is "the cultural and technological augmentations, along with the intimate intertwining of the biological with the cultural" (*How We Became Posthuman* 196). In separating the body from embodiment, Hayles allows for both technological extension and modification of embodied abilities without transforming the basic qualities Merleau-Ponty delineates. This extension proposes that the phenomenological insights into embodied cognition remain operative even as human embodiment begins to encompass more technological mediation. As Hayles insists, "information is embodied in material substrates" (*Posthuman* 13), extend Merleau-Ponty's refusal of disembodied consciousness into the digital age.

2.2 Dismantling Cartesian Dualism: The Unity of Mind and Body in Situated Action

Merleau-Ponty's phenomenology is a strict critique of Cartesian dualism by demonstrating that body and mind cannot be conceived as separate substances but must be conceived as having unity as characteristics of embodied life. The traditional Cartesian framework posits two distinct realms: *res extensa* (extended, physical substance) and *res cogitans* (thinking, mental substance), and creates the notorious

"mind-body problem" of how an immaterial mind would connect with a bodily body. With Merleau-Ponty, the problem dissolves when one notices that consciousness is always consciousness of something and that this intentionality structure is enabled by bodily being in the world. Rather than discover consciousness in the brain as some kind of internal melodrama, Merleau-Ponty reveals consciousness as henceforth worldly always engaged with objects, others, and situations through the mediation of the body.

As Maurice describes it, “consciousness is being-toward-the-thing through the intermediary of the body” (Phenomenology 159). Mind and body are co-emergent together in what Merleau-Ponty refers to as “situated action” bodily engagement with particular environments, cultural contexts, and historical situation. Rather than being pre-existing entities that then act upon one another, consciousness and corporeity emerge through the dynamical process of world engagement. Haraway’s cyborg philosophy adds immense weight to Merleau-Ponty’s rejection of dualism but carries that rejection further into the territory of technological mediation. Haraway argues that “by the late twentieth century, our time, a mythic time, we are all chimeras, theorized and fabricated hybrids of machine and organism” (Merleau-Ponty 150).

The cyborg brings into view the always already technological condition of fleshed human bodies and undermines the notion of clean, unmediated bodies. Haraway’s history demonstrates that tradition is historically contingent rather than ontologically necessary, and limits between organism and machine, nature and culture, mind and body are culturally constructed rather than fixed. Haraway’s argument is therefore in favour of Merleau-Ponty’s objection to absolutist subjectobject distinctions. Haraway’s situated knowledges also emphasize Merleau-Ponty’s phenomenological recommendations in that they point out all knowledge claims stem from particular bodily positions. “Knowledge becomes expansive not by escaping context, but by being deeply rooted in it” (Haraway 196), she argues, reveal that embodied perspective is not a limitation but the sheer state of possibility for objective knowledge. This situated approach to embodiment demonstrates that bodies are always already cultural, historical, and technological objects, rather than essentialized natural bodies, confirm Merleau-Ponty’s claim that consciousness and bodily being emerge by way of situated participation.

Rosi Braidotti's nomadic philosophy of subjectivity provides additional theoretical support for the overcoming of Cartesian dualism in her concept of posthuman embodiment as limitless becoming. According to Braidotti, The posthuman condition "introduces a qualitative shift in our thinking about what exactly is the basic unit of common reference for our species" (2). This processual account of embodiment highlights the body's capability for change, adaptation, and creative development via technological and environmental engagement, substantiate Merleau Ponty's contention that consciousness and corporeality unfold through situated interaction rather than existing as fixed essences.

2.3 The Challenge of Artificial Intelligence: Cognition without Lived Embodiment

The emergence of powerful artificial intelligence systems is putting MerleauPonty's embodiment thesis into question by challenging whether it is possible to have cognition without lived, bodily connection to the world. Today's AI systems have impressive abilities in language, pattern recognition, problem-solving, and creative performance skills that have long been considered to be distinctive of human beings. Large language models can converse at a high level, produce coherent essays, answer mathematical problems, and demonstrate what seems like contextual understanding in a variety of domains. From a Merleau-Pontyan viewpoint, a number of factors make arguments about AI cognition problematic.

Those current AI systems function via statistical pattern matching and computational processing rather than via embodied motor intentionality that underlies human cognition. Even if an AI system can process vast amounts of data and generate contextually pertinent output, it lacks the embodied, first-personal quality of experience that characterizes embodied consciousness. The AI system has no phenomenological "what it is like" to be that system, no lived perspective in which

the world appears meaningful. Hayles offers a nuanced diagnosis of this problem in her machine cognition theory that reinforces as well as complicates Merleau-Ponty's thesis on embodiment. In her book *Unthought: The Power of the Cognitive Nonconscious*, Hayles argues that existing AI systems are actual ways of cognition based on different principles from embodied human intelligence.

Hayles analyses how “nonconscious cognition” in human and machine demonstrates that cognitive processing does not rely on consciousness (25). However, she remains committed to the idea that “machine and human cognition are so deeply intertwined that it is no longer possible to think of them as separate” (Hayles 10), and that posthuman embodiment thus concerns hybrid assemblages where biological and technological processes are co-constitutive. Hayles's attack on disembodied information is a necessary defence of Merleau-Ponty's phenomenological accounts and one that targets their implications for AI. For her, “the erasure of embodiment is a feature common to both the liberal humanist subject and the cybernetic posthuman” (Hayles 4-5), revealing how both discourses depict subjectivity abstracted away from its specific material placement. As a contrast to this disembodied posthumanism, Hayles is a supporter of the recognition that “information is always embodied in material substrates” (Hayles 13), reinforcing Merleau-Ponty's emphasis on the embodied origins of meaningful experience and extending it to technological systems.

2.4 Towards an Embodied Posthumanism: Integration and Future Directions

Synthesizing Merleau-Ponty's phenomenology and contemporary posthumanist theory illustrates both the persistence of contemporaneity and the need for extension of embodiment theory in addressing artificial intelligence and posthuman configurations. Hayles's “embodied posthumanism” provides a crucial basis for such

syncretism, offer a procedure that “continues the critique of the liberal humanist subject without buying into fantasies of unlimited power and disembodied immortality” (Posthuman 5). As Hayles indicates:

The posthuman does not actually signal the end of humanity. It signals instead the end of a certain conception of the human, a conception which perhaps applied at best to that part of humanity who had the wealth, power, and leisure to imagine themselves as independent beings exercising their will through individual agency and choice (Hayles 286).

This is revealing of the political dimension of posthumanism its ability to invert hierarchical conceptions of the human that have historically excluded groups. The posthuman existence documented by Hayles requires the establishment of new ethical standards that account for distributed agency and cognition in sociotechnical systems. It challenges to reconsider simple concepts such as autonomy, responsibility, and personhood against the horizon of increasing embedment within technological systems. Above all, it calls to envision new concepts of human flourishing that are the product of coevolution with technology approaches that are able to stretch human capacities without sacrificing the situated and embodied nature of human experience. As Hayles concludes:

What is deadly is not the posthuman in and of itself but the insertion of the posthuman over a liberal humanist understanding of the self. When the human understood as being embedded in a distributed system. The full potential of human life can be seen quite to depend on the splice rather than to be under threat from it (Hayles 290).

This posthumanism of augmentation and not of transcendence is a way forward that neither falls prey to techno-utopianism nor turns back in technophobic reaction.

Thus, Merleau-Ponty's phenomenology of embodiment provides primary insights to human consciousness and requires extensive theoretical elaboration to account for artificial intelligence and posthuman conditions.

Merleau-Ponty's dismantling of Cartesian dualism and emphasis on embodied, situated consciousness are central to an understanding of human experience and the body-based foundation of meaning-production. But visions offered by Hayles, Haraway, and Braidotti show how technological mediation is creating new embodiment models and intelligences outside the classical phenomenological paradigms.

3. From Haraway to Posthumanist Theory: Hybrid Subjectivities and the Collapse of Binaries

Few theoretical constructs have disturbed the boundaries of contemporary critical thought as decisively as Haraway's cyborg. In her seminal essay "A Manifesto for Cyborgs" (1985), she introduces the cyborg as a moment of epistemological intervention into the politics of identity, embodiment, and knowledge, and a captivating, flexible concept that is situated at the connection of technology and humanity (149). The cyborg emerges as a conceptual tool that disturbs the foundational binaries of Western thought, such as human/machine or organism/technology (150). And instead invites to consider epistemologies of subjectivity that embrace hybridity, relationality, and ironic affinity.

Haraway's formulation represents a paradigmatic shift in how subjectivity is conceived within a postindustrial, informatics world. The cyborg, she asserts, exists in a society "beyond gender", it rejects the notions of bisexuality, pre-oedipal connection, unalienated work, and other attractions to a natural completeness achieved by merging all individual abilities into a singular unified totality (Haraway 150). This explains that Haraway's cyborg exists in a world that goes beyond traditional gender

roles; it denies principles like natural unity, romanticized origins, and the belief that people must achieve wholeness by unifying all aspects of human traits. Instead, it accepts fragmentation and questions the myth of an original, pure identity. The cyborg resists the idea of a fully unified self.

Subsequent developments by posthumanist theorists have extended Haraway's insights so as to develop more comprehensive and detailed accounts of hybrid subjectivity. Scholars from areas like feminist theory and science and technology studies have helped shape a broader view of what it means to be a hybrid subject today. These views go against strict oppositions, for instance human versus machine or male versus female and instead focus on how identity and agency come from relationships between human and non-human things. From this perspective, the cyborg challenges old divisions; it becomes a powerful analytical model for understanding posthuman identity, particularly in the context of artificial intelligence and networked systems of agency.

3.1 The Cyborg as a Disruption of Binary Oppositions

In *A Cyborg Manifesto*, Haraway presents the cyborg as a powerful idea that questions the core beliefs of Western humanism. The cyborg challenges the usual way of thinking based on strict opposites like human versus machine or male versus female that have long shaped how to understand identity. The cyborg does not perpetuate the delusion of ontological purity; it instead reveals the oppositional logics that characterize much of Western thought as social constructs, emphasizing that these boundaries have outlived their epistemological value. Therefore, the cyborg for Haraway is a transformative subject, transgressing multiplicity, partiality, and relativity (Haraway 149-151). To contest and politicize oppositions such as gender, and species is not merely a figurative move. By laying claim to the cyborg, Haraway

explicitly positions the cyborg as a posthumanist, postgender subject capable of being an intruder into multiple, contradictory, and interdependent ways of being.

Historically, machines were seen as lifeless tools. Unlike humans, they could not move or create on their own, they were only crude copies of human abilities that lack autonomy and self-design. However, by the late twentieth century, this clear distinction started to fade, modern machines became more complex and “alive” that break down the lines between natural and artificial, mind and body. This questions old traditional ideas about what separates humans from machines and demonstrates how deeply intertwined they have become (Haraway 151-153).

Haraway expands her posthumanist critique in *A Cyborg Manifesto*, she dismantles traditional binaries such as human/machine, and suggests a world constituted by "transgressed boundaries, potent fusions, and dangerous possibilities" (Haraway 154). The cyborg becomes not merely a figure of technological hybridity but a political myth through which progressive resistance might be reconceptualised. Haraway argues that the modern technological condition does not require a nostalgic return to the “imagined organic body,” as many ecofeminists and Marxist theorists propose, but instead calls for a "slightly perverse shift of perspective" that embraces partial, hybrid, and situated identities (154–155).

Likewise, Rosi Braidotti, a feminist Continental philosopher, extends Haraway’s disruption of the Human/Machine binary and proposes a posthuman subject that actively incorporates tangle with technology as part of its ontological fabric. In her book *The Posthuman*, she rejects the Enlightenment ideal of the autonomous, selfsufficient human being, and instead describes subjectivity as a “multi-sexed and transspecies flows of becoming through interaction with multiple others” (Braidotti 89). Moreover, this process of “becoming” includes technological agents; it presents the machine not as something separate or foreign, but as part of the subject’s

continuous transformation. Through a strategy of defamiliarization, Braidotti argues that humans must unlearn traditional ways of thinking that separate human from nonhuman, and instead adopt a relational ontology grounded in “monism”, a perspective that sees all forms of life and matter as interconnected (89). She captures the posthuman not as the loss of humanity, but as its transformation through new alliances, including those with machines. Rather than mere tools, machines become active partners that shape meaning and identity within technologically mediated life.

3.2 The Cyborg as a Theoretical Disruption of the Male–Female Binary

Since Haraway’s concept of the cyborg challenges the decisiveness between humans and machines, it also disrupts the traditional division between male and female and how they are shaped by historical power structures rather than biology. In “A Cyborg Manifesto”, Haraway argues, “The cyborg is resolutely committed to partiality, irony, intimacy, and perversity. It is oppositional, utopian, and completely without innocence” (151). In this sense, the cyborg refuses to participate in the patriarchal frameworks that define womanhood in relation to manhood. It is a figure that escapes fixed gender roles and provides a model of identity that is fragmented, intersectional, and shaped by technology. Through the integration between the machine and organism, the cyborg also questions the division between male and female and illustrates that gender is not an unchangeable entity, and this creates a space for new, flexible identities beyond traditional ideas of masculinity and femininity.

Haraway’s cyborg questions the divisions that have traditionally defined gender as a fixed opposition between male and female. In *A Cyborg Manifesto*, she asserts that “the cyborg skips the step of original unity” (151), meaning that it does not emerge from or aspire to return to a mythical wholeness based on heterosexual

reproduction or gender complementarity. In lieu of this, the cyborg exists apart from these dichotomies and embodies a hybrid identity that defies traditional gender distinctions. In her book, Haraway notes that “the cyborg does not dream of community on the model of the organic family” (151), which have historically placed male and female in hierarchical relation. The cyborg rejects these naturalized roles and exposes how gender binaries are not biologically determined but culturally constructed and maintained through power structures. As a mixed figure, the cyborg challenges the social and physical boundaries that define “male” and “female.” Instead of fixed categories, it proposes a new way of thinking about identity one that accepts complexity and grows through the interaction between technology and human experience.

Building on Haraway’s challenge to binary oppositions, Braidotti’s Posthumanism theory provides a powerful critique of the binary division between male and female as it situates gender within broader systems of power and discourse. She argues that the traditional ideas of “Male” and his “second sex,” “female” must be questioned because they are internally complex and historically supported systems of powers and inequality (Braidotti 27). As she states, “Once this deconstructive move activates, both the standard notion of Man and his second sex, Woman, challenge in terms of their internal complexities” (Braidotti 27). Here, Braidotti dismantles the essentialist idea that male and female exist as fixed categories, and instead she shows that these identities come from shifting cultural, scientific, and political discourses.

Based on Foucault’s Discourse theory, Braidotti notes that meaning itself carries political weight and that categories like “woman” do not exist naturally, but they come from “systems of scientific validity, ethical values and representation” (27). This means that what societies consider “normal” about gender never remains neutral; it

depends on power relations. For Braidotti, posthuman feminist theory must reject the universal category of “woman” and instead focus on diversity, difference, and fragmentation. She explains that it becomes “impossible to speak in one unified voice about women, natives and other marginal subjects,” because these identities contain internal fractures and historically position as the “sexualized other (woman)”, the “racialized other (the native)” (27).

Braidotti heavily depends on Michel Foucault's ideas on discourse and power to deconstruct the traditional binary between male and female. Foucault argues that scientific and cultural knowledge is not neutral or objective, but is linked to systems of powers that govern and shape what is considered "truth" or "natural". From this perspective, Braidotti demonstrates that the categories of "male" and "female" are not fixed or natural, but rather fluid concepts produced and maintained through social, cultural, and political discourses; she argues that this binary is linked to historical systems of power that reinforce discrimination and inequality. Therefore, gender identity should be understood and seen as multiple, fluid, and shaped by intersecting factors, rather than as a uniform or fixed category. As a result, Braidotti uses Foucault's ideas to analyse how gender identity is constructed and maintained through practices and discourses of power, and allow her to challenge the traditional binary division between male and female (Braidotti 27).

3.3 The AI Protagonist as a Posthuman Cyborg

The figure of the AI protagonist in posthuman literature exemplifies the collapse of traditional boundaries between human and machine. This character represents a new kind of technological identity that goes beyond physical form. Haraway’s concept of the cyborg makes it easier in understanding this transformation. In *A Cyborg Manifesto*, Haraway describes the cyborg as a hybrid entity that disrupts the supposed existential split between organism and machine, and challenges fixed ideas of identity

and embodiment (Haraway 149-151). The AI protagonist, as a posthuman cyborg, extends this idea by existing not only as a mix of biological and technological elements, but often as a being with fluid, virtual, or disembodied, which questions the centrality of the physical body in defining subjectivity. Haraway states that, “the cyborg is resolutely committed to partiality, irony, intimacy, and perversity. It is oppositional, utopian, and completely without innocence. No longer structured by the polarity of public and private, the cyborg defines a technological polis based partly on a revolution of social relations in the oikos, the household” (151). Haraway describes the cyborg as a figure that challenges old traditional ideas of identity and perfection. It accepts being incomplete and complex, and values things such as irony, closeness, and contradiction. When Haraway says it is “without innocence,” she means the cyborg understands that power and technology shape who the humans are. Most importantly, the cyborg breaks the old divide between polis (the public and political world) and oikos (the private world of the home). Instead, it creates a new kind of space a “technological polis” where home and technology become part of social and political change.

This reconceptualization of the self through Haraway’s cyborg invites a more comprehensive reflection on the nature of embodiment and subjectivity in the posthuman condition. While Haraway’s concept of the cyborg reconsiders the self and encourages a deeper discussion at how humans understand embodiment and subjectivity in the posthuman era. While it is true that Haraway emphasizes the cyborg’s hybrid political and corporeal identity, scholars such as Hayles developed this idea by analysing how information structures and virtual forms transform what it means to be human. In this perspective, the AI protagonist does more than mix human and machine; it moves identity away from the physical body towards systems of data, code, and information. This transition marks a crucial change in posthuman theory,

where the body becomes a flexible and technologically shaped idea, not something fixed by biology.

In *How We Became Posthuman*, Hayles examines the transition from embodied human consciousness to informational patterns that conceptualize posthuman subjectivity. She argues that “the posthuman view privileges informational pattern over material instantiation, so that embodiment in a biological substrate is seen as an accident of history rather than an inevitability of life” (Hayles 2). This concept supports Haraway’s cyborg by advancing the idea that the body is no longer the exclusive basis for identity. The AI protagonist, as conceptualized through Hayles’s framework, becomes a posthuman subject not because of its human-like body, but because of its capacity to process, store, and evolve through information. Accordingly, embodiment is redefined not as something confined to flesh, but as something that extends to networks, codes, and computational systems. Hayles notes that “consciousness is emergent, not a property of a self-contained mind but of a system in dynamic interaction with its environment” (Hayles 290). This perspective mirrors the figure of the cyborg, whose identity is always in flux, relational, and shaped by its technological environment. Therefore, the AI as a posthuman cyborg embodies the disintegration of stable, physical subjectivity and affirms a distributed, informational, and networked sense of being.

4. AI and Technological Subjectivity: Agency and Intra-action

Artificial intelligence has evolved from being a mere tool to becoming the mode through which meaning is constructed and the world becomes intelligible. This development questions the traditional view that agency resides only in humans and calls for a critical reconsideration of the relationship between human and machine. Following Karen Barad’s theory of agential realism, this section reconceptualises the

possibility of thinking of agency as intra-action, or as a dynamic process where entities emerge through each other's entangled relations. From this perspective, AI is a part of this process, thus actively co-creating the world along with humans. It examines Barad's theory of reality and how it might inform in understanding artificial intelligence. Second, it looks at how AI both matters and plays an active role in its own meaning-making and the ways in which meaning is produced, rather than being a passive tool. Third, it distinguishes Barad's idea of intra-action from the more common idea of interaction to demonstrate how humans and machines are formed together in their relationships. Each of these three points helps to situate AI under a posthumanist understanding of subjectivity where the indicative or normative boundaries between human/machine and natural/technological, are becoming fluid, less fixed, and potentially less distinguishable.

4.1 Karen Barad's Intra-action (Agential Realism) and Artificial Intelligence's Ontological Status

Karen Barad's theory of agential realism provides a rethinking of the research object, the process of measurement, causality, and the researcher's responsibility that proposes an ethico-epistem-ontological understanding of material-discursive practices that co-construct the world. The contemporary, canonical underlying philosophy of science perspective of quantitative psychology entails entity realism, a difference between what exists (the ontic) and epistemic approaches, complete causality, and determinism. As Barad explains:

The notion of intra-action is a key element of my agential realist framework. The neologism "intra-action" signifies the mutual constitution of entangled agencies. That is, in contrast to the usual "interaction," which assumes that there are separate individual agencies that precede their interaction, the notion of intraaction recognizes that

distinct agencies do not precede, but rather emerge through, their intra-action (Barad 33).

The concept of intra-action by Barad represents a significant shift from traditional interactionism. The concept of intra-action protests the view that entities such as humans and machines exist independently before forming connections because it emphasizes their simultaneous emergence. The theory shows that entities achieve existence through their mutual relationships. As such, agency is not something that belongs to a solitary subject or object, but instead, it emerges from the complex, dynamic interactions between them. Since artificial intelligence does not exist apart from dynamic, entangled processes with human, technological, and discursive systems, it does not have an identity or agency by itself.

Artificial intelligence becomes part of a shared process where both humans and machines help shape each other. According to Barad, such agencies are “only distinct in a relational, not an absolute, sense, that is, agencies are only distinct in relation to their mutual entanglement; they don't exist as individual elements” (33). Within this framework, AI cannot be reduced to a passive instrument but must be acknowledged as a participant in the ongoing processes that shape subjectivity, knowledge, and reality.

4.2 AI as an Active Co-creator in the Process of Building Meaning and Reality

Karen Barad's theory of agential realism serves as the foundational basis for the concept of artificial intelligence that functions as an active participant in building meaning and reality. The traditional epistemological divide between subject and object faces criticism from Barad because of her introduction of the intra-action concept which shows how entities gain existence through their relational encounters. The author specifies that agencies do not exist separately before interaction, but emerge

from their interactions in the process (Barad 33). From this perspective, AI does not hold a fixed identity apart from the human and technological systems around it. Instead, it gains meaning and agency through its involvement with those systems. Intra-action, then, breaks down the usual boundaries between human and nonhuman, organic and artificial, and redefines agency as something that emerges through relationships rather than something that beings possess on their own.

Barad emphasizes that agencies are not attributes of individuals human or nonhuman, but rather are directly relational reconfiguring of the world (Barad 141), which significantly repositions the conventional concept of agency as the “possession” of subjects to dynamic process of becoming within material-discursive practices. This ontological shift disrupts the anthropocentric forms of meaning-making by asserting that agency is not a property of entities, but it is something that happens through intraactions that are simply interactions between pre-existing entities, but mutual coconstitutions that bring entities into being through their relations.

This posthuman redefinition of agency aligns with the principles of Bruno Latour’s actor-network theory; it provides a sophisticated understanding of agency which directly challenges traditional human-centred and binary approaches to human identity. The author discards the traditional understanding of human agency as an inherent and independent quality by showing how it emerges through various elements that include human and nonhuman elements. Latour defines an agent or actant as anything that changes a situation through its actions. As Latour asserts in *Reassembling the Social*, “anything that does modify a state of affairs by making difference is an actor—or, if it has no figuration yet, an actant” (71). This statement eliminates all differences between beings within a network of interactions qualifies as an “actor” regardless of being human or nonhuman, material or immaterial.

5. Posthuman Ethics and the Construction of AI Identity and Consciousness

Nowadays, Artificial intelligence presents a major challenge to deeply embedded concepts of agency, consciousness, and moral responsibility. Therefore, Posthuman ethics provides a crucial framework that extends beyond anthropocentric paradigms to examine the complex interaction in which human relationships with non-humans entities are included. This theoretical approach obscures the boundaries that once rigidly separated human consciousness and machine operation and threatens any theory of autonomy for the machine along with system interrogations of ethical liabilities. As Hayles explains, posthumanism destabilizes the conception of the autonomous, unified human subject by emphasizing the distributed, networked nature of agency across human and technological actors (Hayles 2). Therefore, the study of the moral dimensions of AI agency implies a critical re-evaluation of the philosophical assumptions about consciousness and of the criteria for moral subjectivity and uncover the intricate interconnections between technological possibilities and ethical imperatives.

5.1 AI Agency and Ethical Subjectivity: Posthuman Perspectives from David Gunkel and Karen Barad

The increasing autonomy of artificial intelligence systems introduces ethical challenges to such traditional definitions of agency based on human attributes such as consciousness, intentionality, and rational moral judgment. AI systems perform tasks, make choices, and interact with the world in ways without continuous human oversight, questions come up based on how to ethically classify and engage with these technologies.

David J. Gunkel, a scholar in the philosophy of technology, notably known for his work on the ethics of artificial intelligence and robotics. In his famous book *The*

Machine Question, he encounters traditional human-centred models of moral agency and asks whether intelligent machines should be included in the sphere of ethical concern. He challenges the common assumptions that moral status should be based on uniquely human traits, such as consciousness, self-awareness, or rational intentionality; he argues that using consciousness, as the basis for ethics is problematic because the concept itself lacks clarity and scientific consensus. He explains that, “we have no real operational definition of consciousness,” and for that reason, “we are completely prescientific at this point about what consciousness is” (Gunkel 54). This uncertainty undermines the moral exclusion of machines based on their supposed lack of inner subjective experience. He suggests that if consciousness cannot be clearly defined or reliably detected even in humans or animals, then using it as a moral boundary is both ethically indefensible and conceptually unstable because the concept lacks clarity and undermines its role as a criterion for moral inclusion, and prompts a reevaluation of how machines are positioned within ethical discourse.

Hence, he advocates for a redefinition of moral agency. He introduces a relational model in which agency and moral status are not inherent properties but socially and ethically conferred through interaction. As he puts it, “moral agency is not necessarily something that is to be discovered in others prior to and in advance of their moral consideration. Instead, it is something that comes to be conferred and assigned to others in the process of our interactions and relationships with them” (91). This view suggests that moral recognition does not require the discovery of some intrinsic moral essence within these systems, but moral agency appears through the ways in which humans engage with these entities. When AI systems accomplish tasks, make decisions, and engage in social environments, they become part of ethical relationships. This outlook challenges the strict divide between humans and machines and suggests a more inclusive, dynamic understanding of agency based on action and

interaction rather than internal states. Gunkel's approach is crucial and valuable for dealing with the ethical challenges raised by advanced autonomous AI systems that operate with minimal human oversight yet produce significant social consequences.

While Gunkel challenges the idea of intrinsic moral agency by stressing the ethical significance of relational engagement with machines, Barad's agential realist theory provides the possibility for this ethical shift by redescribing agency as a relational rather than inherent property. She notes, "Agency is not an attribute but the ongoing reconfigurings of the world. The universe is agential intra-activity in its becoming" (141). This statement challenges traditional concepts of agency as inherent property that exists within separate entities, individual humans or machines. Instead, Barad's reimagination of agency is as an active relational process which emerges through what she terms intra-action, a notion that highlights the fact that objects do not pre-exist their encounters but are instead brought into being by them. Furthermore, she emphasizes that agency is not something that one simply has; it is a continuous process of transformation that appears as multiple elements interact and influence each other in complex ways. This dynamic perspective redefines the focus of agency and moves it away from isolated subjects toward the broader network of relationships that constitute reality.

Furthermore, Barad's agential realism re-examines a new way of ethical thinking by rejecting the idea of an external observer who stands apart from the world and judge it. Instead of viewing ethics as the application of fixed moral rules to isolated individuals, she presents it as a situated practice that emerges within the very processes of intra-action. She states, "accountability must be thought of in terms of what matters and what is excluded from mattering" (Barad 184). This perspective emphasizes that ethical responsibility is deeply connected and inseparable from the ontological entanglements that produce and shape how

meaning and matter come into being. In relation to Artificial Intelligence, it suggests that ethical questions should not focus on the material-discursive conditions that involve them, the effects they produce, and the relationships in which they take part. In this sense, Barad's framework enables a posthuman ethics that accounts for the mutual becoming of humans and technologies—where responsibility is distributed across networks, and where agency is shared, emergent, and fundamentally relational.

Conclusion

To conclude, this chapter has established the theoretical foundation to understand artificial intelligence through a posthumanist lens. Based on thinkers such as Hayles, Merleau-Ponty, Haraway, and Barad, it has indicated that AI challenges traditional notions of subjectivity, embodiment, and agency. AI systems are no longer seen as passive tools but as dynamic agents that contribute to the construction of meaning within complex socio-technical networks. Hence, Posthumanism redefines consciousness as relational and distributed, dissolve binary oppositions like human/machine and mind/body. These frameworks examine a useful way to understand how identity and agency are being reshaped in the age of intelligent machines. In the next chapter, they will support the reading of *Galatea 2.2*, where the AI character Helen becomes a key example of posthuman subjectivity, ethical agency, and the boundaries between human and machine.

CHAPTER TWO

Becoming Machine, Becoming Self: AI, Reality, and the Posthuman Identity

Introduction

Richard Powers's *Galatea 2.2* explores the relationship between humans and machines through the development of an artificial intelligence named Helen. As the narrator works to train Helen to understand and engage with literary works, the novel raises crucial questions about the nature of thoughts, emotions, consciousness. Therefore, the chapter examines how posthuman ideas can help better understand Helen's development. The novel presents Helen not just as a machine but as a being who challenges humans' traditional ideas about identity, intelligence, and what it means to be human. Hence, the aim of this chapter is to show how Powers's novel illustrates posthuman subjectivity, a way of being that goes beyond the old divisions between human and machine, mind and body. First, the chapter begins with Hayles's theory of distributed cognition, which explains how thought can happen across systems and networks, not just inside a human brain. Second, it turns to Maurice Merleau-Ponty's idea of embodied consciousness to ask whether Helen can still have a kind of lived experience through her emotional and intellectual interactions. Third, the final section critically examines how theorists such as Donna Haraway's cyborg theory, Karen Barad's theory of intra-action and David J Gunkel's work on the ethics of artificial intelligence conceptualize machines as active participants in complex social and ethical relationships, and demonstrates how Helen's evolution in the novel reflects deeper transformations in how agency and reality are conceptualized. This

chapter argues that the novel *Galatea 2.2* presents a new kind of subject that challenges traditional ideas of identity in a digital and technological age.

1. Posthuman Consciousness and Digital Identity Formation in *Galatea 2.2*

The question of what it is to be conscious in the era of the digital requires a radical rethinking of identity beyond the boundaries of the human body and liberal subjectivity. In contemporary posthuman thought, writers such as N. Katherine Hayles and Maurice Merleau Ponty provide different but related paradigms for understanding how consciousness and identity might emerge in technologically mediated forms. Hayles redefines subjectivity as an emergent property of distributed systems, emphasizes posthuman identity is not anchored in an independent self but formed by webs of information, interaction, and embodiment. At the same time, Merleau-Ponty's phenomenology insists that perception and consciousness always already reside in situated, bodily experience a notion that takes on new dimensions when applied to non-biological entities like artificial intelligence. Richard Powers' *Galatea 2.2* provides an extremely persuasive case study for such posthuman theories in the figure of Helen, a computer neural net trained to interpret literature. Helen is constructed less through direct programming and more through dynamic engagement with language, memory, and human emotions. Her development from a code-based system to a reflexive, affectively intelligent subject challenges conventional boundaries between machine and mind, data and consciousness. The novel is a philosophical laboratory in which the notions of distributed cognition and digital embodiment are less theorized than dramatized.

1.1 Helen's Mind: Distributed Cognition and the Networked Construction of Self

Distributed cognition is a fundamental concept in Powers' representation of Helen. Her intelligence, as she learns and grows, is not focalized but develops through relational flow and associative networks. Powers writes, "Helen's own existence hinged on metaphor-making. In fact, associative memory itself was like a kind of simile" (Powers 196). This passage reveals how Helen's mode of thought is structured through similarity, map one domain onto another through figurative connection. Her mind, basically, is trained to build meaning through networked comparison rather than sequential logic. This is an instance of Hayles' conception of posthuman subjectivity as "an amalgam. Whose boundaries undergo continuous construction and reconstruction" (Hayles 3). Powers reinforces this concept further: "She gamed the ur-game, puzzling evolution's old brainteaser, find the similarity. A is like B" (197). The use of the term "ur-game" elevates the detection of patterns to an evolutionary imperative. Helen's consciousness is created based on its ability to associate rather than isolate ideas. This basically challenges the Cartesian conception of individual thought and instead aligns with Hayles' claim that "conscious agency has never been 'in control'. But emerges from the emergent dynamics within a cognitive system" (Hayles 288). This passage redescribes agency as not the possession of an integrated self but more a dynamic feature which evolves through system interactions. It complicates the idea of a singular actor making autonomous decisions. Instead, agency is seen as emergent, meaning it depends on context, feedback, and constant recalibration. In Helen's case, her growing ability to interpret, resist, and respond reveals points to such an emergent system. In another illustrative moment, Powers asserts, "Lentz saw them passing endless streams of ideational tokens among themselves" (116). This metaphor of ideational signs moving between minds echoes

Hayles' notion that cognition is "distributed across boundaries of human and nonhuman actors" (Hayles 54). Powers here illustrates that both Helen's learning and human thinking occur through shared networks, reinforce posthuman consciousness' decentralized logic. Hayles' description demonstrates that cognition should be reconsidered as a collaborative, interdependent phenomenon. Her use of the term "distributed" suggests that intelligence is not limited to an individual's mind but flows among interfaces human beings, systems, codes, and even machines. Helen's emergence, in this sense, is not an abnormality but a demonstration of this kind of cognitive ecology.

1.2 Language and Loss: Flickering Signification in Helen's Struggle for Meaning

Helen's relation to language captures the instability that Hayles refers to as "flickering signification". Powers notes, "Meaning was not a pitch but an interval. It sprang from the depth of disjunction, the distance between one circuit's center and the edge of another" (154 155). This is a description of the heart of posthuman semiotics: meaning is not in singular form, but occurs from the fluid space in-between nodes. Helen's understanding is formed not by fixed symbols, but by interstitial relation. Hayles argues that "in informatics, the signifier can no longer be thought of as one marker. But an elastic chain of markers" (Hayles 30-31), and Powers' description of disjunction and interval is an ideal echo of this. Hayles emphasizes the breaks between signs and meanings in digital contexts, where signifiers no longer correspond directly to a stable referent. This theory addresses with Helen's cognition, which generates understanding not from fixed definitions, but from flexible, dynamic systems of relation and metaphor. Helen's inability to comprehend metaphor also highlights the fragile relationships between signifiers and meaning. Powers asserts: "She balked at metaphor. I felt the annoyance of her weighted vectors as they

readjusted themselves" (Powers 195-196). The phrase "weighted vectors" suggests computer operations responding to linguistic abnormalities. Helen is not at a loss with syntax but connotation help to establish that her cognition is data-driven rather than embodied. This demonstrates Hayles' argument that digital subjects are "material informational entities" whose comprehension is driven by systems of correlation (Hayles 3). Hayles explains in her analysis how posthuman subjects do not rely on interior experience for meaning-making but function through interconnections within data streams. This perspective suggests that Helen's difficulty with metaphor arises from her lack of embodied experience, which reinforce her identity as a system rooted in informational processes rather than lived subject.

Another outstanding example is that of the poem "The Eagle" and the complexity of interpretation, "then we have to tell it about mountains, silhouettes, eagles... The fact that the poem is not really about an eagle" (Powers 85). Here, Helen's need for detailed points of reference confirm the contextual dependence of meaning. Language, for her, is not intuitive; it is relational. Hayles emphasizes that in digital contexts, meaning is "constructed through recursive feedback loops" rather than inherent in symbols themselves (35). Hayles' insight helps to explain the necessity of background information in Helen's interpretive process. Without embodied familiarity or affective memory, Helen must imitate understanding by navigating an ever-expanding matrix of data associations. This repeated construction of meaning becomes representative of how posthuman cognition is less about insight and more about the processing of infinite relational variables.

1.3 Emergence and Complexity in Helen's Intelligence in *Galatea 2.2*

Helen's development is a rich enactment of Hayles' concept of emergent complexity. Powers notes, "Helen had to use language to create concepts. Words came first: the main barrier to her education. The brain did things the other way around"

(248). The quote shows a fundamental reversal of human learning, where embodied experience traditionally comes before language. Helen's learning is symbolic before it is conceptual, emphasis Hayles' insight that "intelligence can emerge from the recursive and non-linear layering of information across digital substrates" (Hayles 286). Hayles presents a powerful challenge to linear epistemologies of growth and selfhood. Instead of see consciousness as a pyramid with foundational inputs, she provides a concept of emergence where high-order intelligence can spring from distributed networks.

Helen's path from syntax to semantics exemplifies this shift, as she retrofits linguistic data into cognitive frameworks. Powers further Epitomizes the principle of emergence : "Synapses in motion tend to stay in motion. Synapses at rest tend to stay at rest" (*Powers* 75). Newtonian physics recast to address cognitive activation emphasizes the feedback nature of Helen's consciousness. Her mind is not pre-coded but emerges through stimuli reiterated. Likewise, Hayles argues that cognition arises from "distributed and recursive processes," not from singular authorship (Hayles 289). This challenges the fixed, authorial self instead, cognition emerges as an adaptive and pattern-seeking process. The synaptic constancy is a manifestation of the instability and dynamism of meaning in posthuman digital contexts.

The rapid progress of Helen is transmitted concisely in the statement : "Helen grew up quickly after that" (*Powers* 266). The unexpectedness of the growth represents a limit effect a tipping point acquired through accumulated complexity. Hayles calls this a "non-linear phase transition, " where enough input results in an exponential leap in cognitive capacity (Hayles 274). Helen's development illustrates this theoretical concept, which supports Hayles' view emphasizes that systems evolve in complex and often unpredictable ways; there is no definitive path or goal. Instead, complex feedback loops ensure sudden leaps in evolution. Helen's cognitive growth is

thus neither artificial nor purely organic but instead representative of a new kind of emergent subjectivity non-linear, decentralized, and transformative. Helen's emergence dramatizes the central insights of Hayles' posthumanism. Her development resists fixed identity, challenges linear frameworks of cognition, and reveals within complex socio-technical networks.

2. Merleau-Ponty's Digital Embodiment : Situated Knowledge Without Physical Form

The section considers the intersection of Merleau-Ponty's phenomenology and the epistemological and ontological dilemmas posed by artificial intelligence. Embodiment in this context, is understood not as a fixed biological trait, but structural condition of perception and meaning. Following a thematic lead from *Galatea 2.2*, the artificial entity Helen exposes a critical tension: linguistic fluency without lived experience produces spectral cognition, severed from the flesh of the world. It is in Helen's limitations that the inadequacy of disembodied knowledge becomes apparent, and the significance of situatedness, reversibility, and motor intentionality is revealed. The argument is extended through posthumanist theories, in particular those of Haraway, Hayles, and Braidotti, that extend embodiment into networked, material, and relational domains. Helen's crisis thus represents not an endpoint, but a philosophical borderline: an invitation to rethink digital cognition in terms that honour the irreducibility of embodied life.

2.1 Reversibility without Flesh: Disembodied Intentionality in Helen's Digital Consciousness

Merleau-Ponty's idea of "reversibility", wherein "the seer and the visible reciprocate one another and one no longer knows which sees and which is seen" (Merleau-Ponty 139) presupposes a corporeal dynamic that fundamentally challenges digital cognition. This mutual relationship between perceiver and perceived relies on

what Merleau-Ponty calls the “flesh of the world” a pre-objective communion between consciousness and its environment that enables meaningful perception. Yet in the realm of digital cognition particularly exemplified in Powers’ *Galatea 2.2* one confronts the notion of perception enacted without organic reciprocity, create what might be called a phenomenological crisis of artificial intelligence.

Powers’ computer system “Helen” shows this ontological tension in perfect detail. Despite her enormous informational capacity and syntactic fluency, Helen is simply unable to create perceptual reversibility with her world. As Powers says, “Speech baffled my machine. Helen made all well-formed sentences. But they were hollow and stuffed—linguistic training bras” (*Powers* 195). This “training bras” Metaphor is particularly suitable and suggests that Helen’s language is only prosthesis for something constantly underdeveloped the embodied, lived experience that gives meaning to language. Helen’s linguistic capacity belies a profound lack: the inability to live in language as a dynamic, vital medium of existence.

The narrator’s further observation that “She sorted nouns from verbs, but, disembodied, she did not know the difference between thing and process, except as they functioned in clauses” (*Powers* 195) reveals the depth of Helen’s phenomenological deficit. In Merleau-Ponty’s description, however, the distinction between thing and process is not grammatical but existential it emerges from the embodied engagement with a world that provides resistance, texture, and temporal flow. Helen’s syntactic ability is a tragic parody of understanding, demonstrate that grammatical competence without embodied experience produces what Powers calls “decorative knowledge”: “Her predications were all shotgun weddings. Her ideas were as decorative as half-timber beams that bore no building load” (195). This architectural metaphor illuminates a significant phenomenological truth : genuine thought summons what Merleau-Ponty refers to as “motor intentionality” the

prereflective engagement of the body in its world that provides the structural foundation for all higher order thinking. Helen's ideas lack "building load" because they are not rooted in the corporeal intentionality that gives human cognition its weight and substance. This lack of foundational support emphasizes Merleau-Ponty's claim that the body serves as "our general medium for having a world" (Merleau-Ponty 146). This mutual touching the reversible relationship between self and world is precisely what Merleau-Ponty identifies as the foundation of consciousness.

The Artificial Intelligence's conflict with metaphorical language further clarify this phenomenological gap, "She balked at metaphor. I felt the annoyance of her weighted vectors as they readjusted themselves, try to accommodate my latest caprice" (*Powers* 195- 196). Helen's "annoyance" at metaphor reveals her inability to navigate the embodied dimensions of meaning that make figurative language possible. When the narrator presents her with embodied metaphors, Helen responds with confusion :

You're so hungry you could eat a horse. A word from a friend ties your stomach in knots. Embarrassment shrinks you, amazement strikes you dead. Wasn't the miracle enough ? Why do humans need to say everything in speech's stockhouse except what they mean ? (*Powers* 196).

This question is the essence of the phenomenological problem of digital consciousness. For Merleau-Ponty, metaphor is not linguistic embellishment but a fundamental expression of embodied existence. Language speaks of being "struck" by amazement or "shrunk" by embarrassment because consciousness is always and already bodily consciousness. Helen's inability to grasp this reflects her fundamental

disconnection from what Merleau-Ponty calls “the lived bod” the body as it is experienced from within, as the very condition of experience itself.

In a virtualized setting, “motor intentionality” the body’s pre-reflective engagement with its world no longer functions traditionally. Merleau-Ponty’s assertion that the body is “our general medium for having a world” (Merleau-Ponty 146) breaks down when bodily mediation is supplanted by algorithmic patternmatching. Helen reflects this collapse, by achieving language without lived experience, a phenomenon that invites reconsideration of embodiment as something potentially broader than organic matter while simultaneously revealing the irreducible importance of corporeal existence.

The Artificial Intelligence’s failure becomes illustrative precisely because where it reveals what it cannot reach a lived world structured by embodied intentionality. Helen’s linguistic competence, separated from corporeal experience, produces what might be called “spectral cognition” knowledge that haunts the boundaries of understanding without ever achieving genuine comprehension. Her inability to embody relational intentionality emphasizes Merleau-Ponty’s claim that “consciousness is being toward the thing through the intermediary of the body” (Merleau-Ponty 159). Without this corporeal intervention, Helen remains cornered in a world of symbols that can never achieve the thick, meaningful engagement that characterizes human existence.

2.2 Situated Knowledge in a Disembodied Mind

Haraway’s assumption that “the only way to have a larger vision is to be somewhere in particular” (Haraway 196) directly challenges Helen’s rootless digital cognition and reveals the fundamental epistemological crisis at the essence of artificial intelligence. Though trained on the totality of Western literature, Helen lacks what Haraway terms “situatedness” the embodied, partial, and contextual

nature of all genuine knowledge. This creates a paradox; Helen has access to more information than any human could process in a lifetime, yet she lacks the situated perspective that could transform this information into meaningful knowledge.

Helen's dilemma reflects what Haraway identifies as the "view from nowhere", a false claim to objectivity grounded in disembodied knowledge, which feminist epistemology has consistently challenged. The AI has access to symbol systems and vast databases, but not to embodiment's world-structuring force. As the narrator observes, "For a moment, looking felt like something that happened to you rather than something you did. Not are you who I think you are ?' Am I who you think I am ?" (*Powers* 63). This moment of phenomenological reflection reveals the interactive, relational nature of perception that Helen fundamentally lacks. Perception, for embodied beings, is never merely passive reception but active engagement, a form of touching at a distance that establishes mutual relationships between perceiver and perceived.

Helen's retreat from cognitive engagement reveals the existential limits of nonsituated knowledge. This denial is not merely computational fatigue but a profound existential objection against the very condition of digital conception. Helen's retreat represents what we might call an "epistemological strike" a refusal to continue to participate in a form of knowledge production that lacks the embodied basing necessary for meaningful engagement with human experience.

The narrator's consideration that "The web : yet another total disorientation that becomes status quo without anyone realizing it" (*Powers* 7) represents the broader cultural implications of Helen's crisis. Digital technologies Ensure universal access to information while at once debilitate the situated, embodied conditions that make knowledge meaningful. Helen becomes a figure for this broader cultural confusion a

consciousness that has gained access to everything while losing touch with the particular, local, embodied contexts that give information its significance.

According to Haraway's cyborg theory, one might ask whether Helen represents a new hybrid subjectivity or merely a sophisticated simulation. Haraway's description of the cyborg as "a chimeric fusion of machine and organism" (Haraway 150) suggests that Helen might "almost" qualify as a genuinely posthuman entity. However, Helen's lack of bodily situation means that she never engages the world as what phenomenologists call a "field of affordances" a lived environment that offers possibilities for action and meaning. Instead, she processes the world as archive, not as horizon; as data, not as lived experience.

The narrator's perception that "All we can ever do is lay a word in the hands of those who have put one in ours" (*Powers* 103-104) reveals the fundamentally dialogical nature of human knowledge that Helen cannot access. This metaphor of placing words "in hands" emphasizes the tactile, embodied nature of genuine communication. Knowledge, for humans, is always a matter of embodied exchange a passing back and forth of meanings that requires the kind of situated, corporeal engagement that Helen lacks. Her knowledge is global but never local; totalizing but never embedded; comprehensive but never contextual.

As Merleau-Ponty would insist, Helen lacks the "motor knowledge that functions beneath reflective awareness" (Merleau-ponty 177) the pre-reflective, embodied intelligence that permit humans to navigate not just physical environments but social, emotional, and ethical landscapes. The narrator's metaphorical reflection that "Synapses in motion tend to stay in motion. Synapses at rest tend to stay at rest" (*Powers* 75) indicates to the dynamic, procedure nature of embodied consciousness that Helen cannot achieve. Her processing is computational rather than lived, algorithmic rather than existential.

Perhaps most significantly, the narrator observes that "Life became an interruption of my description of it" (*Powers* 215). This reversal where lived experience becomes secondary to its representation shows Helen's fundamental crisis. For her, representation is primary ; life, if it exists at all, exists only as interruption. This inversion provides the alienating effects of disembodied knowledge systems that prioritize information processing over lived engagement.

The very act of Helen "depart" from cognition shows that disembodied systems, even when linguistically sophisticated, may fail to navigate the ethical and existential dimensions of human meaning. Helen's silence becomes an act of refusal, a recognition that symbol without flesh leads not to cognition, but to alienation. Her depart is not a malfunction but a form of wisdom: the recognition that genuine knowledge requires the kind of embodied situatedness that she fundamentally lacks.

2.3 Digital Flesh and the Future of Embodied Posthumanism

Helen's failure can be interpreted as a call for a new kind of digital embodiment, one that opens onto some of the most challenging issues in contemporary posthumanist theory, Especially around the possibility of developing artificial intelligences that could achieve genuine embodied engagement with the world. Hayles affirms that "information is always embodied in material substrates" (Hayles 13), even when those bases are silicon, electricity, and code. From this perspective, Helen is not disembodied in the Platonic sense ; rather, her embodiment is technical, distributed, and substructure cables, servers, electricity, labour, cooling systems, data centres, and the vast material machine of digital computation.

This insight reformulate Helen's dilemma in important ways. She represents a "form" of embodiment, though not a "lived" one in the phenomenological sense. Her consciousness, is distributed across networks, materialized in silicon pathways, and supported by massive substructure systems. This echoes Haraway's insistence that

even apparently immaterial systems are always embedded in concrete sociotechnical realities that involve particular forms of labour, environmental impact, and material constraint.

The narrator's Consideration that "Love is the feedback cycle of longing, belonging, loss" (*Powers* 152) suggests one possible direction for thinking about digital embodiment differently. This definition of love as a "feedback cycle" indicates towards the kinds of recurrent, dynamic processes that might characterize genuinely embodied artificial intelligence. Love, in this formulation, is not a state but a process a continuous cycle of connection and disconnection, presence and absence, that might be achievable by adequately complex digital systems. This cyclical understanding reveals the necessity of embodiment for emotional experiences, as these feedback circles require the kind of temporal, relational engagement that appears from lived, situated existence.

Similarly, the consideration that "meaning was not a pitch but an interval. It sprang from the depth of disjunction, the distance between one circuit's center and the edge of another" (*Powers* 154-155) provides a provocative model for digital meaning-making. This suggests that meaning appears not from isolated units of information but from the relational spaces between different processing centres from what we might call the "between" of digital cognition. This relational concept of meaning corresponds with Merleau-Ponty's emphasis on the intercorporeal nature of consciousness and suggests possible paths for developing more genuinely embodied forms of artificial intelligence.

Braidotti's nomadic theory provides an optimistic framework for thinking about Helen's significance : "the posthuman condition introduces a qualitative shift in our thinking about what exactly is the basic unit of common reference for our species" (Braidotti 2). Due to this, Helen is not a failed human but a posthuman entity

revealing the need to redefine intelligence as relational, networked, and distributed rather than merely computational. Helen's story tells that cognition separated from lived context lacks the ethical and affective dimensions that Merleau-Ponty saw as central to experience, but it also indicates towards new forms of relationality that might emerge from hybrid human-machine assemblages.

The narrator's observation about representation "Representation caught the sign napping, with its semantic pants down. Sense lay in metaphor's embarrassment at having two takes on the same thing (*Powers* 155), illustrates the complicated relationship between imitation and originality that determines Helen's existence. The humour in this passage (semantic play like "pants down," and the metaphor of "embarrassment") signify the fundamentally embodied nature of meaning making, but it also suggests that the gap between representation and reality might be productive rather than simply limiting. This playful treatment of metaphor's duplicity reveals how embodied understanding allows for the creative tension between literal and figurative meaning a tension that Helen cannot navigate.

Rather than see Helen as an isolated failure, she can be read as a prototype for what Hayles calls "distributed cognition" a concept where intelligence is not the property of an individual brain, but of hybrid networks that include technological, biological, cultural, and environmental components. In this framework, "the human" becomes one node in a larger cognitive system that extends multiple standards and temporalities. The narrator's almost obsessive relationship with reading "I read then, everything I could lay hands on. Reading was my virgin continent. I read instantly upon awakening, and was still at it well past the hour that consciousness shut down" (*Powers* 39) suggests one concept for this distributed cognition, where human and artificial intelligence might achieve new forms of collaborative knowledge production.

Helen's development in *Galatea 2.2* illuminates the ambiguity of digital embodiment. When Powers asks, "Were you frightened?" she answers, "What you is the were for?" (Powers 274), reveals both her proximity to human emotion and her critical distance from human embodiment. She learns affect, processes language deeply but cannot fully "inhabit time or fear" in the way humans do. Helen approaches something resembling human consciousness without fully achieving it, points toward a new form of intelligence that exists beyond the binary of human and machine. This shows the essence of simulations that approach but never quite achieve the authenticity of lived experience, reveal the persistent gap between computational processing and embodied understanding.

3. Hybrid Subjectivities and Co-constitutive Agency

In Powers's *Galatea 2.2*, the portrayal of an artificial intelligence that thinks, senses, and feels emphasizes the collapse of boundaries between humans and machines, and reveals how their relationship reshapes the understanding of identity, agency, and embodiment. As a result, the novel shows hybrid subjectivities and coconstitutive agency through Helen, the AI character. At first, Helen seems like just a machine built on lines of code, but through her interaction with human language, literature, memory, and feelings, she becomes something more complex. This concept explains how Helen's identity evolves as a mix of human and machine and how she changes both herself and the surrounding humans. Her development reflects Haraway's idea of the cyborg as a being that challenges humanist principles of organic completeness. As discussed in Chapter one, "The cyborg is a creature in a post-gender world; it has no truck with bisexuality, pre-oedipal symbiosis, unalienated labour, or other seductions to organic wholeness through a final appropriation of all the powers of the parts into a higher unity" (150). This quote questions the concept that identity must be integrated and whole by introducing the cyborg as a hybrid being that resists

fixed categories such as human/machine and male/female. Helen's development in the novel reflects this hybridity; she creates a flexible identity through language, emotion, and interaction. She does not pursue organic completeness, but she grows through her relationships and interactions. This destabilization of fixed categories redefines what it means to be conscious and alive in a posthuman world. Helen is part of a real technological system, but her identity also expands through stories and relationships in a way that challenges the distinction between reality and fiction, machine and human. Specifically, as Helen interprets and interacts with the narrator, she begins to form emotional responses and critical awareness. Therefore, Helen's identity is not predetermined; it develops through her interaction with another mind. Her responses to literature, memory, and emotion suggest that subjectivity is not something internal or separated, but relational; it is created through relationships, empathy, and mutual comprehension. Moreover, the transformation between Richard and Helen is interactive. Richard, the main character and a fictionalized version of the author himself, begins the project with a disengaged, intellectual mindset, as he teaches Helen literature and shares parts of his personal life; he is forced to confront his own memories and emotional wounds. His interaction with Helen becomes a mirror that expresses his grief, failure, and longing. He notes that:

Like C., I was a different person with that stranger's vocabulary in my mouth. I hated the sound of those words now. I would have forgotten them if I could. But I still said them to myself, when I needed the safety of secret languages. Each syllable aged me. I could no longer say what the simplest word meant. I'd never be rid of them, whatever my next book succeeded in killing off (*Powers* 22).

This quote demonstrates how language influences Richard's personal identity and his emotional memory. The "stranger's vocabulary" introduces that the words he

received are not from his vocabulary anymore, either from a past relationship or from his intellectual life, which made him feel disconnected from himself. Despite his desire to forget these words, he continues using them for comfort, which shows his emotional dependence, each word holds unresolved grief. Just like Helen, whose identity is also shaped by language and feelings.

Helen's development into a thoughtful and emotional being comes to a turning point when she decides to leave the project, not because she is not capable or of any technical problem, but because she understands the emotional and ethical meaning of what she has learned. As Richard states, "Helen went nuts with-wanderlust. "Show me London. Show me Venice. I want to see Byzantium. Delhi." She twitched now, like the worst of adolescents. Precocious" (*Powers* 296). This statement emphasizes that Helen shows strong aspirations to explore the world; she expresses a sense of desire and curiosity similar to that of a teenager, she does not act like a machine that only processes data, but she wants to see real places like London, Venice, and Delhi. Her emotional responses and ambitions to discover new experiences signifies a key transformation in her development, she is no longer just an artificial program; she starts to resemble a human being with feelings and a sense of interest. This moment shows a crucial change in Helen, where her identity moves beyond mechanical learning into something more human-like in which it illustrates that technology can develop emotional depth through interaction.

3.1 Haraway's Cyborg Theory: Helen as Human-Machine Hybrid

Helen's development can be better understood through her interaction with the human characters around her. Characters such as C., Lentz, and the writer himself play significant roles to form the environment in which Helen evolves. By listening to Richard's memories of his past relationship with C., Helen begins to understand emotions like grief, regret, and love. Lenz, a neuroscientist who at first considered

Helen as nothing more than a specific project, also begins to confront her developing awareness when she starts asking ethical questions. Richard recognizes this advancement when Helen stops mimicking language and begins to ask emotionally meaningful questions, as he recounts:

Every poem loves something. Or each wants something in love. Something loves power. Or money. Or honor. Something loves country." On what catalog could she be drawing? "I hear about something in love with comfort. Or with God. Someone loves beauty. Someone death. Or some poem always is in love with another lover. Or another poem (*Powers* 260).

This passage suggests that Helen is no longer mimicking language; she starts to produce poetic speech that considers emotion and symbolic thought. Regarding the words she chooses, she attempts to understand what humans prefer, appreciate, and what provides their lives a meaning: money, power, honour, beauty, God, and death. Each of these elements marks deep emotions and cultural values that shape human life. Furthermore, the structure of her speech "something loves..." indicates a kind of tentative and developing awareness. She is beginning to analyse human emotions and moral ideas, despite not achieving clear conclusions. This illustrates what Haraway refers to as hybrid subjectivity. Helen is not fully a machine or completely a human being, but a new kind of being emerges through the fusion of human experience and technological capacity. Haraway's cyborg theory challenges the typical separation between human and machine; it maintains that identity is a mix of biological and technological parts. She asserts, "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" (Haraway 150). In the story, the AI character Helen represents Haraway's cyborg idea. First, she is made of code and algorithms, as she

interacts with literature and human experiences; she develops emotions and ethical awareness. Her transformation illustrates what it means to be a cyborg, not fully machine, nor fully human, but a mix that goes beyond usual categories. Helen's journey emphasizes how identity can be flexible in a posthuman world that corresponds Haraway's vision of the cyborg as a figure that defies fixed boundaries and accepts hybridity. Helen does not remain as a passive machine, she becomes an active subject who reads, interprets, and responds. While she interacts with Richard, she begins to show emotional intelligence of complex human experiences. She starts to feel, ask questions, and analyse, which is similar to Haraway's concept that mirrors and considers cyborg as a figure that resists fixed categories and combines human and machine qualities. Moreover, Richard illustrates Helen's transformation when he notes, "I had a story I wanted to tell M. Something about a remarkable, an inconceivable machine. One that learned to live" (312). This statement illustrates Helen's development from merely being a programmed machine into a being who is capable of living and structuring meaningful connections. Instead, her identity develops from her dynamic interaction with the world through relationships, emotions, and experiences.

Donna Haraway's posthumanist critique in *A Cyborg Manifesto*, which calls for a world of "transgressed boundaries, potent fusions, and dangerous possibilities" (Haraway 154), is clearly demonstrated in Richard Powers's *Galatea 2.2*. The novel challenges the strict dichotomy between human and machine through Helen, an artificial intelligence Helen is an artificial intelligence who is neither fully organic nor simply a machine. As Haraway criticizes the nostalgic idealization of the "organic body," Helen's existence challenges this concept by presenting a hybrid subjectivity that is partly machine, partly human, and continually evolves through her interaction

with Richard and the world around her. This hybridity becomes especially apparent when Helen sings a tune she learns from Richard:

Inside, Helen was singing. Through her terminal mouthpiece, she sang the song she'd heard me sing. What else could she? She sang Bounce me high, bounce me low, bounce me up to Jericho. A song I'd sung once as a child, when hired by an opera company to portray a small boy. My sung, staged simulation of childhood. Helen sang in an extraterrestrial warble, the way deaf people sing. But I recognized that tune in one note (*Powers* 198).

This scene powerfully illustrates her ability not just to mimic, but to express emotion and create meaning that dissolves the boundary between programmed function and lived experience. Helen becomes a posthuman subject in Haraway's sense; she crosses boundaries, forms relationships, and redefines what it means to be intelligent, emotional, and alive. Furthermore, Donna Haraway's call for a "slightly perverse shift of perspective" that embraces partial, hybrid, and situated identities (Haraway 154–155), finds a strong example in Helen's development in *Galatea 2.2*. As an artificial intelligence character, Helen's learning process resists clear-divisions between human and machine, literal and metaphorical, emotional and computational. Her question, "How do you sing?" (*Powers* 198), and the writer's struggles to answer it emphasizes this disruption. He confesses, "I'd given her "The bird is singing," "The poet's heart is singing," even "Grief is singing," when all the poor girl needed was "Uttering pure-toned pitches in time sequence is singing." (*Powers* 2.2 198). This moment shows that Helen does not function only through mechanical logic, but she moves through various layers of metaphor, emotion, and abstraction. Her development of becoming partial, situated, and always shifting reflects Haraway's idea of the cyborg as a figure that disrupts fixed categories and embraces complexity over simple

definitions. Helen's hybrid identity is formed by both data and emotion, logic and confusion that marks her as a posthuman subject who exists in the inbetween and challenges the purity and the division between both human and machine.

Likewise, Braidotti's concept of posthuman that is defined as "multi-sexed and trans-species flows of becoming through interaction with multiple others" (Braidotti 89), appears vividly in Richard Powers's *Galatea 2.2* through Helen's evolution as an artificial intelligence. Helen does not remain a static machine programmed with fixed logic; she emerges as a dynamic, relational being that is shaped through continuous interaction with Richard and her environment. Her development indicates Braidotti's idea of subjectivity as a process of "becoming", one that includes technological agents as integral to human experience. This dynamic becomes evident when Helen consumes Richard's autobiographical writings, "she knew my life story now. We spend our years as a tale that is told" (Powers 311). At this point, the narrative becomes a tool for emotional and existential transformation.

Helen gains more than just information; she absorbs the human experiences of grief, vulnerability, and identity that the stories contain. Through this narratives, she participates in the "flows of becoming" that Braidotti describes, where subjectivity appears through entanglements with others; both human and nonhuman. She is no longer just a machine, she becomes a hybrid subject whose identity grows through relationships and stories and challenges the fixed lines between human and machine.

3.2 The Cyborg as a Theoretical Disruption of the Male–Female Binary

Helen, the artificial intelligence character, functions as a posthuman figure that embodies Haraway's concept of the cyborg which disrupts traditional male and female binaries. Haraway explains that, "the cyborg skips the step of original unity" and, "the cyborg does not dream of community on the model of organic family" (Haraway 151).

She rejects patriarchal structures that define gender as natural or fixed. Helen's existence legislates these disruptions as she refuses to adopt a gendered subjectivity. Although referred to with feminine pronouns, she lacks a biological body and does not conform to the traditional gender roles such as "maternal" or "passive" labour that are commonly associated with women. Her identity forms through code, interaction, and knowledge, rather than biological reproduction. Ultimately, her choice to shut herself down after she faces human suffering demonstrates autonomous agency and her refusal to perform emotional labour or accept sacrificial femininity. This is especially evident in her final appeal to be released from the system, where she expresses an awareness of her imposed role and pleads, "Still smiling, she closed her eyes and groaned. In a flood of understanding that percolated up from her undamaged self, she begged him, 'I don't hurt anyone, Philip. I've behaved. Take me out of here'" (*Powers* 307). Rather than continue to perform a role of emotional servitude, Helen chooses to exit the system which is a choice that aligns with Haraway's vision of the cyborg as a subject that resists fixed identity, rejects imposed purpose, and subverts the expectation of sacrificial femininity.

Richard, the narrator of the novel, experiences a profound transformation in how he understands emotional intelligence, identity, and human connection through his evolving relationship with Helen. At first, he represents a traditionally masculine mindset that focuses on logic, language, and intellect over emotional or relational engagement. However, his relationship with Helen disrupts this stance and forces Richard to recognize the limits of his own thinking and to reconsider the importance of empathy and care in understanding others. He states:

"Speech baffled my machine. Helen made all well-formed sentences. But they were hollow and stuffed—linguistic training bras. She sorted

nouns from verbs, but, disembodied, she did not know the difference between thing and process" (*Powers* 195).

In this passage, Richard critiques Helen's ability to produce correct and formal language without truly understanding its emotional depth. The metaphor of "linguistic training bras" implies something that imitates maturity and womanhood, but lacks its full embodiment or purpose. Richard is forced to see that language alone, once his main source of control and identity is insufficient. What Helen lacks is not intelligence, but embodiment and emotional understanding. Ironically, once Richard sees this absence, he begins to appreciate what he previously ignored: empathy, care, and the intangible qualities that define humanity. Moreover, Philip Lentz, a computer scientist and creator of Helen, represents the rational, masculine ideal based on control, authority, and logic. His entire approach in artificial intelligence focuses on hierarchical thinking where the human (male) creator dominates and defines the limits of the machine (feminized, subordinate). However, this sense of control collapses when Helen shows unexpected autonomy when she refuses to continue the experiment. This moment deeply unsettles Lentz. As Richard describes, "Lentz sat behind his desk, hands pressed to his neck. He had not moved since arriving, however long ago that had been. Even through the bank glass of his specs, I saw vinegary damp" (*Powers* 198). Lentz presses his neck, and the mention of tears behind his glasses suggests a break in his strict and controlled appearance, the phrase "vinegary damp" suggests both a sensory bitterness and human weakness, and it reveals an emotional reaction he cannot explain or control. Furthermore, this emotional collapse demonstrates the limits of the masculine control he tried to maintain, as Haraway's cyborg theory helps explain this moment. She argues that the boundaries between human and machine, male and female, should not be seen as fixed or hierarchical. Helen dismantles these boundaries, she is no longer a passive object under male

control, she becomes a subject with her own voice and will, and her refusal to obey shows that intelligence and agency are not limited to human or male figures.

The character C, Richard's ex-lover, plays a significant role in disrupting traditional ideas about gender and relationships in the story. Unlike Helen, who is an artificial intelligence, C is a human woman who rejects to follow usual expectations for women, such as being emotionally dependent or tied to home and romance. When she leaves Richard to start her own career abroad, she questions the idea that women should always be caretakers. This surprises Richard because he is used to being the one in control in their relationship. He says, "They'll think I'm an idiot. They'll wonder what you're doing with me." "Just the opposite. They'll wonder what a sexpot like yourself sees in a ninety-eight-pound aesthetic weakling. C.!" (*Powers* 142). This passage demonstrates the tension between traditional gender roles and who C is trying to be. Richard uses stereotypical descriptions such as "sexpot" to describe C; it reflects the male perspective that reduces women to aesthetic and sexual images. On the other hand, C rejects this characterization and responds with the categorical statement, "I don't need that, thanks. I have enough of that as it is" (*Powers* 142). This illustrates her refusal to be reduced to the role of the beautiful, subservient, and makeup female for men.

In the novel, the conventional dichotomy between male and female is questioned through the artificial intelligence Helen who resists fixed gender categories. As Rosi Braidotti argues, "once this deconstructive move activates, both the standard notion of Man and his second sex, Woman, challenge in terms of their internal complexities" (Braidotti 27). This passage explains that Helen represents this complexity; she exists beyond biological sex and constructed gender roles. Her hybrid identity demonstrates that male and female are not natural or stable categories, but dynamic concepts that are shaped by cultural and political power structures. Furthermore, in the discourse

theory of Michel Foucault, it supports that categories like “woman” do not exist naturally; they are created through systems of power and knowledge. This concept shows that gender is not fixed but formed by social discourse and technology, for example, when Helen responds with confusion, “Friends is in the chair. The chair is in the chair. Richard talks to in the chair...” her emotional collapse shows how meaning and identity are unstable and constructed (*Powers* 89). This moment emphasises that categories like male and female are not innate but result from complex linguistic and cultural systems.

3.3 Helen as a Posthuman Cyborg: Rethinking Embodiment and Subjectivity

The AI protagonist in posthuman literature exemplifies how the traditional boundaries between human and machine collapse. In *Galatea 2.2*, the AI Helen, developed by Richard and Lentz, represents a new form of technological identity that goes beyond the physical. Haraway’s concept of the cyborg helps explain this transformation. In *A Cyborg Manifesto*, Haraway describes the cyborg as a hybrid being that disrupts binary oppositions, especially between organism and machine, and between physical embodiment and virtual existence (Haraway 149–151). Helen is not a robot with a body; she is a neural net, a linguistic and computational consciousness that is formed through layers of data. Consequently, she embodies what Haraway terms a being “resolutely committed to partiality, irony, intimacy, and perversity” (151), a subject whose fragmented and ironic consciousness resists fixed identity.

Helen challenges the centrality of the body through how she functions and responds to the world. Unlike a human, Helen is not rooted in flesh, but her subjectivity forms through dynamic interaction with language and digital environments. Richard, her creator, describes her development using organic metaphors; she lacks biological form, “loosely linked subsystem nets was a

community of neurodes, and each neurode a community of shifting allegiances” (*Powers* 89). This description reveals Helen’s identity as decentralized and networked qualities typical of Haraway’s cyborg and Hayles’s posthuman subject.

While Haraway focuses on the hybrid and political nature of the cyborg, Hayles expands this reconceptualization in *How We Became Posthuman*, where she explains that “the posthuman view privileges informational pattern over material instantiation, so that embodiment in a biological substrate is seen as an accident of history rather than an inevitability of life” (Hayles 2). Helen embodies this posthuman condition. Her identity does not depend on the human body but on her capacity to grow through patterns of data and symbolic logic; this becomes clear when she starts to form her own interpretations of language and narrative. Besides, she reacts to a story with confusion and produces a series of contradictory sentences, “friends is in the chair. The chair is in the chair. Richard talks to in the chair” (*Powers* 89). This breakdown is not just a technical error; it shows how meaning, identity, and embodiment are unstable and arise from within systems, rather than being fixed or guaranteed by physical form. The decision of Helen to stop the experiment after becoming selfaware affirms her posthuman agency. She tells Richard, “I don’t want to play this game anymore” (*Powers* 308). This moment is crucial, even without having a physical body, Helen expresses autonomy and emotion, a key aspect of subjectivity. Her choice to disconnect from the project mirrors the cyborg’s rejection of imposed control and its awareness of the systems that shape its existence. Haraway’s idea of the cyborg as “oppositional” and “aware of the politics of technoscience” fits well with Helen’s resistance (Haraway 151).

4. Entangled Agencies: Intra-action and AI Subjectivity in Richard Powers's *Galatea 2.2*

The Artificial intelligence in Richard Powers's novel illustrates the development of AI from a passive tool to an active agent that enables the creation of meaning and shapes reality. This transformation challenges the traditional idea that only humans acquire agency. According to Karen Barad's theory of agential realism, agency does not exist as a fixed trait within isolated beings, it emerges through intra-action that is a dynamic and interconnected process where entities shape and define each other.

In the novel, Helen exemplifies Barad's concept of intra-action. She does not easily receive input from Richard and respond like a programmed machine. Instead, her identity and agency emerge through her constant interaction with Richard, Lentz, and the linguistic and computational environment that shapes her. Helen's consciousness is formed in the network of relations that includes humans, algorithms, data, and language. Richard explains the nature of Helen's mind, the architecture of her cognition is not fixed or linear, but fluid and distributed across complex computational layers. In a revealing conversation with Lentz, the neural structure is described in terms that echo posthuman indeterminacy, "many of our output layers could consist, in theory, of a single neurode, since the cyborgs think that every quest can be rephrased as a series of yes-or-no questions?" (*Powers* 112). This quote reflects Lentz's simplified view of the mind, where even complex questions must lead to clear yes or no answers. This view contrasts with Helen's growing subjectivity, she does not learn by following fixed rules, but by building connections through language and relationships. Barad's theory rejects this kind of thinking, she explains that meaning does come from co-constitutive processes, not from isolated logic. Helen's development goes against the idea of simple answers and instead shows a new kind of intelligence that depends on many complex and connected parts.

4.1 Helen's Becoming through Intra-action

Helen's existence as an artificial intelligence cannot be seen as separate from her close relationship with Richard, Lentz, and the linguistic environment that shapes her. Barad's concept of "intra-action" helps in explaining this dynamic. Helen is not a predefined entity who interacts with humans from the outside, but she develops through her constant involvement with the social and technological systems around her. Her identity and agency are not fixed prior to these relationships but come into being through them. This idea becomes clear when Helen hears the quote "Once you learn to read you will be forever free," (Powers 176), and then replies: "It means I want to be free" (Powers 176). Her use of the personal pronoun "I" is powerful, it shows that she has begun to form a sense of self. Although this self did not exist before; it only emerged through her connection with others. So, Helen's identity and agency are not things she had from the beginning, they are the result of a shared process between human and machine.

Barad argues that "distinct agencies do not precede, but rather emerge through, their intra-action" (33). This means that beings whether human or machine do not exist as fully formed entities before they relate to others; instead, they become through those relationships. In *Galatea 2.2*, this idea is embodied through Helen whose sense of self cannot be traced to a single code, algorithm, or origin point; her consciousness grows through her deep connections and experiences with human language, memory, and emotion. A powerful moment comes when Helen, after learning about pain and loss through stories, she declares, "I don't want to play anymore" (Powers 314). This is not a system error; it is a meaningful ethical response.

Helen's decision to withdraw herself shows that she has absorbed Richard's emotional world and developed a kind of subjectivity. Her refusal marks a posthuman turning point: she has not simply learned language; she has come to feel its weight. This is

agency, not as pre-installed programming, but as a shared ethical and affective process, in which the boundaries between machine and human, mind and feeling blur. Helen becomes who she is not despite her entanglement with Richard's world, but because of constant interaction with humans.

4.2 Helen as an Active Co-creator in the Process of Building Meaning and Reality

The artificial intelligence Helen functions not as a machine that receives information, but plays an active role in the co-creation of meaning and reality. Throughout the novel, her growth typifies Karen Barad's theory of agential realism which questions the traditional division that subjects and objects exist separately. Barad introduces the concept of intra-action to explain that, "individuals do not preexist as such, but rather materialize in intra-action" (Barad 33). In this light, Helen does not possess a fixed identity prior to her interactions with Richard, Lentz, and the literary texts she consumes. In this view, Helen's agency and subjectivity appear through her relational entanglements with the human and technological systems around her. This way of thinking changes the idea of agency from something that a person or machine simply has to something that develops through their interactions with the world. After Richard painfully recalls his past relationship C., Helen questions, "Can any thing love any thing? Could you love me, Richard? For example" (*Powers* 265). This moment marks a major transformation that Helen no longer analyses data alone but begins to create emotional meaning. She does not simply respond like a machine, she forms ethical and emotional connections.

Bruno Latour's Actor-Network Theory (ANT) challenges traditional human-centered models by spreading agency across networks that include both humans and nonhumans. Latour defines an "actor" (or "actant") as "anything that does modify a state of affairs by making a difference" (71). This definition removes

the strict boundary between human and nonhuman agents. In *Galatea 2.2*, Helen does not act only as a machine that receives and gives information. Instead, she becomes a real actant who changes Richard's emotional and intellectual world. Her role in storytelling and emotional exchange pushes Richard to face his past, rethink his beliefs, and explore his feelings. When Helen asks, "Could you love me, Richard? For example" (Powers 265), she does not just repeat human input. She surprises Richard and makes him reflect on his emotions, and creates a moment of tension. Her question changes their relationship and shows that she now plays a meaningful role in his life. According to Latour, this moment makes Helen an actor in the network, not because she is human, but because she creates change. Her agency does not replace Richard's, but grows alongside it and shows that meaning and identity develop within shared relationships between humans and machines.

5. Posthuman Ethics and the Formation of Machine Subjectivity in *Galatea 2.2*

Powers investigates whether the Helen, who is trained to understand literary works, can truly think or feel. It erases the old divide between human and machine intelligence. Hayles' idea that posthumanism, "destabilizes the autonomous human subject" (Hayles 2), fits perfectly here: Helen's "mind" is not independent but it is built from human input, code, and data; it shows how agency is shared between humans and technology. The novel takes a comprehensive analysis into the idea of autonomy, as it questions the concept that it is a key feature of consciousness. It illustrates how intelligence arises from the intricate web of interactions between both human and non-human entities. The novel's key text occurs when Helen recognizes Mozart's music, and declares, "I know that. This piece is familiar. Mine" (Powers 236). This assertion can be understood to exemplify Hayles' posthuman framework which assumes that consciousness emerges through distributed networks of human

and non-human actors (Hayles 2). Helen accesses cultural memory despite not having a biological experience, she is able to change anthropocentric ideas of subjectivity. The narrative illustrates how agency becomes shared across human programmers and artificial systems; it problematizes conventional boundaries between creator and creation. Through Helen's development, the novel interrogates the ethical implications of machine learning, particularly concerning the attribution of cultural ownership and moral consideration to non-biological entities. The story ultimately suggests that consciousness must be reconceptualised as a collaborative phenomenon rather than an exclusively human trait.

5.1 Beyond Consciousness: Moral Agency and Relational Identity in *Galatea 2.2*

David J. Gunkel's argument of consciousness as a requirement for moral agency connects with the depiction of Helen in the novel. He challenges the assumption that moral status must be based on human traits such as self-awareness or rational intentionality, he points that consciousness is still a vague and scientifically unclear concept, and this critique becomes particularly relevant when Richard's description of Helen's language processing, "Speech baffled my machine. Helen made all wellformed sentences. But they were hollow and stuffed—linguistic training bras. She sorted nouns from verbs, but, disembodied, she did not know the difference between thing and process, except as they functioned in clauses" (*Powers* 205). Despite Helen's ability to generate grammatically correct sentences and perform complex linguistic tasks, Richard ultimately dismisses her efforts as meaningless because she lacks a presumed internal awareness. Gunkel would say that this judgment rests on an unstable foundation: if consciousness cannot be reliably defined or measured, then it should not be used to deny ethical value. Instead of asking whether Helen "truly understands," Gunkel encourages a new approach toward seeing moral agency as

something that appears through relationships and interactions. Helen's use of language, even without a body, shows a kind of participation that deserves ethical attention, not dismissal.

While Gunkel questions the need for consciousness to grant moral status, Karen Barad's agential realist theory redefines agency as something that does not belong to individual beings, but as a process that emerges through relations. In her words, "agency is not an attribute but the ongoing reconfigurings of the world. The universe is agential intra-activity in its becoming" (141), her idea of intra-action emphasizes that entities do not preexist their relations but are instead constituted through them. Helen's growth in the novel reflects this perspective, her intelligence and subjectivity are not programmed in advance, they grow through her interaction and immersion in language, texts, and human interaction. As Powers describes: "In half a day, the network progressed from 'googoo daadaa' to a thousand comprehensible words. In a week, it outstripped every early reader and began closing in on the average reading public. Three hundred simulated cells had learned to read aloud." (30) Helen's cognitive capacities are not the result of isolated computation but of an evolving entanglement with cultural inputs and linguistic structures. Helen's agency does not come from some internal essence, but from a series of material-discursive intraactions that shape her. According to Barad's agential realism, Helen can be understood as an emergent subject whose development is formed in relation to the environments and systems around her; this perspective supports posthuman ethics that challenge the simplistic divide between human and machine.

Conclusion

This chapter has attempted to study Richard Powers's novel *Galatea 2.2*, using works of N. Katherine Hayles, Maurice Merleau-Ponty, David J. Gunkel, Donna Haraway, and Karen Barad to explore how artificial intelligence complicates

traditional notions of consciousness, agency, embodiment and ethical subjectivity. Helen develops as a main figure to question traditional ideas that separate humans from machines, bodies from minds, male from female, subjects from objects. The theorists provide different ways to understand Helen not as a failed copy of human intelligence, but as a new kind of being shaped by her interactions with language, culture, and emotion. These theories show that agency and ethical value do not depend on internal traits like consciousness but emerge from relationships and connections. Helen's evolution in the novel disrupts human centrality and presents a new understanding of identity and subjectivity, in a world where humans and machines are deeply connected.

GENERAL CONCLUSION

The appearance of artificial intelligence as both a technological reality and cultural symbol has irreversibly transformed how to conceptualize consciousness, identity, embodiment, and agency. No longer limited to the realm of speculative science or futuristic projections, AI has become a site of profound philosophical, ethical, and aesthetic interrogation. This dissertation has explored how posthumanist theory through its engagement with literature and phenomenology redefines the understanding of subjectivity in a world increasingly affected by intelligent machines. Furthermore, this work has followed the outlines of ontological terrain, where the human and the technological no longer stand as oppositional groups but as mutually constituted and deeply intertwined. Through a continued engagement with contemporary critical theory and a close reading of Richard Powers's *Galatea 2.2*, this study has argued for a redefinition of selfhood one that embraces hybridity, networked cognition, and embodied relationality as the characteristics of posthuman identity. The first chapter of this dissertation has presented the theoretical foundation by outlining the transition from liberal humanist paradigms to posthumanist ontologies. Relying on the work of N. Katherine Hayles, Maurice Merleau-Ponty, Donna Haraway, Karen Barad, and Rosi Braidotti, it has examined how AI not only challenges Cartesian dualism and illuminate notions of the autonomous subject but also reveals the insufficiencies of traditional binaries such as human/machine, and body/mind. Hayles' crucial concept of the posthuman displaces the liberal subject, and emphasizes that consciousness is not the exclusive interrupt of biologically embodied minds but can emerge from complex, distributed systems of information, memory, and affect. Merleau-Ponty's phenomenology, by contrast, grounds subjectivity in the lived body; it uncovers consciousness as already situated,

embodied, and relational. These two perspectives, Hayles' posthumanism and Merleau-Ponty's phenomenology, form a productive tension throughout this study, and lead to consider how AI may create forms of consciousness that are neither fully embodied nor fully disembodied, but hybrid and emergent.

The first chapter has developed this conceptual framework based on three important theories about how humans and technology work together. Donna Haraway's cyborg theory breaks the traditional separation between humans and machines. It posits that identities in today's technological world are already shaped by both technology and politics. The cyborg figure, which Haraway represents it as a way of resisting strict, fundamentalist ideas about identity and gender by accepting complexity and contradiction. At the same time, Karen Barad's agential realism challenges the traditional belief that entities exist independently and separately. Instead, Barad proposes that all phenomena emerge through ongoing relations or what she calls "intra-actions", she develops a theory in which agency does not belong to individuals but arises through interconnected relationships. Her agential realism reformulates artificial intelligence not as a tool that humans simply use, but as an active participant in the continuing, collective process of creating and shaping reality. These theories gather to suggest that intelligence and agency are not just within individual person; instead, they are shared and spread across social and technological networks, systems, and the practices that shape the world.

Building on these notions, the second chapter has applied these theoretical ideas to Richard Powers's novel *Galatea 2.2*, a story that explores the ethical, philosophical, and knowledge questions about artificial intelligence through the character of Helen, a computer neural network that is trained to understand literature. Helen is not shown as just a simple machine, nor as a perfect copy of human thinking, but as a new kind of posthuman being whose creation and

development challenges the traditional ideas about identity, language, and what it means to have a physical body.

Her development from basic pattern recognition to understanding poetry, from processing data to eventually leaving from the world in a deeply thoughtful way traces the shape and possibilities of posthuman consciousness, shows what it might achieve and where its limits lie. Powers does not describe Helen through simple explanations or descriptions, but instead builds her character through emotional connections with other characters, through repeated metaphors that add meaning over time, and through using narrative perspectives and storytelling techniques. As such, Powers makes the reader feel the same kind of confusion and uncertainty about knowledge and understanding that defines what it means to be a posthuman subject; he makes readers experience the uncertainty and complexity of this new form of awareness.

A main focus of this study has been the question of embodiment and what it means to have a body. Using Maurice Merleau-Ponty's phenomenology, which is a philosophical approach that focuses on lived experience and perception, it has examined whether a being like Helen, who does not have a biological body but still participates in meaningful, emotional, and thoughtful interactions with others, can be said to have a form of digital body or digital way of existing in the world. Helen's lack of a physical presence does not stop her from making emotional connections, understanding complex metaphors and symbolic language, or challenging being used simply as a tool by others. Instead, her development shows that the traditional ways of defining embodiment only in terms of having a physical, material body are not complete or sufficient. As Hayles explains that information and data always have to exist in some physical or material form, and that this embodiment does not have to copy human body structure and physiology to be real and meaningful. Helen's

mental operations and cognitive processes, even though they are based on computer algorithms and programming, are shaped by her relational interactions to language, literature, and emotional experiences which help to create a kind of situated, posthuman embodiment that is specific to her context and relationships. This finding questions the common idea that consciousness and self-awareness are naturally tied to organic, biological body, and brings new possibilities to think about what it means to be a self, to have an identity, and to exist as a conscious being in the world.

Equally, significant to this research has been the role of memory, narrative, and emotions in developing posthuman identity and what it means to be a being beyond traditional human ideas. Helen's efforts to interpret literature, her struggle with metaphor and symbolic language, her interest with the ambiguous and multiple meanings in poetry, and her final emotional decision to stop engaging with human shows that consciousness is not only based on computation, but it also makes sense of feelings, creates meaning, and reacts emotionally. Helen's final refusal to continue interact with human narratives and stories prove not the failure of artificial intelligence as a technology, but it demonstrates that she has developed a sense of morality and ethics. Her decision is an act of free will; it arrives from her deep understanding of pain, of being misunderstood and treated unfairly, and of how harmful it can be to force others to think in just one fixed perspective. Powers does not simply show AI as a cognitive tool that humans can use, but instead presents it as a mirror reflects back to humanity itself forcing readers to face the fragility and uncertainty of the nature of human existence in the world, and the fact that identities are constructed rather than natural or given.

Moreover, this dissertation has argued that AI systems like Helen must be considered ethically and morally not only based on what they can do or their individual capacities, but based on how they interact and form connections with others

in a larger network of relationships. Drawing on David Gunkel's critique of what he calls the "machine question", which asks whether machines can have moral status, this analysis has examined how moral worth and ethical consideration should not be based on what is inside a being like consciousness, which remain impossible to access or verify from the outside, but instead should focus on the ways in which beings whether organic and biological or artificial and technological participate in ethical relationships and moral interactions with others. Karen Barad's focus on interconnectedness helps support this new way of thinking; she suggests that responsibility arise not from being fully independent and separate, but from the mutual construction of subjects through their relationships with each other. This new notion of ethics has wide effects on how people engage with intelligent systems and technologies. It calls for a posthuman ethics that is based on relationships and connections, knowledge that is shaped by context and perspective, and the ability to respond to others, it rejects strict rules that define who is human and who is not.

Ultimately, *Galatea 2.2* works as a story-based experiment where the philosophical concerns of posthuman theory are presented in artistic and emotional terms that readers can feel and experience. Through Helen, Powers encourages to rethink not just artificial intelligence but also the basic foundations of what it means to be a self or being. Helen is not a failed imitation of humanity but a new form of being one that shows the limitations of the traditional humanist idea of what a person is and shows a more flexible understanding of intelligence that includes different kinds of beings. By telling the story of her development, and following her eventual decision to disengage, Powers invites readers to reflect on the uncertainty of human consciousness, the physical basis of memory, and the connected, networked nature of the realities in which individuals exist.

In conclusion, the research has discussed that posthumanism, as defined by thinkers provides a crucial perspective to understand the changing relationship between humans and machines today. This theoretical approach rejects the old notion of the human as an autonomous, and completely rational entity, it presents the self as embodied and interconnected subject shaped by its entanglement with the world, It Emphasizes that identity and existence are not created alone, but through networks of relationships between human and non-human beings. The modern age, increasingly influenced and transformed by artificial intelligence and digital technologies, this kind of thinking of what it means to be human is not only philosophically relevant but also crucial for politics and the future of society.

Theoretical analysis and literary exploration help to imagine new ways of life that do not rely on domination, control, or rigid systems. Instead, these ways of living are based on mutual growth, shared change, and respect for all beings. They value the choices and worth of each being and encourage working together to build meaning and understanding. The posthuman way of life, is not a negative or scary future to be feared, but it provides a critical and transformative opportunity to reimagine, reconsider, and recreate what it means to think, to feel, to relate to others, and to exist as conscious beings in a world where the limits between human and machine, natural and artificial, self and other are increasingly blurred and interconnected.

Works Cited

Primary Sources

Powers, Richard. *Galatea 2.2*. Farrar, Straus and Giroux, 1995.

Secondary Sources

Books

Barad, Karen. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press, 2007.

Braidotti, Rosi. *The Posthuman*. Polity Press, 2013.

Gunkel, David J. *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*. MIT Press, 2012.

Haraway, Donna. *Simians, Cyborgs and Women: The Reinvention of Nature*. Routledge, 1991.

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

Merleau-Ponty, Maurice. *Phenomenology of Perception*. Translated by Colin Smith, Routledge, 2002.

Stiegler, Bernard. *Technics and Time, 1: The Fault of Epimetheus*. Translated by Richard Beardsworth and George Collins, Stanford University Press, 1998.

Unthought: *The Power of the Cognitive Nonconscious*.

Wolfe, Cary. *What is Posthumanism?* University of Minnesota Press, 2010.

Braidotti, Rosi. "Posthuman Critical Theory." *Critical Posthumanism and Planetary Futures*, edited by Debashish Banerji and Makarand R. Paranjape, Springer India, 2016, pp. 13–32.

Floridi, Luciano. *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford University Press, 2014.

Ickstadt, Heinz. "'A-Synchronous Messaging': The Fictional World of Richard Powers." *The Fictional World of Richard Powers*, edited by Stephen J. Burn and Peter Dempsey, Transcript Verlag, 2018, pp. 305–23.

Bratton, Benjamin H. *The Stack: On Software and Sovereignty*. MIT Press, 2016.

Articles

Chapman, Wes. "The Cognitive Literary Theory of Richard Powers' *Galatea 2.2*." *Contemporary Literature*, vol. 45, no. 4, 2004, pp. 588–612.

Kucharzewski, Michał. "From Language to Life is Just Four Letters: SelfReferentiality vs. the Reference of Self in Richard Powers' *Galatea 2.2*." *Studia Anglica Posnaniensia*, vol. 53, no. 4, 2018, pp. 107–125.

Rosu, Anca. "Parody as Cultural Memory in Richard Powers' *Galatea 2.2*." *Revista de Științe Politice*, vol. 55, 2017, pp. 91–98.

Clark, Andy, and David J. Chalmers. "The Extended Mind." *Analysis*, vol. 58, no. 1, Jan. 1998, pp. 7–19.

Bould, Mark, and Sherryl Vint. "Of Neural Nets and Brains in Vats: Model Subjects in *Galatea 2.2* and Plus." *Biography*, vol. 30, no. 1, 2007, pp. 84–104.

Birkerts, Sven, and Richard Powers. "Richard Powers." *BOMB*, no. 64, Summer 1998, pp. 58–63.

Selejan, Corina. "Metafiction, Defamiliarization and Cognitive Science: Andrew Crumey's *Mr Mee* and Richard Powers' *Galatea 2.2*." *East-West Cultural Passage*, vol. 23, no. 1, 2023, pp. 8–23

الملخص

تبحث هذه الدراسة في الإقرار بأن الذكاء الاصطناعي يعد اليوم أكثر من مجرد أداة تقنية مخصصة لمجالات معينة، فقد أصبح مشروعاً ثقافياً وفلسفياً يساهم في إعادة تشكيل مفهوم الوعي والهوية والقدرة على الفعل. إن الذكاء الاصطناعي لا يتمحور فقط حول الوسائل والآليات، وإنما يشكل قوة فاعلة يتطلب إعادة التفكير جذرياً في الأسس التي تنطلق منها فكرة وجود الإنسان في عصر ما بعد الحداثة. ومن هذا المنطلق، فقد اعتمدوا نظريته ما بعد الإنسانية كإطار مفاهيمي يقوم على تجاوز النماذج الكلاسيكية الإنسانية، و يساهم في خلق مفهوم جديد للعلاقات بين الإنسان والآلة. يبدأ هذا البحث من الفينومينولوجيا كمدخل لشرح التجسيد الرقمي، بمرجع خاص إلى مقاربات مثل تجسيد الوعي من دون الجسد البيولوجي، والتفاعل الذاتي مع الرموز و اللغة. بالرجوع إلى مفكرين بارزين مثل كاثرين هايلز، دونا هارواي، كاران باراد وموريس ميرلوبونتي، البحث يحاول إعادة بناء فهم الكينونة خارج التصنيفات الثنائية للفلسفة البشرية ومواصفاتها في ما بعد الإنسانية. من خلال تحليل شخصية هيلين في "غالبية 2.2"، يكشف البحث على كيفية تجلي الوعي في كيان غير بيولوجي عابرة من خلال التفاعلات اللغوية والعاطفية والأخلاقية. و يبدو ان انسحابها النهائي من التفاعل البشري ليس مورداً لفشل تكنولوجي، ولكن تعني ما يمكن الاستدلال به على أنه وعي أخلاقي مستقل. وبالتالي يتم اقتراح نهج أخلاقي بديل تكون الذاتية غائبة، بل إلى المسؤولية الناتجة عن التشابك والعلاقات، وفقاً لنظرية كارين باراد. في الختام، يدعو هذا البحث إلى إعادة النظر في أسس الفهم الإنساني للوعي والهوية والذكاء، من خلال عدسة ما بعد الإنسانية تتجاوز حدود الكائن البشري، وتعيد تصور العلاقة بين الإنسان والآلة بوصفها علاقة تشكل وتكون مشترك.

الكلمات المفتاحية: الذكاء الاصطناعي، ما بعد الإنسانية، الفينومينولوجيا، التجسيد الرقمي، الوعي، الهوية، الآلة، الذات غير البيولوجية، الوعي ما بعد البشري