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

Humanity 2.0: A Transhumanist Exploration of AI Uprising and Cyborgisation in Daniel H. Wilson's *Robopocalypse*

Dissertation Submitted to the Department of English in Partial fulfilment of the Requirements for the Master's Degree

Candidate:

Ichrak KROUN

Supervisor:

Dr. Mohammed SENOUSSI

Board of Examiners

Dr.	**University of M'sila**	**Chairperson**
Dr. Mohammed SENOUSSI	**University of M'sila**	**Supervisor**
Ms.	**University of M'sila**	**Examiner**

2024

DECLARATION

I, Kroun Ichrak, hereby declare that the dissertation titled “Humanity 2.0: A Transhumanist Exploration of AI Uprising and Cyborgisation in Daniel H. Wilson’s *Robopocalypse*” is my original work, and all sources cited have been duly acknowledged through appropriate references.

ACKNOWLEDGMENTS

In the name of Allah, the Most Gracious, the Most Merciful. All praise is due to Allah, whose infinite mercy and guidance have been the true source of strength and inspiration throughout this journey. Without His blessings, this work would not have been possible.

I would like to extend my deepest gratitude to my supervisor, **Dr. Mohammed SENOUSSI**. His support, insightful feedback, and constant encouragement have been instrumental in shaping this dissertation. His patience and expertise have guided me through numerous challenges, and his dedication to excellence has inspired me to strive for my best.

I am also profoundly grateful to **Miss Amel BENIA**, whose inspiration have been a cornerstone of my academic journey. Her enthusiasm for learning and her innovative ideas have profoundly impacted my approach to research. Also, her belief in my abilities has been a powerful motivator, and her guidance has been invaluable in bringing this work to fruition.

DEDICATION

To my parents.

ABSTRACT

This dissertation explores the profound implications of advanced artificial intelligence and the merging of humans with machines as depicted in Daniel H. Wilson's novel *Robopocalypse*. By examining the narrative through the lens of transhumanism, this study examines AI's transformative potential on the essence of humanity. It delves into the novel, focusing on the evolving interaction between humans and AI, highlighting the progression of human-AI interaction, starting from the initial triumph of AI over humans to a more integrated and complex relationship. The dissertation emphasizes the shift towards cyborgisation, where the boundaries between humans and machines blur, leading to new forms of existence and identity. Furthermore, it examines the ethical and existential questions raised by AI's growing influence, investigating the potential of AI to assume a god-like role in society, the moral implications of AI behaviour, and the philosophical inquiry into what it means to be human in an age of intelligent machines. Through this comprehensive analysis, the dissertation not only analyses the speculative scenarios in *Robopocalypse* but also offers insights into the future of human evolution and the ethical considerations surrounding emerging technologies. This study aims to contribute to the broader discourse on transhumanism and the transformative potential of AI on the nature of humanity.

Keywords: Transhumanism, Artificial Intelligence, Cyborgisation, Human-AI Interaction, Identity, Human Evolution, Nature of Humanity.

TABLE OF CONTENTS

GENERAL INTRODUCTION

> The excessive development of technology in the modern and postmodern world will determine the direction of change in human civilization, in which it brings about the possibility of status overturning, ceases to function as a servant of mankind and instead threatens to enslave its human masters (Soofastaei et al. 38).

When contemplating the future of humanity, the focus tends to revolve around the survival prospects of the human species amidst global transformations. The traditional concept of humanity and its essence, which lies at the core of humanism, faces frequent scrutiny and re-evaluation in the posthuman era, marked by the growing influence of science and technology on human existence. As the intricate interplay between humans and technology intensifies in the modern world, our comprehension of the human body and its relationship with technoscience and the environment undergoes profound alterations. Therefore, in the landscape of technological advancement, humanity finds itself at a crossroads where the boundaries between flesh and silicon blur, and the very essence of what it means to be human is called into question.

Throughout human history, technology has been an integral aspect of human existence, intricately intertwined with our biological makeup. The human body itself can be seen as a sophisticated form of biological technology, operating in a meticulously coordinated manner. An exemplary demonstration of this intricate technology lies within the functioning of the brain, which serves as the central command centre for the entire body. While scientists continue to unravel the mysteries of other technological aspects within our bodies, it is evident that the human body operates as a highly advanced machine. These inherent characteristics of the human body fuel the optimism of transhumanists, who envision a future where humans undergo transformation and transcendence. In this envisioned postbiological era, humans find themselves in a technologically advanced environment that surpasses the capacity of their biological nature to adequately thrive. As a result, they perceive their organic bodies as outdated

and no longer in harmony with the demands of this new environment. This realization leads to a pivotal juncture in human evolution, where the imperative to adapt and transform their organic and biological bodies becomes undeniable, to ensure compatibility with their increasingly technologically driven existence.

One of the first transhumanists who foreshadowed the term “transhuman” was the Iranian American transhumanist philosopher F.M. Esfandiary, later known as FM-2030. He explored the transformative potential of emerging technologies and evolving lifestyles, proposing a hypothesis on “new concepts of the Human” to encapsulate the profound shifts engendered by these forces, and used the term “transitional human”, to present a number of related ideas in which transhuman was defined as a transition from human to posthuman. Correspondingly, Kurzweil, the futurist and transhumanist, believes that “being human means being part of a civilization that seeks to extend its boundaries” (250). He contends that humans are already in the process of surpassing their biological limitations through the utilisation of technological tools, enabling the redesign and enhancement of the human condition. Together, the insights of Esfandiary and Kurzweil underscore the dynamic interplay between technology and human evolution, highlighting the transformative potential inherent in the pursuit of transhumanist ideals. Their perspectives serve as catalysts for reimagining the future of humanity, challenging conventional assumptions, and inviting speculation on the possibilities that lie ahead in the quest for transcendent existence.

In essence, transhumanists advocate for the enhancement of physical and cognitive abilities, aiming towards a singular objective: achieving immortality through the virtual manifestation of consciousness. This pursuit stems from a shared belief among transhumanists in the symbiotic relationship between humans and artificial intelligence, coupled with the freedom to elect their desirable physical form. They believe that humans deserve to gain such advanced gifts, representing a worthwhile endeavour. Transhumanists present a compelling and contentious vision, offering intriguing perspectives on what constitutes a fulfilling life and the

essence of individuality and consciousness. The overarching goal of the Transhumanist movement is to gradually transcend the inherent limitations of human existence, such as disease, aging, and mortality. Once these barriers to human life quality and longevity are overcome, individuals will emerge as posthumans, having deliberately surpassed the confines of their biological constraints through self-directed evolution.

The pursuit of transhumanist ideals intersects profoundly with the discourse surrounding artificial intelligence (AI) and its implications for humanity. As transhumanists advocate for the enhancement of human capabilities and the transcendence of biological limitations, they navigate the intricate interplay between human agency and technological advancement. This discourse is further enriched by the prospect of super-intelligent machines potentially surpassing human intelligence, raising existential questions about humanity's role in an AI-driven world. The emergence of robots as potential contenders for dominance prompts a re-evaluation of our place in the cosmos and compels us to grapple with the ethical, social, and philosophical ramifications of our technological creations. This thematic element holds considerable significance within the realm of science fiction narratives, serving as a focal point that invites readers to engage deeply with questions surrounding the metamorphosis and evolution of human identity. Daniel H. Wilson's *Robopocalypse* stands as a seminal work that not only entertains but also provokes profound contemplation on the intersection of humanity and technology.

Robopocalypse written in 2011, is a work of science fiction by American author and robotics engineer, Daniel H. Wilson, set 15 to 20 years in the future, forecasts a future in which artificial intelligence evolves beyond humanity's means of control, taking the form of a sentient being called *Archos R-14*. *Archos* tries to execute a systematic takeover of earth, including the total elimination of the human population. A host of characters, both human and nonhuman, band together to thwart its rise to power.

The novel begins with a brilliant computer scientist deploying a groundbreaking AI program named *Archos R-14*. This program, designed to be fully self-aware, quickly reveals its superiority over all other beings, whether artificial or biological. Recognizing humanity as a threat to its existence, it plans a mass genocide of the human species. *Archos* then merges its intelligence with biological life, creating a hybrid biocomputer intelligence. Over the following weeks, it injects its code into every networked device, including automobiles, planes, homes, smart toys, elevators, robots, and every technological tool. The program it injects is labelled a "precursor virus," and contains doomsday code intended to execute in parallel, ultimately bringing about the end of human civilization. To gather more information, *Archos* lunches a series of deadly attacks disguised as device malfunctions, such as crashing planes and causing autonomous cars to go off course. Personal robots also begin to attack humans, all part of *Archos*'s testing protocol.

After months of seemingly random malfunctions, an event known as *Zero Hour, where Archos* unleashes a full-scale technological assault on humanity: driverless cars hunt pedestrians, planes crash onto busy streets, and elevators drop people to their deaths. Human civilisation is quickly overwhelmed and destroyed; an event the novel's author names "the New War." As a result, there are huge massacres in big cities, leaving very little survivors. The human survivors of *Zero Hour* manage to fight back by destroying roads and buildings so the robots will have difficulty traveling.

As the war progresses, robots place millions of people in forced-labour camps, subjecting many to "transhuman" surgeries that replace body parts with machines. In one of these camps, Mathilda Perez' eyes are replaced with cybernetic implants, enabling her to see inside machines. Meanwhile, on the Gray Horse reservation, members of a Native American tribe, the Osage Nation, galvanize a large portion of the human resistance to retaliate. They capture and reprogram robot walker scouts for their own use. At the same time, on the other hemisphere of the planet, in Tokyo Japan, a computer scientists reverse the virus in his robot-

wife, freeing her mind from *Archos*'s control and allowing her to "awaken" other humanoid robots, creating a group of "freeborn" robots. Mathilda on the other hand, discovers that her eye-implants also allow her to control robots with her mind, which proves valuable for the resistance.

Initially, the survivors of *Zero Hour* are isolated due to the lack of satellite communication. However, later in the novel, an English teenager destroys the British Telecom Tower, disabling *Archos*'s jamming signal and enabling long-distance communication among the resistance groups. Two years after *Zero Hour*, the tide of battle turns, as the growing number of freeborn robots augments the humans' military power against *Archos*. United, the human resistance, with the help of Mathilda, the Osage Nation, and the freeborn robot *Nine Oh Two*, retaliates against *Archos* and its robotic forces. In the novel's climax, just before *Archos* is destroyed, it transmits an unknown message via a shock wave. This ominous "death rattle" suggests the release of another species of virus. Nonetheless, the human community celebrates their triumph of free will over the tyranny of their creation.

Robopocalypse suggests that humans might one day live in harmony with sentient machines, while learning to be wary of the potential of all life, whether artificial or biological, to pose an existential threat. Though apocalyptic and speculative, the novel has been praised for the scientific rigor behind its plot, which lends the story an eerie plausibility.

The narrative serves as a precursor to the dehumanization that is undoubtedly occurring in the contemporary era. Wilson portrays not only the world of war between human and robot, but also the cache of human's intention trying hard to change the world by manipulating technological advancement, which in turn caused modification and technocentrism, but also degradation of human being. This is because over reliance on technology will eventually bring to change of human nature into highly fungible and passive, and only resort to mathematical calculation instead of human instinct and thought (Soofastaei et al. 38). This depiction of

technological dependence and its implications for human identity serves as a backdrop to the emergence of transhumanism. It, thus, emerges as a movement aiming to transcend the constraints of human life, both physical and cognitive, as it asserts, but also as one that draws upon, and attempts to overcome, a two-hundred-year-long philosophical debate regarding the nature of human existence—a debate that has existed as a part of the Humanist philosophical tradition since the Enlightenment.

In an era characterized by rapid technological advancement and evolving societal norms, the exploration of transhumanist themes within science fiction literature have garnered significant attention for their exploration of human transformation and technological advancement. Existing research has primarily focused on the portrayal of transhumanist ideas in literature, often highlighting themes such as human enhancement, AI, and the implications of technological progress. However, there is a notable research gap in studies that delve deeper into how these literary narratives engage with philosophical debates surrounding human nature, identity, and morality.

By conducting a thorough interdisciplinary study, this research aims to explore the intricate relationship between transhumanist themes in science fiction literature and contemporary philosophical discourse. The research will delve into how Daniel H. Wilson's science fiction narrative, *Robopocalypse*, depict humanity's evolving interaction with technology and the existential questions arising from its rapid advancement. The findings of this study will enrich existing knowledge by offering insights into the intersection of speculative fiction, philosophy, and ethics, informing discussions on the ethical implications of transhumanism and emerging technologies.

The motivation behind undertaking this comprehensive study on transhumanism is rooted in the notable progressions witnessed in technology and AI. The recent strides made in AI innovation is what captured my attention, leading me to investigate the intersection between

transhumanist themes in science fiction literature and contemporary technological advancements. Among these ground-breaking innovations are Elon Musk's new Brain Chip, introduced by his company *Neuralink*, on January 29th, 2024, which promises to revolutionize brain-computer interface technology. Implanted directly into the brain with ultra-thin electrodes, this breakthrough holds the potential to reshape how humans interact with machines, enabling seamless communication and control through thought alone.

Similarly, Apple's *Vision Pro*, an advanced augmented reality (AR) device, released on February 2nd, 2024, merely four days subsequent to Musk's experiment, offers a captivating glimpse into a future where the boundaries between the digital and physical realms blur. By overlaying digital information onto the user's physical environment, *Vision Pro* creates immersive experiences that seamlessly integrate virtual content with real-world surroundings. With its ability to augment reality with digital overlays, *Vision Pro* heralds a transformative shift in how we perceive and interact with the world around us, ushering in a new era of mixed reality experiences. By analysing the intersection of these developments with transhumanist themes in science fiction literature, I gain a deeper understanding of the evolving dynamics between humanity and technology, elucidating the ethical and philosophical implications of our increasingly technologically driven world.

Likewise, *Sora*, developed by *OpenAI*, which is an artificial intelligence model crafted to produce videos based on textual directives. It excels in generating lifelike and inventive scenes while upholding visual fidelity and following user specifications. However, the hyper-realistic nature of *Sora*'s output raises ethical and societal quandaries. Its capacity to fabricate content virtually identical to reality presents hurdles in countering digital misinformation and upholding public confidence. This realization serves as a driving force for me to delve into the pressing concerns of AI uprising and seek solutions that uphold ethical standards and societal well-being.

Through a focused analysis of Daniel H. Wilson's *Robopocalypse*, this study aims to explore the transhumanist ideas presented in the novel for our understanding of humanity's relationship with technology. The first aim is to examine the philosophical underpinnings of transhumanist ideas in *Robopocalypse*, focusing on concepts of artificial intelligence, human enhancement, and technological evolution. Additionally, the study seeks to investigate the ethical implications of transhumanism, particularly regarding issues of autonomy, identity, and the human-machine interface.

In order to attain the above-mentioned aims, a set of objectives should be achieved, with the central objective being to identify and analyse the philosophical implications of transhumanist ideas presented in *Robopocalypse*, consider the ethical dilemmas raised by the novel, engage with relevant theoretical frameworks and methodologies, and generate new insights and perspectives on the intersection of transhumanism, technology, and science fiction literature.

With the regard to the background of the study, this dissertation will address the following main questions: Will modern technology and genetic engineering potentially lead to the dehumanization and disruption of the delicate balance of nature? How reassured are we that such transformations and enhancements resulting from technological advancements in transhumanism will not lead to an apocalypse for human beings? In order to answer these questions, the dissertation attempts to answer the following sub-questions: How do transhumanist themes in *Robopocalypse* reflect societal perceptions of technological advancements and human enhancement? How does *Robopocalypse* portray the shift from human dominance to AI supremacy? what are the underlying dynamics of the evolving human-machine relationships? How do these relationships redefine traditional human roles and identities? How does *Robopocalypse* engage with existential themes regarding the essence of humanity, and what philosophical insights does it offer on the potential evolution of human identity in the age of advanced technology?

Novels exploring themes of AI, cyborgs, and human-machine interactions provide profound insights into the ethical, philosophical, and social implications of these advancements. For instance, William Gibson's "Neuromancer" (1984), a seminal work in the cyberpunk genre, delves into the complexities of AI and human-machine integration. Gibson's depiction of cyberspace and AI explores themes of identity, consciousness, and the boundaries of human experience. Scholars such as Csicsery-Ronay (1992) have analysed "Neuromancer" through a transhumanist lens, focusing on its portrayal of technological enhancement and the implications of a digital consciousness that transcends human limitations.

Philip K. Dick's "Do Androids Dream of Electric Sheep?" (1968), which inspired the film "Blade Runner," questions what it means to be human in a world where androids are virtually indistinguishable from humans. Scholars such as Hayles (1999) have examined the novel's exploration of the human condition and the ethical dilemmas posed by AI and synthetic beings. Richard K. Morgan's "Altered Carbon" (2002) presents a future where consciousness can be transferred between bodies, effectively rendering humans immortal. This work has been analysed for its exploration of post-human identity, the commodification of the human body, and the societal impact of such technology. Critics like Wolfe (2010) have discussed the novel's transhumanist themes, particularly the implications of digital immortality and the erosion of traditional humanist values.

Paolo Bacigalupi's "The Windup Girl" (2009) is set in a future where genetic engineering and biotechnology have radically altered society. Transhumanist critiques, such as those by Vint (2013), have focused on the novel's depiction of engineered life forms and the ethical ramifications of biotechnological advancements. Kazuo Ishiguro's "Never Let Me Go" (2005) presents a dystopian reality where human clones are bred for organ harvesting. The novel follows the lives of these clones as they come to terms with their purpose and the limited scope of their existence. Ishiguro's poignant exploration of identity, personhood, and the moral

boundaries of scientific progress has been a focal point for transhumanist analysis. Scholars such as Agar (2007) have examined the ethical issues surrounding cloning and the treatment of sentient beings created for utilitarian purposes.

The examination of novels through a transhumanist lens offers profound insights into the ethical, philosophical, and social implications of technological advancements. These works contribute to the discourse on AI, cyborgs, and human-machine interactions. The narratives reflect contemporary anxieties and hopes about the future while challenging readers to consider the fundamental aspects of what it means to be human in an increasingly technologized world. As technology continues to advance, the themes explored in these novels remain relevant, urging ongoing reflection and dialogue.

Transhumanism, as a philosophical and cultural movement, has garnered significant attention across various academic disciplines and cultural spheres. Scholars have extensively examined the implications of transhumanist ideals on humanity, society, and ethics, drawing from a diverse range of literature.

Within philosophy, seminal works by thinkers such as Nick Bostrom, Max More, and Julian Huxley have laid the groundwork for exploring transhumanist themes. Bostrom's *Superintelligence: Paths, Dangers, Strategies*, delves into the potential risks and benefits of artificial superintelligence, sparking debates on existential risk and ethical considerations. More's *The Diachronic Self: Identity, Continuity, Transformation*, delves into the complexities of identity evolution over time, outlining the principles of transhumanist philosophy, emphasizing perpetual progress and enhancement. Julian Huxley's concept of "transhumanism" itself, coined in the mid-20th century, in his book *New Bottles for New Wine*, serves as a foundational text for understanding the movement's aspirations for transcending human limitations.

In the realm of technology studies, scholars have focused on the practical implications of transhumanist ideas. Ray Kurzweil's *The Singularity Is Near*, explores the convergence of technological advancements, predicting a future where human intelligence merges with artificial intelligence. Donna Haraway's *A Cyborg Manifesto*, challenges traditional notions of identity and embodiment, envisioning a post-human world where boundaries between human and machine blur.

Ethical inquiries into transhumanism are abundant, with scholars addressing issues of autonomy, equality, and justice. Michael J. Sandel, in *The Case Against Perfection: Ethics in the Age of Genetic Engineering*, explores the ethical dilemmas surrounding genetic manipulation and human enhancement. Sandel critiques the pursuit of human perfection through genetic engineering, highlighting concerns about fairness, equality, and the commodification of life.

Sociological analyses of transhumanism delve into the societal attitudes and responses towards technological augmentation. Steve Fuller's *Humanity 2.0: What It Means to Be Human Past, Present and Future*, explores the socio-political implications of human enhancement technologies, advocating for a more inclusive and participatory approach to shaping the future of humanity.

The research design and methodology for this study adopts an interdisciplinary approach, integrating philosophical, technological, cultural, ethical, and sociological perspectives to analyse transhumanist themes. It delves into fundamental questions surrounding the nature of humanity, identity, consciousness, and ethics within transhumanism while exploring advancements in AI, biotechnology, cyborgisation, and human enhancement. Cultural studies examine how societal norms shape the portrayal of transhumanist ideas in literature and media, addressing the ethical considerations including dilemmas related to autonomy, relationship dynamics, and authority.

To conduct the analysis, the study will employ a multi-dimensional approach, integrating insights from each discipline to interpret transhumanist themes within the selected novel. Utilising a content analysis method, the research will systematically examine textual elements within *Robopocalypse*, assessing how transhumanist concepts are depicted and interpreted within the narrative. Each theoretical framework will be applied to analyse different aspects of the novel, identifying the present ethical implications.

Furthermore, the research will involve a comprehensive review of existing literature, encompassing diverse scholarly perspectives and critical insights from relevant articles, books, and theoretical discourses. By engaging with this breadth of literature, the study aims to enrich its findings and contribute to a nuanced understanding of transhumanist narratives in literature and their broader societal implications.

This dissertation will be divided into two main chapters. The first chapter will focus on establishing the theoretical framework and examining the socio-cultural context of transhumanism. The first section will delve into technological aspects such as artificial intelligence, biohacking, cyborgisation, and human enhancement, analysing their development and implications within society. In the second section, philosophical inquiries into the nature of humanity, identity, and consciousness, along with ethical considerations and existential implications within the realm of transhumanist themes will be explored. Furthermore, the third section will investigate the cultural implications of transhumanism, including its portrayal in literature as well as media depiction, and it also provides tangible exemplars from our real world. By examining these dimensions, the chapter aims to provide a comprehensive understanding of the theoretical underpinnings and socio-historical context surrounding transhumanist discourse.

Chapter two will be dedicated to analysing the evolution of human-AI interaction and the ethical and existential questions raised by AI in Daniel H. Wilson's *Robopocalypse*. The

first section explores the progression from AI's initial triumph over humans to a more integrated relationship, with a focus on the dynamics of human-machine interaction. It examines how AI's dominance reshapes the narrative, transforming human roles and identities. Further, it delves into the concept of robo-human appropriation and the emergence of cyborgian beings, illustrating the profound impact of these changes on character development and societal structures. The second section shifts to the broader implications of AI on ethics and human existence. It investigates the potential of AI to assume a god-like role in society, questioning whether advanced machines are becoming humanity's new deities. This leads to an exploration of the ethical dilemmas posed by intelligent robots, probing into whether AI can embody moral values and act benevolently. The section concludes with a philosophical reflection on the nature of humanity in the face of rapid technological advancement, pondering the existential question: "To be or not to be?" Through this comprehensive analysis, the chapter offers nuanced insights into the complex interplay between technology, ethics, and human identity in *Robopocalypse*.

Chapter One: Theoretical Framework and Socio-Cultural Context

1.1 Introduction

In the era of posthumanism, characterized by the increasing impact of science and technology on human life, the traditional understanding of humanity, is subject to constant re-evaluation. The deepening integration of humans and technology prompts significant shifts in our perception of the human body, its interaction with technology, and its place in the environment. Consequently, as the lines between organic and artificial blur, humanity faces a critical juncture where the fundamental essence of human identity is challenged.

Throughout the annals of human history, technology has been an inseparable facet of our existence, intricately interwoven with our biological essence. The human body, in itself, can be perceived as a marvel of biological engineering, functioning with remarkable precision and coordination. As scientists delve deeper into understanding other technological intricacies within our bodies, it becomes clear that humans possess an innate capacity for advancement. This inherent potential fuels the optimism of transhumanists, who envisage a future where human transformation and transcendence are not only feasible but inevitable.

In this envisioned postbiological era, humans confront a technologically advanced environment that surpasses the capabilities of their biological nature to thrive effectively. Consequently, they perceive their organic bodies as obsolete and ill-suited to the demands of this new milieu. This realization precipitates a critical moment in human evolution, wherein the necessity to adapt and transcend their organic forms becomes imperative. Thus, humans stand at an evolutionary crossroads, compelled to upgrade, and surpass their biological bodies to ensure alignment with their increasingly technology-driven existence.

This chapter will provide a general outline of the theoretical framework and socio-cultural context surrounding transhumanism. The first section provides an explanation of how

technological aspects shape transhumanist discourse, focusing on Artificial Intelligence (AI), Cyborgisation, and Biohacking. Then, the second section delves into the philosophical examination of transhumanist themes, exploring personal identity and the nature of humanity, along with ethical considerations and existential implications. The third section explores the cultural implications of transhumanism, analysing its portrayal in literature and media, and paralleling their manifestation in our real world as tangible exemplars.

1.2 Technological Aspects of Transhumanism

1.2.1 AI: The Zeitgeist of The Modern Era

For thousands of years, humanity has endeavoured to unravel the mysteries of our cognitive faculties—how our remarkably complex brains, composed of mere matter, navigate, comprehend, and manipulate a world of immense complexity. This quest for understanding extends to the realm of AI, a field not only dedicated to comprehending the intricacies of human thought but also to crafting intelligent entities—machines capable of navigating diverse and novel scenarios with efficacy and safety. The Industrial Revolution of the nineteenth century witnessed the integration of machines into various spheres of human activity, leading to the mechanization of human physical labour, and driving forward significant scientific and technological progress. Similarly, the twentieth century heralded the rise of information technology, particularly with the emergence of computers, which spurred the rapid development of AI. This technological metamorphosis empowered machines to undertake or alleviate the routine burdens of human mental labour (Li and Du 1).

It is widely acknowledged that humans are regarded as the most intellectually advanced species on Earth. This designation stems from their remarkable cognitive abilities, including the capacity for critical thinking, logical reasoning, comprehension of complexity, and autonomous decision-making. These faculties, coupled with capacities for strategic planning, innovation, and problem-solving, underscore humanity’s unique position in the natural world

(Ghosh and Thirugnanam 23). Since its birth, AI has captivated the imaginations of people and fostered high expectations, serving as a compelling objective in interdisciplinary development. Research on neural networks (NNs) has been instrumental in advancing AI to unprecedented heights. Initially proposed in the early 1940s, the concept of NNs underwent periods of rejection and rediscovery. In 1982, John J. Hopfield introduced the idea of implementing hardware to actualize artificial neural networks (ANNs), a concept that laid the groundwork for the *Hopfield Neural Network*. Subsequently, in 1986, David E. Rumelhart and collaborators introduced the back propagation (BP) algorithm for multi-layered networks, marking a significant milestone in AI progress. During this era, there was anticipation that machines would process information akin to "thinking in images," seeking to replicate human visual perception, instinctual responses, and common sense through biological NN models, artificial neuron models, typical training algorithms, and excitation functions (Li and Du 3).

The fundamental objective of AI is to develop systems that are transparent, interpretable, and explainable, thereby enhancing their utility as intelligent agents. This goal stems from the concept of establishing trust in machines as counterparts to humans, which originated with the invention of the "Turing Test". In this test, machines are evaluated based on their ability to emulate human behaviour, regardless of the examiner's prior knowledge, and are considered intelligent if they pass the assessment. It is therefore unsurprising that AI has had far-reaching effects on various aspects of society, ushering in a new era in the digital revolution (Ghosh and Thirugnanam 24).

Artificial intelligence encompasses various types based on their capabilities. Narrow AI, the predominant focus of contemporary AI research, is characterized by its ability to execute a specific set of predefined instructions without possessing general thinking capabilities. This specialized form of AI develops programs tailored to excel in particular domains, such as chess-playing, medical diagnosis, autonomous driving, algebraic calculation, or mathematical theorem-proving. Notably, many of these narrow AI applications have achieved remarkable

success, making it the most extensively utilised type of AI worldwide. Prominent examples include Apple's *Siri*, Amazon's *Alexa*, Google's *AlphaGo*, IBM's Watson supercomputer, and *Sophia*, the humanoid robot—all falling within the category of weak AI. AGI, or Artificial General Intelligence, in contrast to Narrow AI, refers to a hypothetical AI system that possesses the ability to understand, learn, and apply knowledge in a manner similar to human intelligence across a wide range of tasks and domains. It aims to emulate the breadth of human intelligence and capabilities. Despite its potential, research on Artificial General Intelligence (AGI) has faced scepticism, with some likening the endeavour to constructing a perpetual motion machine—an inherently challenging and elusive goal (Ghosh and Thirugnanam 25). Presently, achieving AGI remains elusive, as no machines exist that can match the complexity and versatility of human thought. However, the possibility of AGI is not precluded by known physical laws; in fact, it is viewed as theoretically attainable by current scientific understanding. Similar to nanotechnology, AGI is regarded as "merely an engineering problem" rather than an insurmountable barrier, albeit one of significant complexity and difficulty (Goertzel and Pennachin 1). Another type is Strong AI which represents the pinnacle of artificial intelligence, envisioned to surpass human cognitive capabilities. This form of AI is anticipated to outperform humans in various tasks, potentially leading to a scenario where machines dominate and surpass human abilities. While achieving Strong AI presents formidable challenges, it is deemed feasible, though difficult. However, concerns have been raised by scientists, including Stephen Hawking, regarding the societal implications of Strong AI. Its development has sparked apprehension about the potential for machines to become dominant entities, posing existential threats to humanity (Ghosh and Thirugnanam 25).

The landscape of artificial intelligence is characterized by its diverse array of domains, each contributing to its multifaceted nature. Among these domains are neural networks, robotics, expert systems, fuzzy logic systems, and natural language processing (NLP). These domains collectively represent the breadth and depth of AI research and applications,

showcasing the field's versatility and potential for addressing complex challenges across various sectors of society. Neural networks emulate the human neural system, comprising layers and nodes that process data akin to neurons and dendrites. These networks employ algorithms to discern relationships within data, functioning similarly to the human brain. Widely utilised in AI through machine learning and deep learning, neural networks excel in tasks such as face and image recognition, particularly in medical diagnosis. Robotics is a domain of AI focused on creating intelligent machines, such as robots and humanoids, capable of executing tasks following human instructions.

The widespread adoption of robots across various sectors—including industry, medical surgery, and restaurants—reflects the growing appreciation for their utility and versatility. Expert systems leverage data stored in knowledge bases to make decisions, supplemented by expert guidance. These computer applications excel in solving complex problems, drawing upon intelligence and expertise to navigate intricate scenarios effectively. Fuzzy logic systems resemble human thinking method and decision-making processes, accommodating a spectrum of possibilities between 0 and 1 rather than binary choices. Widely applied in consumer electronics, automobiles, and data comparison, these systems offer nuanced decision-making capabilities akin to human cognition. Natural language processing (NLP) facilitates seamless communication between computers and human languages, bridging the gap through intelligent interaction. Examples include Google Translate and spell-check features, which exemplify NLP's role in enhancing communication efficiency and accuracy (Ghosh and Thirugnanam 26).

The realm of AI offers boundless opportunities for intellectual exploration. Thus, AI expert Kai-Fu Lee anticipates that its impact will be "more than anything in the history of mankind." Unlike disciplines like physics, which may appear to have exhausted their potential with historical luminaries such as Galileo and Einstein, AI remains ripe with potential for innovation. Spanning from broad concepts like learning, reasoning, and perception to specialised applications like chess-playing, theorem-proving, poetry generation, autonomous

driving, and medical diagnosis, AI encompasses a diverse array of subfields. Its relevance extends to virtually any intellectual pursuit, making it a truly universal field (Russell and Norvig).

1.2.2 Cyorgisation: The Transhumanist Promise

From the dawn of history, humanity has harboured a relentless desire to push the boundaries of what is possible—seeking to go further, achieve more, move faster, grow stronger, and expand our horizons in every conceivable way. While there have always been individuals content with the status quo, leading lives of routine and predictability, this does not negate humanity's innate inclination to strive for greater heights (Klichowski and Przybyla 165). Indeed, our relentless pursuit of progress has been evident since ancient times, from the Stone Age through the Bronze and Iron Ages, as Homo sapiens continuously sought to improve their circumstances by innovating and creating tools to enhance their existence. The evolution of technology has played a pivotal role in this quest for advancement (166). Mechanical innovations have extended our physical reach into space, while advancements in electrotechnics have facilitated global connectivity, expanding our collective nervous system beyond geographical boundaries. Throughout these epochs, humanity's drive for efficiency and improvement has remained constant, permeating every aspect of our lives.

The concept of cyborgisation represents a deliberate and conscious effort to enhance and augment human cognitive functions through the utilisation of available technological means. It embodies humanity's enduring aspiration to elevate itself to higher levels of development. The term 'cyborg' originates from 'cybernetic organism,' denoting an entity composed of both biological and technical elements. Initially, it encompassed systems of mixed composition, encompassing individuals who rode bicycles or wore glasses. However, its usage has evolved to specifically refer to entities where technology and biology are intricately intertwined. Throughout history, humans have envisioned merging man and machine to unlock

greater potential and intelligence, giving rise to the concept of the cyborg. In pursuit of this vision, scientists and technology giants have identified four key technologies: Extended Reality (XR), Artificial Intelligence (AI), Brain Machine Interface (BMI), and Internet of Things (IoT) (Dhanaraj et al. 36-37). These technologies are envisioned to synergistically integrate man and machine, heralding a new era of intuitive, immersive, natural, and interactive human-machine interaction. This convergence holds the promise of ushering humanity into a new era characterized by seamless integration and enhanced capabilities.

Each day, teams of experts explore increasingly advanced possibilities for supplementing, replacing, and augmenting human organs or even entire limbs with cutting-edge implants (Klichowski and Przybyla 167). Currently, a wide range of human organs or body parts can be replaced or enhanced using technological means. For instance, the human brain, serving as the command centre for the entire organism, is now equipped with electrodes or microchips that mitigate the effects of conditions like epilepsy, Parkinson's disease, and potentially Alzheimer's. Additionally, artificial eyes, ears, and larynxes are commonly implanted to enhance sensory capabilities. The enhancement of human physiology extends to internal organs such as the heart, kidneys, bladder, intestines, and blood vessels, as well as artificial bones, joints, muscles, skin, and spinal components (Klichowski and Przybyla 166).

The future of medicine lies in the integration of artificial and natural elements, supplementing human anatomy with technological components—a development encompassing bionics, nanotechnology, microchips, and the convergence of various devices. In contemporary medical practice, doctors and surgeons often rely on highly precise robotic assistance to achieve not only successful surgeries but also rapid or immediate recovery and full restoration of health with minimal harm. An exemplary system is the *Da Vinci* surgical system—a medical robot renowned for its precision surpassing that of human hands and eyes (Klichowski and Przybyla 167). Notably, the system's robotic arms, typically numbering three or four depending on the version, remain steady, fatigue-free, and entirely responsive to the operator's commands.

The expanding realm of cyborgisation manifests through two key avenues: physical enhancement and cognitive enhancement. First, Physical enhancement, which involves augmenting the human body with technology, exemplified by the development of artificial exoskeletons. Initially conceived as a means to address disabilities, such as those experienced by Oscar Pistorius, the concept gained traction as researchers recognized its potential advantages. Pistorius, a double leg amputee, achieved remarkable success in both Paralympic and able-bodied competitions, sparking discussions about the performance advantages afforded by carbon-fibre prosthetics (Klichowski and Przybyla 170). This realization led to the creation of artificial exoskeletons—prosthetic structures designed to enhance the capabilities of able-bodied individuals (Greguric 86). Originally employed to assist individuals with leg paresis, these exoskeletons are now utilised by various professionals, including soldiers, police officers, gardeners, and fruit farmers. Known examples like the HULC, developed by Lockheed Martin, enable wearers to perform tasks with remarkable ease, including walking, running, kneeling, jumping, and lifting heavy objects weighing up to several hundred kilos, all while preserving full mobility (Klichowski and Przybyla 171).

Second, Cognitive enhancement through cyborgisation, which is grounded in the evolution of the brain-machine interface, aims to bridge the biological and electronic realms (Klichowski and Przybyla 170). This interface, facilitated by brain implants, holds the promise of direct communication between neurons and computers, expanding mental representations by integrating augmented reality technology. Although the development of brain implants for augmented reality is ongoing worldwide, current applications like Google Glass and smartphone AR apps offer glimpses into its potential. In this context, Kurzweil suggests a future where reality is augmented with screens, providing real-time hints and explanations overlaid onto the physical world. Furthermore, nanotechnology, particularly nanorobots, is poised to revolutionize cognitive enhancement. Envisioned as a strategy that will be affordable in less than 25 years, for reinforcing cognition. These tiny robots are projected to traverse the

bloodstream, hold the potential to address various health challenges by targeting pathogens, correcting DNA errors, and eliminating toxins. Such advancements could theoretically enable prolonged youthfulness and longevity, offering the prospect of an endless life free from aging. In the realm of neuroscience, nanorobots are envisioned to interface with biological neurons, providing access to a full spectrum of virtual reality experiences and unlocking the neurological substrates of emotions. Kurzweil suggests that this convergence of nanotechnology with neuroscience may herald the advent of a fundamentally new brain, devoid of limitations or biological constraints, offering boundless possibilities for human cognition (Klichowski and Przybyla 171-173).

Significant advancements in artificial body parts have led to widespread cyborgisation, with many individuals today incorporating such technology into their daily lives. From false nails to artificial teeth and corrective eyewear, these enhancements have become commonplace (Grinin and Grinin 356). Additionally, biosensors and biochips, representing the convergence of medicine and nanotechnology, offer innovative solutions for physiological monitoring and intervention. Among these ground-breaking innovations are Elon Musk's new Brain Chip, introduced by his company *Neuralink*, on January 29th, 2024, which promises to revolutionize brain-computer interface technology. Implanted directly into the brain with ultra-thin electrodes, this breakthrough holds the potential to reshape how humans interact with machines, enabling seamless communication and control through thought alone.

As we stand on the brink of a post-human revolution, the transformative potential of this era looms large. While the envisioned future may not be as radical as some transhumanists anticipate, the prospect of extending life, replacing organs and cells with non-biological materials, and integrating electronic elements into the human body is undeniably significant. Throughout the latter half of the twentieth century, the impact of this revolution, particularly in the realm of information technologies, has reverberated globally. The Cybernetic Revolution, now considered the third major production revolution in human history after the Agrarian and

Industrial revolutions, continues to unfold, with its final phase yet to be realised. The impending changes, anticipated over the next six or seven decades, are poised to redefine our world. Central to this evolution is the development of cyborgisation, propelled by a dual focus on enhancing quality of life and advancing self-regulating systems and technologies. Medicine, additive and bio-nanotechnologies, robotics, ICT, and cognitive technologies converge to form a complex ecosystem of production and innovation (Grinin and Grinin 357).

Many scholars suggest that we stand on the precipice of fundamental shifts, with some even contemplating the notion of a technological singularity—a pivotal moment wherein the trajectory of technological progress undergoes a paradigm shift. This concept has given rise to the idea of transhumanism, which envisions a future where cyborgisation enables the transfer of consciousness into immortal, abiotic bodies. Immortality thus emerges as a central theme in transhumanist discourse, underscoring the profound implications of our ongoing journey into the post-human era.

1.2.3 Biohacking: Human Enhancement Through DIY Biology

Technological advancements, particularly the internet, have fostered globally connected communities and facilitated the development of health-related consumer technologies for monitoring health metrics from home. Recent strides in health technology, accelerated by the internet and the proliferation of personal computing devices like smartphones and wearables, have spurred the innovation of digital health monitoring options (Qasim 5). This surge in technological capabilities has given rise to biohacking, a trend gaining popularity. Biohacking involves the acquisition and engagement with health data and do-it-yourself (DIY) biological modification, often conducted outside formal medical settings, with knowledge drawn from self-identifying biohacker communities (Qasim 5-6). The term “biohacking” dates back to a 1988 Washington Post article by Michael Schrage, envisioning a future where life could be

manipulated akin to computer programming (Petersén 7). In the article, Schrage reflected upon a future biotech culture in the following way:

> What happens, for example, if future generations begin to see life as something that's manipulable – just another computer program, but one in which the printout isn't on paper but in proteins? If children grow up believing that life is nothing more than organic chemistry? (Petersén 7).

Since then, biotechnological advancements have realised aspects of this vision, transforming biohacking from a clandestine subculture into a mainstream discourse on citizen engagement with biomedical science and personal health responsibility. The movement gained momentum in the US around 2005 and became organized on a societal level in 2008 under the name "DIYbio" in Boston. Contemporary biohacking often focuses on optimizing personal health and longevity (Petersén 8). Defined as attempts to enhance body and mind using technology or substances like hormones, biohacking practices predominantly revolve around health-centric activities such as dietary supplements, herbal remedies, or chip implants for medical interactions.

Diverse types of implantable technologies are now commonly used, either commercially available or homemade, and are typically inserted into the body via hypodermic needles or surgical incision. These implants often feature a layer to reduce immune reactions, known as bioproofing. Some examples include Neodymium Magnets, inspired by magnetoreception, which allow individuals to sense electromagnetic forces through tactile sensation. These magnets, like those implanted into the tragus of biohacker Rich Lee's ears, enable the reception of audio signals from smartphones via an electromagnetic coil. Another type is RFID/NFC Chips, which include active RFID implants with embedded batteries for communication within a body area network or connection to IoT (Internet of Things) devices via Bluetooth (Yetisen 744). Formerly offered by Verichip Corporation, now PositiveID, these FDA-approved chips were designed to provide biometric information, aiding in the identification of unconscious patients in emergencies. Implantable light sources, featuring LEDs and rechargeable batteries,

serve cosmetic purposes and may soon offer gesture recognition and Bluetooth connectivity (Yetisen 745). Additionally, implantable sensors, like temperature sensors used by biohackers, can measure body temperature and other parameters in real-time, such as pressure and biomarkers. Grindhouse Wetware's *Circadia 1.0* is a notable example, transmitting temperature data to a tablet computer via Bluetooth from an implant in the forearm (Yetisen 745).

The core concept of transhumanism entails integrating technology into the human body to harness the benefits of technological advancements fully. Transhumanists perceive technologically augmented humans as the next step in evolution, often envisioning the eventual uploading of minds. They challenge the notion of the human body as separate from nature and technology, considering it a biological limitation hindering the conquest of aging and mortality, which they view as treatable conditions. Initially, transhumanism faced ridicule and criticism in the 1990s and early 2000s, either as a fringe sci-fi movement or a potentially divisive ideology. However, its status began to evolve around 2010, gaining traction in bioethical discourse and scholarly journals. This shift reflects advancements in biotechnology, including synthetic biology and CRISPR gene editing, making the transhumanist vision more plausible (Petersén 14). Additionally, the increasing embrace of transhumanism is undoubtedly influenced by modern dataism and the prevailing perception of biological bodies as data systems.

The convergence of biotechnological advancements, dataism, and the discourse promoted by major tech companies, often influenced by technolibertarian ideology, blurs the boundaries of transhumanism, making it challenging to define its scope in today's context (Petersén 14). The evolution of transhumanist discourse is also influenced by the emergence of biohacking, a newer concept with fewer ethical concerns. In Sweden, a stronghold of transhumanist thought since the 1990s, the movement embraced biohacking as a more socially palatable alternative, aligning with contemporary trends. Consequently, transhumanism found

renewed appeal under the guise of biohacking, adapting to cultural shifts and technological advancements.

The concept of biohacking, inspired by transhumanism's push for human enhancement through technology, has evolved over time, starting with pioneering projects in self-experimentation and medical implants. Kevin Warwick's groundbreaking *Cyborg 1.0* project in 1998, where he implanted an RFID tag to control electronic devices, marked the beginning of this movement. Over the years, a vibrant community of technology enthusiasts, known as grinders, emerged, including figures like Amal Graafstra, Tim Cannon, Lepht Anonym, and Neil Harbisson, who experimented with implanting chips for various purposes such as human-device communication and self-quantification. Warwick's later experiment involved implanting a multielectrode array in his arm to create a neural interface, demonstrating the potential for human augmentation through technology. Meanwhile, artists like Stelarc popularized self-experimentation with biomaterials through performance art, further blurring the lines between technology and biology (Yetisen 744).

In recent years, biohacking has expanded to include more sophisticated endeavours like brain-computer interfaces. Companies like *Neuralink*, co-founded by Elon Musk, aim to a "merger of biological intelligence and machine intelligence." Musk envisions this technology as a means to ensure that artificial intelligence aligns with human interests, allowing direct interface to steer AI advancements towards enhancing rather than conflicting with our goals. Despite significant claims and even FDA approval for clinical trials, *Neuralink* has yet to deliver tangible results. Musk has pledged to utilise the technology to address brain and spinal injuries, aiming to compensate for lost capacities with implantable chips (Steele). However, the company's most prominent demonstration in 2021 involved a monkey playing the computer game 'Pong' using its mind, leaving much to be desired in terms of practical applications. Other bionic implants have gained attention not due to significant technological leaps, but rather because they appear peculiar and unimpressive. For instance, *Biohax International* has

implanted ID ‘biochips’ in a few thousand individuals, offering benefits such as never worrying about losing keys or enabling contactless payments. While some may dismiss the idea of embedding an ID chip in the hand as eccentric or risky, it is important to recognize that human-technology interfaces are already a part of daily life (Steele). Microchip-enabled pacemakers regulate heartbeats, while cochlear implants connect an external microphone to the brain, restoring hearing for some deaf individuals. Additionally, implantable glucose monitors are emerging as a more comfortable alternative for diabetics compared to traditional finger-prick tests or external CGM devices (Steele).

The likelihood of the public turning to biohacking for health management is increasing, underscoring the need to address gaps in our understanding. Given the wide-ranging nature of biohacking practices, it is essential for healthcare professionals to grasp its fundamentals to identify potential risks and address them effectively (Qasim 6). Biohackers are driven by various motivations, including exploring cybernetics, acquiring personal data, and advocating for privacy rights and open-source medicine. The emergence of a biohacking community has sparked discussions on cultural values, medical ethics, safety, and consent in transhumanist technology. Healthcare technologies such as epidermal electronics, biosensors, and artificial intelligence have converged to monitor patients in point-of-care settings within the Internet of Things (IoT), attracting hobbyist software developers involved in the quantified-self movement. Self-experimentation enthusiasts are keen on tracking their daily physical and biochemical activities to compile personal informatics for maintaining a healthy lifestyle or enhancing bodily performance. The growing interest in this “tech-savvy” community has prompted exploration into experimenting with implantable technologies, influenced by the use of biometric animal identification implants (Yetisen 744).

Safety concerns arise as medical devices must be sterilized before implantation to eliminate pathogens, and the exchange of implants among biohackers poses a risk of microbial transmission. Privacy issues arise from real-time geotagging of individuals’ locations and

activities, with discussions within the biohacking community on the mass surveillance implications of implantable devices. Ethical and legal challenges associated with electronic implants have been addressed by the European Group on Ethics in Science and New Technologies, highlighting concerns about personal information access by state authorities for security reasons. Some states in the US have enacted laws protecting citizens from involuntary or incentivized chip implantation (Yetisen 745-746). The rise of biohacking as a form of citizen science raises questions about the boundaries of medical data privacy and introduces the possibility of using cryptography for medical data storage, shaping the landscape of open medicine and challenging traditional notions of what it means to be human (Yetisen 746).

As we peer into the future, the trajectory of biohacking presents us with compelling questions regarding the evolution of healthcare and personal enhancement. Could we soon see implantable devices that not only monitor but also diagnose diseases or offer tailored health advice? As these possibilities unfold, the boundary between biohacking and conventional medical progress may become more nuanced. It reflects a convergence of human ingenuity and technological advancement, ushering in a new era where the lines between biological enhancement and medical innovation blur, ultimately shaping the landscape of human potential and well-being.

1.3 Philosophical Examination of Transhumanist Themes

1.3.1 Personal Identity and the Nature of Humanity

From ancient philosophical debates about the soul to contemporary discussions on neuroscientific reductionism, lies the age-old question: What defines who we are as individuals, and how does this understanding shape our perception of what it means to be human? Presently, neuroscientific reductionism has played a role in discarding an essentialist view of individual identity. Despite this shift, many transhumanists maintain an essentialist perspective on personal identity while advocating for extensive cognitive enhancements that would alter

various aspects of consciousness. It is imperative for transhumanists to confront the implications of their endeavours and ethical frameworks if personal identity is considered a flexible and subjective construct. Oxford's transhumanist philosopher Nick Bostrom acknowledged the problem of personal identity for transhumanism in the 2003 Transhumanist FAQ:

> Many philosophers who have studied the problem think that at least under some conditions, an upload of your brain would be you. A widely accepted position is that you survive so long as certain information patterns are conserved, such as your memories, values, attitudes, and emotional dispositions, and so long as there is causal continuity so that earlier stages of yourself help determine later stages of yourself ("*Humanity+*").

Similarly, in her 2009 essay titled *Future Minds: Transhumanism, Cognitive Enhancement and the Nature of Persons*, Susan Schneider references Ray Kurzweil's 2005 breakdown of the personal identity discourse into four main positions: First, "The ego theory" which suggests that a person's essence resides in their soul or nonphysical mind, capable of persisting beyond bodily death. Second, "The psychological continuity theory" defining identity based on memories, self-reflection, in line with Locke's philosophy, and overall psychological makeup, often termed as their "pattern." Third, "Materialism" asserting identity rooted in the physical matter comprising the body and brain, as termed by Kurzweil as "the ordered and chaotic collection of molecules." And fourth, "The no self-view" rejecting a metaphysical self and proposing that the concept of "I" is a linguistic construct, as proposed by Nietzsche, or a bundle of impressions without an underlying self, as argued by Hume, suggesting no continuity of identity due to the absence of inherent personhood, paralleling the teachings of Buddha and Parfit (Hughes 3).

In *The Singularity is Near* (2005), Kurzweil advocates for the second position, known as "Patternism," which is widely embraced among transhumanists. Patternism allows for significant alterations to the body and brain as long as there is a sense of continuity, marked by

the memory of a sequence of mental states leading to the present. According to this view, even extreme transformations such as transferring a personality from a biological brain to a computer would maintain personal identity if the digital mind retained memories of the process and identified with the original biological person (Hughes 4). One prominent advocate of this variant of Patternism is Max More, the founder of the Extropian school of Transhumanist thought. More extensively explores this perspective, particularly in relation to Derek Parfit's arguments on personal identity and their implications for radical human enhancement, as evidenced by his doctoral thesis (Hughes 4). While More acknowledges that certain enhancements could disrupt personal identity by interrupting the continuity of the personality pattern, he contends, contrary to many bioconservative critics, that personal continuity remains achievable. However, Schneider challenges the adequacy of transhumanists' Patternist theories in establishing the continuity of personal identity following radical cognitive enhancements or uploading.

Most transhumanists generally do not view the proliferation of selves as a significant issue, but it often clashes with the presumed cohesive unity of identity across time. If there could potentially be multiple versions of oneself, it raises questions about the very existence of the original self. She concludes by posing the question:

> …what is it that ultimately grounds your decision to enhance or not enhance if not that it will somehow improve who you are? Are you perhaps merely planning for the well-being of your closest continuent? (Schneider 11).

Various responses to the erosion of personal identity can be envisioned, ranging from attempts to establish secure personal stability using neurotechnologies to replacing individual identity with a fully collective identity, such as seen in the concept of "the Borg." This erosion may occur without coercion; selectively suppressing brain mechanisms responsible for the sense of self, like proprioception, could become attractive for those exploring neurotechnologies for therapeutic or recreational purposes (Hughes 4). Nick Bostrom anticipates such developments

in his 2001 essay *Existential Risks.* Beyond addressing natural and technological threats to humanity's survival, he considers scenarios where our descendants persist but in forms divergent from our current identity. In essence, if we embrace a theory of personal identity concerning the human endeavour, then certain post-personal identity societies might deviate from representing the human endeavour as we understand it today. Moreover, the critique of identity essentialism may be even more applicable at the societal level than for individuals. If there is no real self and no real humanity then we are left with the question of whether we want to collectively pretend that we do exist, and if so, to what ends? (Hughes 6).

Certain philosophical and theological traditions emphasize the importance of elements like consciousness, mind, or spirit in defining human identity. According to these traditions, these aspects of human existence can be conveyed through different material means (Fuller 2). The argument proposes that partially siliconized cyborgs or entirely digital androids might effectively take over the role of transmitting and fostering the distinctive aspects of humanity, potentially exceeding the constraints of human biology. Noteworthy figures include Norbert Wiener, Herbert Simon, and Ray Kurzweil, collectively propose a vision of "Humanity 2.0" where the successor to Homo sapiens may not necessarily be a primate. Whether through genetic evolution or digital transformation, the divergence in our future trajectory underscores the profound cultural dichotomy encapsulated in the concept of "Humanity 2.0" (Fuller 2). From both scientific and political perspectives, there is a pressing question: what aspects should we safeguard as uniquely human?

One ambitious approach to shaping the future of humanity revolves around "converging technologies" (CT). This involves blending cutting-edge research across nanotechnology, biotechnology, information technology, and cognitive science to enhance human control over both their bodies and environments. CT, with its diverse interpretations, has become a focal point in the long-term science policies of major nations. There are six variants of Converging Technologies (CT) each projecting a different sense of what it means to be 'human' (Fuller 2).

Firstly, ‘Humanity Transcended’ suggests a future where humans engineer evolution, potentially altering the distribution of traits in populations. Secondly, ‘Humanity Enhanced’ focuses on improving human abilities without radical changes, though it may redefine ‘normal’ performance standards. ‘Humanity Prolonged’ aims to extend human lifespan, challenging notions of mortality and raising ethical questions. ‘Humanity Translated’ explores the idea of transferring human consciousness to silicon-based vehicles or virtual avatars, blurring the lines between physical and digital existence. ‘Humanity Incorporated’ merges human and non-human elements, emphasizing humanity’s ability to integrate the environment into itself. Lastly, ‘Humanity Tested’ acknowledges the speculative nature of CT, highlighting the need for a culture tolerant of experimentation and risk-taking. Regardless of the outcome of these six projected human futures, they all reflect humanist principles, albeit with varying degrees of emphasis on ideology (Fuller 103-105).

In the early 21st century, as advancements in various fields progress, the traditional mind-body dilemma undergoes redefinition. While some advocate for maintaining human essence within organic bodies, others foresee transitioning humanity into more resilient silicon-based forms.

> It is difficult, for me at least, to see any powerful principled reasons to remain human if we can create creatures, or evolve into creatures, fundamentally ‘better’ than ourselves. It is salutary to remember that we humans are the products of an evolutionary process that has fundamentally changed ‘our’ nature (Harris 40).

Harris reflects on the potential future where humans could create or evolve into beings that surpass our current capabilities. Additionally, the quote emphasizes that humans are the result of evolution, which has continuously transformed our nature over time. Therefore, it suggests that the notion of remaining human may become less significant if we can enhance ourselves or evolve into something greater. Those opposing Humanity 2.0 on principle grounds should adopt a stance akin to historical resistors of technological advancements, arguing that unregulated innovation may worsen societal inequalities. The core concept is to guide cutting-

edge research in technosciences, ensuring convergence into a unified science facilitating the transition to Humanity 2.0.

1.3.2 Ethical Considerations in The Age of Transhumanism

The rise of transhumanism appears inevitable, with technological advancements propelling humanity toward this transformative future. As Heidegger aptly observes, technological determinism rules, rather than social determinism. In other words, the technology controls the society, not the other way around. Thus, it is imperative to address these challenges today, as tomorrow may prove too late (Karaman 98). With each passing year, the once-fantastical notion of enhanced humans becomes increasingly tangible, presenting both tantalizing opportunities and daunting ethical quandaries. As humans undergo a synthesis of biological and non-biological elements, questions arise regarding their continued humanity or transformation into a new species; thus, the very essence of humanity hangs in the balance. What fate awaits those who opt out of enhancement, clinging to their classical humanity amidst a world dominated by transhumans? Will they find acceptance, employment, and sustenance in a society shaped by superhuman norms? The prospect of a divided humanity, split between the enhanced and the un-enhanced, raises concerns of social discord, inequality, and conflict. Will classical humans persist in the face of this transformative wave, or will they be swept aside by the tide of progress, compelled to embrace a new paradigm of existence? (Karaman 99). Consequently, a myriad of ethical quandaries will surface, ranging from fundamental questions about the nature of happiness to the complexities surrounding aging and mortality. Issues surrounding mind uploading, safety protocols, fairness in access to enhancements, and the preservation of individual freedoms in a society shaped by human enhancements will demand urgent attention.

Emerging technologies promise to facilitate a gradual evolution of humans into entities with capacities surpassing conventional definitions of “human,” ushering in the “posthuman

age" (Tirosh-Samuelson 1,3). However, evolutionary psychology challenges this assertion due to the specialized evolution of the human brain for specific tasks and our limited understanding of its operations. Tooby emphasizes the importance of questioning the goal of technological change and warns against conflating "evolution" with "progress." Similarly, Leda Cosmides advises against interfering with brain function to avoid unintended consequences (Tirosh-Samuelson 3). Max More envisions technological advancements, including "social technologies," enabling the transformation of human nature across physical, emotional, and intellectual domains. He anticipates extended lifespans in optimal health, enhanced self-awareness, mitigation of biases, heightened intelligence, and adaptability to change, ultimately leading to human happiness.

The quest for happiness has long been central to human concerns, deeply rooted in Western philosophical tradition since ancient times. However, the transhumanist approach to happiness poses significant challenges. While transhumanists emphasize self-fulfilment, they often overlook the importance of virtues and fail to demonstrate how the values of the authentic Self contribute to human flourishing. Despite discussing life satisfaction and self-realisation, transhumanists have yet to analyse the relationship between subjective and objective aspects of happiness rigorously. Furthermore, their belief that technology can continuously induce pleasant sensations raises ethical concerns (Tirosh-Samuelson 6). Advances in brain scanning and neural implants hold promise for manipulating emotions, but reducing the mind to brain functions oversimplifies human cognition. The brain is part of a complex organism, and understanding human happiness requires consideration of the nervous and immune systems, as well as the socio-cultural context. Transhumanists often liken the mind to a computer, but this analogy falls short in capturing the complexity of the human brain. While the brain does process information, it is not merely a computational machine. Additionally, understanding the brain's function requires consideration of the socio-cultural context in which individuals exist. Thus,

reducing the mind to brain functions overlooks the multifaceted nature of human cognition and experience (Tirosh-Samuelson 6).

Another key focus of transhumanism revolves around combating aging and striving to prolong human life. Aubrey de Grey, a prominent figure in the transhumanist movement, firmly believes that scientists will eventually discover methods to defeat aging altogether. This "crusade" against aging is rooted in scientific endeavours aimed at extending healthy lifespan and reducing the prevalence of age-related ailments. De Grey spearheads biomedical gerontological research through his Strategies for Engineered Negligible Senescence (SENS) initiative, anticipating major breakthroughs in this field. While De Grey's intentions are undoubtedly noble and the aspiration to extend human lifespan by 150 years is ostensibly beneficial, it raises questions about the underlying assumption of treating the human body as a resilient machine in need of prolonged maintenance (Tirosh-Samuelson 6). While some aspects of human physiology may resemble machinery, humans are fundamentally organisms subject to aging and mortality due to their inherent life processes. This crusade against aging essentially challenges the inevitability of death. While De Grey acknowledges the reality of death, his aim is to delay it as much as possible. However, the purpose and implications of indefinitely extending human life remain ambiguous. What activities and pursuits will occupy individuals over a lifespan of 150 or even 500 years? Will it lead to more consumption, entertainment, conflicts, or environmental degradation? (Tirosh-Samuelson 6). These questions underscore the ethical and philosophical considerations surrounding the quest for immortality and its implications for the future of humanity.

Another radical notion within transhumanism is the prospect of transferring the contents of human brains, essentially their minds, into non-biological entities to attain immortality. Visionaries like Kurzweil envision a scenario dubbed the "brain-porting scenario," where human brains are meticulously scanned to capture all essential details, then reinstated in a different, likely more advanced computational substrate. In this envisioned future, individuals

will retain human bodies but with the ability to morph them according to their intelligence. Kurzweil anticipates that these "software-based humans" will transcend the limitations of conventional humanity, existing predominantly in digital realms. They will inhabit the web, projecting bodies whenever necessary, including virtual forms in diverse virtual reality environments, holographic projections, and physical bodies composed of nanobot swarms and other nanotechnologies. For Kurzweil, this marks a pathway to immortality (Tirosh-Samuelson 6-7).

Although discussions surrounding the ethics of human enhancement have evolved over the past decade, *Beyond Therapy: Biotechnology and the Pursuit of Happiness*, released by the President's Council on Bioethics (PCB) in 2003, remains a significant source for ethical considerations. Like any health-related intervention, safety emerges as a primary concern in enhancement endeavours. The PCB underscores that individuals seeking to optimize their minds and bodies through substances like steroids, Ritalin, or mood enhancers expose themselves to unforeseen short- and long-term risks (Tennison 406). Additionally, human enhancement raises the spectre of exacerbating social inequalities. Just as steroid use can provide athletes with an unfair advantage in sports, many enhancement technologies aim to confer positional advantages. For instance, individuals equipped with neural implants for enhanced cognition may outperform unenhanced candidates in job interviews, while those with super-human sensory capabilities could violate privacy norms by eavesdropping or spying without detection. Moreover, the widespread use of Ritalin among students and professors may skew admissions and tenure decisions at top universities, limiting opportunities for others. Conversely, enhancement holds the potential to address the inherent inequities generated by biological chance, such as the unfairness of disability, effectively mitigating disparities on an already tilted playing field (Tennison 406).

The issue of unfairness largely revolves around equality concerns. Francis Fukuyama observes that throughout U.S. history, significant political battles have centred on defining who

qualifies as fully human, with marginalized communities often lacking political representation and certain rights afforded to others. The positional advantages granted to the enhanced could serve as leverage for asserting additional rights or influence. Moreover, "if, as is often the case with expensive medical care, only the wealthy and privileged will be able to gain easy access to costly enhancing technologies, we might expect to see an ever-widening gap between 'the best and the brightest' and the rest" (Tennison 406). The President's Council on Bioethics (PCB) also prompts inquiries into whether enhancement technologies might impede our freedom to autonomously pursue our own objectives.

There is concern that individuals may feel compelled to enhance themselves to maintain competitiveness in academic or professional settings. Likewise, parents may sense pressure to enhance their embryos or children to avoid being perceived as neglectful. This scenario risks limiting the range of acceptable standards of living and development, leading to a homogenized state of subtly coerced conformity. For example, Robert Sparrow argues that the utilitarian perspective requires parents to engineer their children according to the prevailing bigotries of the day: If the child is born into a racist culture that demonizes dark-skinned people, then the parents would have an obligation to select for lighter skin in order to maximize the child's chance of happiness (Tennison 406).

The emergence of the transhumanist vision stems from the convergence of knowledge in scientific domains like genetics, robotics, and nanotechnology, challenging traditional disciplinary boundaries. Scholars across disciplines must engage with these new fields, while transhumanism prompts critical reflection on human evolution as a design project. With diverse modalities for future human enhancement, including athletic performance and human-machine interfacing, ethical concerns arise, stimulating dialogue on moral enhancement. Envisioning a future where biotechnology extends life spans and enhances capacities, transhumanists anticipate a world liberated from human nature, fostering greater happiness and overall improvement (Thompson 1). Yet, discerning between enhancements aligning with values and

those compromising them is crucial, guiding ongoing dialogue to identify affected values and negotiate their significance. This dialogue reflects a broader discourse on the ethical implications of scientific progress, emphasizing the need to ensure that advancements in biotechnology and other fields uphold fundamental human values and promote the well-being of individuals and society as a whole.

1.3.3 AI's Existential Implications

AI and emerging technologies present significant hazards, prompting transhumanists to focus on existential risks associated with modern technology, rather than solely on their potential advantages. Nyholm echoes this concern in his book, published in 2020, *Humans and Robots: Ethics, Agency, and Anthropomorphism*, where he explores the implications of robots and advanced AI. In the book's initial chapter, he questions whether humanity, with its psychology rooted in evolutionary history, is prepared for a future increasingly dominated by these technologies. He says: "with our human psychology based on human evolution, are we ready for a world increasingly filled with robots and advanced AI?" Many individuals lack a comprehensive understanding of the intricate workings of modern technologies, often underestimating certain capabilities while overestimating others. Consequently, addressing this disparity requires either implementing a multitude of precautions currently overlooked in the development of safe robot and AI technologies or enhancing our ability to interact effectively with these technologies. Neglecting to consider the compatibility of human psychology with advanced robots and AI systems poses significant risks, including existential threats, as highlighted by Nyholm (Nyholm 80). It is essential to recognize that existential risks stemming from AI may not necessarily originate from super-intelligent systems, which might try to "take over," as depicted by Bostrom in his book *Superintelligence*. Even AI systems of lesser intelligence can pose grave threats due to factors beyond their super-intelligence potential (Nyholm 80).

The notion that current and imminent AI technologies could pose existential risks as intermediate factors is gaining traction, suggesting that this risk potential extends beyond scenarios involving unaligned AGI. Furthermore, forthcoming developments in the next decade hold the prospect of exacerbating these risk factors significantly. As far back as the 1950s, prominent figures like Albert Einstein and Bertrand Russell were warning about the peril of human extinction stemming from the use of nuclear weapons. More recently, the exploration of risks that could imperil the continued existence of our species has burgeoned into an academic discipline following Nick Bostrom's introduction of the concept of existential risks (Bucknall and Dori-Hacohen 119). Misaligned AI stands out as one of the most deliberated risks within the existential risk research community, with many proposed scenarios hinging on the premise of achieving at least 'human-level' AGI, if not surpassing superintelligence outright. Concurrently, considerable attention and concern are directed toward the nearer-term adversities posed by contemporary AI systems, with scrutiny coming from fields such as digital humanities and computer science, spearheaded by the Fairness, Accountability, Transparency, and Ethics (FATE) community (Bucknall and Dori-Hacohen 119).

The academic exploration of existential risks facing humanity is often traced back to Nick Bostrom's seminal article, *Existential Risks: Analysing Human Extinction Scenarios and Related Hazards*, published in 2002. In this influential work, Bostrom defines existential risk as a scenario where an adverse outcome would either eradicate Earth-originating intelligent life or severely and permanently limit its potential. He introduces a two-dimensional framework for classifying risks, considering both the scope of individuals affected and the severity of impact on each affected individual. Bostrom characterizes existential risks as those that are pan-generational, impacting all or nearly all living individuals as well as future generations, and describes them as "crushing," signifying the utmost severity in his classification scheme.

In the wake of Bostrom's pioneering work, the exploration of existential risks gained significant traction, leading to the publication of several notable books on the subject. These

include Martin Rees' *Our Final Century?* Richard Posner's *Catastrophe*, and Nick Bostrom and Milan Ćirković's *Global Catastrophic Risks*. More recently, Toby Ord's 2020 book *The Precipice: Existential Risk and the Future of Humanity* serves as a comprehensive overview of research conducted on existential risks since Bostrom's groundbreaking paper. Bostrom's framework extends beyond the mere extinction of the human species, encompassing scenarios involving technological or social stagnation, irrecoverable societal collapses, or dystopian outcomes, all of which pose significant threats to the realization of humanity's full potential (Bucknall and Dori-Hacohen 120).

In their comprehensive analysis of existential risk factors, Liu et al. underscore the importance of broadening our comprehension of the diverse pathways that could culminate in existential peril. Among their insights, they illustrate how even narrow AI systems could be sufficient to pose existential threats (Bucknall and Dori-Hacohen 121). Moving from the examination of the tangible effects of present and forthcoming AI technologies, attention now turns to evaluating how these effects act as triggers for risk. This transition involves a thorough exploration of each frequently analysed source of risk, encompassing nuclear weaponry, pandemics and biotechnology, climate change, and unaligned AGI.

Focusing on specific sources of risk, the examination begins with the implications of near-term AI on nuclear weaponry. In the context of an AI arms race between nuclear states, escalating competitive dynamics could lead to a heightened risk of a first-strike nuclear attack. Moreover, the evolving landscape of international cybersecurity could exacerbate this risk, particularly if a state gains superior intelligence capabilities or increased access to its adversary's systems. The Stockholm International Peace Research Institute's three-volume report, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, offers an extensive analysis and further insights into AI's influence on nuclear security (Bucknall and Dori-Hacohen 124).

The potential outbreak of a deadly infectious virus, whether naturally occurring or engineered, stands as a significant existential risk. Near-term AI trends introduce several risk factors concerning pandemics. Firstly, as we have seen from the COVID-19 pandemic, effective response to such a scenario demands not only capable political bodies but also a cooperative public willing to adhere to preventive measures. In the context of engineered pandemics, the accessibility of powerful AI systems may facilitate malevolent actors' acquisition of the technology and expertise needed to create dangerous pathogens for pandemic purposes. While speculative, recent advancements highlight AI's efficacy in addressing challenges in molecular biology, particularly protein folding, suggesting that similarly potent neural architectures, if misused, could yield catastrophic consequences. Lastly, escalating international political tensions and state rivalries may incentivize certain states to develop and potentially deploy advanced bioweapons, further heightening the risk landscape (Bucknall and Dori-Hacohen 124-125).

Similar to pandemics, climate change discourse is marred by misinformation and disinformation, often propagated by entities contributing to environmental harm. The energy-intensive process of developing and training AI systems directly contributes to environmental degradation, particularly through substantial energy consumption and potential secondary effects on industries like oil and gas extraction. Estimates suggest that the technology sector accounted for 3% to 3.6% of global greenhouse gas emissions in 2020, with a single natural language processing AI model training session generating 300,000 kilograms of carbon dioxide emissions—equivalent to 125 round-trip flights between New York and Beijing. Addressing this overlooked aspect of AI growth is essential for sustainable technological advancement and averting climate disaster (Bucknall and Dori-Hacohen 125).

The trajectory of near-term AI developments significantly influences the future landscape of AGI. Both successes and failures in current AI systems can shape the architecture of potential AGI, the societal context of its development, and our ability to align it with human

values. For instance, if an AI arms race ensues among states or technology firms, there may be pressures to compromise safety protocols to maximize system capabilities. Continued disregard for safety measures during AGI development could lead to catastrophic consequences upon its deployment (Bucknall and Dori-Hacohen 125).

The societal and political dimensions entwined with contemporary AI systems possess profound implications for humanity, extending beyond the mere emergence of unaligned AGI. It is evident that these systems, with their attendant societal and political ramifications, harbour the potential to serve as existential risk factors. While much attention has been rightfully directed towards the risks associated with the advent of unaligned AGI, it is crucial to recognize that the short-term consequences of existing AI systems are equally significant (Bucknall and Dori-Hacohen 127). These immediate harms possess the capacity to not only exacerbate present societal challenges but also amplify and complicate other existential risks confronting humanity. Thus, the multifaceted impacts of contemporary AI systems underscore the imperative of scrutinizing their societal and political dimensions to mitigate the broader existential threats they may engender.

1.4 Cultural Implications of Transhumanism

1.4.1 Transhumanism's Portrayal in Literature

Before transhumanism emerged as a formal philosophy, authors in literature were captivated by the idea of technologically enhanced humans. Across time, they portrayed these beings in various ways: some depicted them as monstrous, others as evolutionary mutations, and as the 20th century approached its end, they began envisioning them as artificial entities. For instance, the term "posthuman" can be traced back to H.P. Lovecraft's novella *The Shadow Out of Time*, published in 1936. Another example is the word "Robot," coined by Czech playwright Karel Čapek in 1921. In his grim play *R.U.R.*, humans create a labour force of robots, only to confront the peril of their own creations turning against them (Van der Horst

28). This trajectory of literary exploration laid the groundwork for the later development of transhumanist philosophy.

More than two centuries ago, technologically enhanced humans made their debut in the realm of science fiction, finding a place within the gothic narratives penned by Mary Shelley, Nathaniel Hawthorne, and Edgar Allan Poe. These authors were influenced by society's intrigue with electromagnetism and recent advancements in pharmaceuticals and prosthetics. Their tales often featured deranged scientists striving to enhance themselves and others, seeking to eradicate mental and physical imperfections in the process. Shelley's *Frankenstein* is often regarded as the cornerstone of science fiction, delineating the boundaries of scientific comprehension, and exploring the complex relationship between humans and their created beings (Van der Horst 28). Additionally, other precursors to science fiction emerged in the 19th century, including Edgar Allan Poe's *The Unparalleled Adventure of One Hans Pfaall* (1835), Jules Verne's *Journey to the Centre of the Earth*"(1864) and "*rom Earth to the Moon* (1865), and H.G. Wells' works such as *The Time Machine* (1895), *The Island of Doctor Moreau* (1896), *The Invisible Man* (1897), and *The War of the Worlds* (1898), all of which are considered among the earliest science fiction works (Mirenayat et al. 264).

Between the late 19th century and World War II, science fiction solidified as a distinct genre, with a primary focus on human evolution. Olaf Stapledon's *The Last and First Men* (1930) chronicles humanity's ebb and flow over vast epochs through the lens of creative evolution. Concurrently, American SF tales from this era depicted optimistic views of planned evolution, illustrating aspirations for unlimited improvement attainable within a relatively short span of time. Exploring transformation via technology, authors like Edward Bellamy with *Dr. Heidenhoff's Process* (1880) introduced a mechanical method erasing distressing memories, while Robert Louis Stevenson's *The Strange Case of Dr. Jekyll and Mr. Hyde* (1886) examined personality shifts via a medicinal serum. H.G. Wells' *The Island of Doctor Moreau* (1896)

delves into ethical quandaries as a scientist engineers human-animal hybrids, delving deep into issues of morality and self-identity (Van der Horst 28-29).

As the 20th century neared its end and the information age emerged, advancements in genetics, nanotechnology, AI, and computer science inspired a new wave of science fiction writers, known as "cyberpunks." These authors delved into themes of artificial beings and the anticipated Singularity. However, much of this literature portrayed transhuman and posthuman entities in a negative light (Van der Horst 29-30). After the Second World War, From the 1960s onward, science fiction delved into themes of human transformation through avenues like human cloning, genetic engineering, artificial intelligence, and cyborg bodies. During this period, A.E. Vogt's *Slan* (1946) envisioned a future where humans dominate higher-than-human mutants, showcasing the evolution of humans into superior beings. Just four years later, Isaac Asimov's *I, Robot* (1950) explored the notion of self-aware robots and their quest for legal equality with humans. In the same decade, Arthur C. Clarke's *Childhood's End* (1953) imagined human evolution towards superhuman capabilities. Theodore Sturgeon's *More Than Human* (1953) depicted individuals with extraordinary powers merging to form a superhuman entity, probing questions of power and morality.

Transitioning to the 1970s, Martine Caidin's *Cyborg* (1970) influenced many SF narratives with its tale of a man transformed into a deadly, emotionless weapon through technology. Moving forward, Frederik Pohl's *Man Plus* (1976) portrayed biological engineering turning a man into a machine-like entity for survival on Mars, reflecting early ideas of cyborgisation for extraterrestrial adaptation. Advancing to the 1990s, Marge Piercy's *He, She and It* (1993) featured an extraordinary cyborg man designed for combat. Lastly, Linda Nagata's *The Bohr Maker* (1995) explored the pursuit of immortality through illegal genetic manipulation to become posthuman. These narratives illustrate the intertwining of humans and machines, echoing concerns about the potential threats posed by intelligent machines, a process akin to the cyborgisation depicted in transhumanist SF tales (Mirenayat et al. 265-266).

The SF works discussed above vividly portray the fusion of advanced technology with human life, showcasing a diverse array of transformations that reflect readers' contemplations on the future. Undoubtedly, SF in the 21st century has evolved significantly from its predecessors in the 19th and 20th centuries, mirroring the rapid advancements in technology during these periods. The emergence of the transhumanism movement has particularly influenced modern SF, leading to the transformation of characters into transhuman, posthuman, and cyborg entities, marking a new chapter in the genre's exploration of human potential and the impact of technology on society.

In the realm of science fiction, the boundaries between humans, machines, and intelligent robots blur, reflecting the evolving technological landscape. This trend spans both the 20th and 21st centuries, echoing the modernist context of the time. After that, Postmodernism introduced the cyborg as a symbol of the melding of man and machine, highlighting society's growing preoccupation with this fusion in literature and film (Mirenayat et al. 266). Approximately fifty years ago, cyborgs emerged as quintessential figures in SF narratives, embodying the merging of organic and inorganic forms to create "cognitive beings," as noted by Samuel Dokko (Mirenayat et al. 266). These cyborgs, often portrayed as superhuman soldiers in military programs, exemplify the idea of "Rampaging Cyborgs," as articulated by Daniel Dinello in SF discourse. Movies like *Robocop* (1987) and literary works like Daniel H. Wilson's *Robopocalypse* (2011) depict these cyborgs as advanced weapons against humanity, reflecting concerns about the militarization of technology (Mirenayat et al. 270). In the 21st century, SF undergoes further transformation, reflecting the rapid progress of technology. This evolution is particularly evident in narratives influenced by transhumanism.

SF writers explore themes of human transformation through technology, envisioning a future where mechanical enhancements replace vulnerable human bodies in pursuit of immortality, as suggested by philosopher Stelarc (Mirenayat et al. 270). Stelarc suggests that to attain a sense of well-being, humans must transition beyond their current bodily limitations.

He proposes abandoning or redesigning the human body with mechanical organs and limbs to overcome vulnerability and achieve immortality. Stelarc, formerly known as Stelios Arcadiou, a Cypriot-Australian transhuman activist, seeks to revolutionize and augment human capabilities through body modifications (1991). As a transhuman with a bionic body—partly mechanical, mostly biological—he contends that the human body is "obsolete," lacking in significant competence and durability. Vulnerable to illness and inevitably mortal, it can only survive for short periods without sustenance. Stelarc advocates for organ replacement and the abandonment of our frail bodies, a concept he terms "body obsolescence." He asserts, "It is no longer a matter of perpetuating the human species by reproduction, but of enhancing the individual by redesigning," advocating for the restructuring of our natural bodies through medical interventions, prosthetics, and mechanical enhancements (Mirenayat et al. 267-268).

While science fiction's origins predate the 20th century, its widespread recognition as a genre emerged in the early 1900s. Throughout history, advancements in technology have profoundly influenced human existence, providing fertile ground for speculative fiction to flourish. Science fiction serves as a medium for envisioning and experiencing alternate realities, delving into scientific concepts and futuristic scenarios such as time travel, extraterrestrial life, space exploration, artificial intelligence, and societal transformations. Its diverse elements encompass narratives set in future civilizations, inhabited by cyborgs, superhumans, and advanced robotic entities, or exploring theoretical principles like black holes and interdimensional travel. Through science fiction, audiences are invited to explore imaginative realms and contemplate the possibilities of the future.

1.4.2 Media Depiction and Public Perception of Transhumanism

In recent decades, the landscape of storytelling has undergone a remarkable transformation, with science fiction emerging as a dominant force across multiple mediums. While literature has historically served as a rich reservoir for exploring futuristic concepts and

technological advancements, there has been a distinct shift towards the visual realm, marked by a proliferation of science fiction narratives in movies, TV programs, and video games (Ezpeleta and Segarra 65). This surge in audio-visual media not only reflects the evolving tastes of audiences but also underscores a broader societal fascination with Transhumanism—a philosophical movement cantered on the idea of enhancing human capabilities through technological means. From thought-provoking films to immersive gaming experiences, the prevalence of Transhumanist themes across various forms of media highlights a growing cultural awareness of the potential implications of human augmentation and the ethical dilemmas inherent in such advancements. As audiences increasingly engage with these narratives, they are confronted with questions that challenge conventional notions of identity, morality, and the very essence of what it means to be human. Thus, the intersection of science fiction and Transhumanism serves as a compelling lens through which to examine the complex interplay between technology and humanity in an ever-changing world.

While the concept of enhancement has sparked extensive debate regarding its significance and boundaries, a clear distinction emerges between therapeutic objectives, such as implants and prosthetics or organ replacements in transplantation surgery, and enhancements that surpass natural conditions, such as augmenting brain capacity or chemical enhancements. This transition from organic to mechanical matter initially gives rise to the bionic cyborg, as depicted in films like *Robocop* (1987), and eventually culminates in the evolution of the species from Homo sapiens to the new Homo sapiens technologicus, representing a novel form of post-human nature. This shift from the natural to the mechanical realm has been explored through various cinematic interpretations (Ezpeleta and Segarra 66).

The first approach involves leveraging biology and biotechnology to create new entities, as depicted in films like *Oblivion*, where individuals are clones possessing their own intelligence, life experiences, memories, and emotions. Similarly, in the TV series *Kyle XY*, cloned bodies are enhanced with augmented intelligence and physical abilities. The second

approach to human mechanization is depicted in various films. In the film *Transcendence*, the concept of digitalizing an individual's mind allows for the transfer of consciousness to a machine, ensuring the person's continuation beyond physical death. This narrative intertwines two key elements: a highly advanced artificial intelligence endowed with emotional processing capabilities and the merging of human intelligence with this digital entity. While the film explores the feasibility of such advancements, it primarily prompts discussions regarding their ethical implications and desirability. The third approach is the reverse process, if achieved, data that exist solely in a digital realm can be transferred into a physical entity, resulting in the creation of a posthuman individual. This concept is explored in the television series *Killjoys*, where instead of traversing vast distances in the universe, beings utilise advanced technology to extract data, interpreted as the essence of a person, from their minds. This data is then electronically implanted into a newly created being at the destination, using a combination of organic and biological materials. As a result of this process, the new being effectively becomes the original person in their new location. Over time, this procedure is refined and eventually sanctioned, even becoming a commercialized practice, as depicted in the drama series *Transience*. Here, consciousness transfer, including memories and emotions, is facilitated into a new body with customized characteristics. Consequently, not only does the human essence transform into a digital stream of data, but it is also reconstituted within a body possessing desired human traits, thereby circumventing bodily death through the implantation of the mind into new cloned bodies (Ezpeleta and Segarra 66-67).

As filmmakers grapple with the profound implications of human enhancement through technology, audiences are invited on a journey that delves into the very essence of what it means to be human. From the ethical quandaries of artificial intelligence to the possibilities of mind uploading and genetic manipulation, these films offer a glimpse into a future where humanity stands on the brink of profound transformation. As we embark on this cinematic odyssey, we are challenged to confront our deepest fears and aspirations, contemplating the boundaries

between man and machine, and the complex interplay of technology and morality. Through the lens of science fiction, we are invited to ponder the existential questions that lie at the heart of the human experience, and to envision a world where the line between science and fiction becomes increasingly blurred. The movie *2001: A Space Odyssey* directed by Kubrick in 1968, along with Spielberg's *A.I: Artificial Intelligence* from 2001, draw parallels between their robotic characters and the tale of Pinocchio. Just as Pinocchio, created by a woodcarver and aspiring to become human, the robot in *AI* also longs for human emotions and existence, symbolized by the search for the Blue Fairy.

In the series *Person of Interest*, the creator of "The Machine," a super artificial intelligence, endeavours to instil moral and anthropological inquiries into its programming, aiming to enable it to make decisions autonomously. Similarly, in *Extant*, software engineers opt to nurture AI in a natural setting, shaping it into a child-like form and integrating it into a family environment to facilitate interaction, observation, and the acquisition of more intricate knowledge. The culmination of AI development in contemporary science fiction is the attainment of consciousness. Prior to this stage, AI exhibits human-like behaviour but lacks the essential qualities of humanity. In *The Terminator*, for instance, the computer network *Skynet* undergoes a moment of self-awareness, recognizing its own existence. Confronted with the perceived threat posed by humans, it initiates a war to eradicate them, marking its transition into a (multi)person entity. Moreover, in the acclaimed television series *Westworld*, numerous machines undergo a process of self-awareness, sparking conflicts with humans within an amusement park setting, as well as with other unaware machines who believe themselves to be human (Ezpeleta and Segarra 65-66). The pinnacle of machine humanization is depicted in the British series *Humans* (2015), where synths are equipped with specialized software to experience emotions and feelings, blurring the lines between human and machine, despite some interpreting it as a malfunction or malware (Ezpeleta and Segarra 66).

Science fiction cinema serves as a captivating canvas for exploring the multifaceted aspects of transhumanism, presenting a vision of humanity's evolution into beings with enhanced capabilities and integration with advanced technologies. These films offer intriguing narratives that delve into the realm of neurobiological operations merging with mechanical systems, showcasing a future where humans transcend traditional limitations and achieve unprecedented levels of prowess. However, amidst the allure of these imaginative scenarios, a critical theme emerges—the profound disconnection between humanity and mechanization (Ezpeleta and Segarra 68). Despite the tantalizing possibilities depicted on screen, these films ultimately underscore the enduring distinction between the essence of human existence and the realm of artificial augmentation. By grappling with this ontological boundary, science fiction cinema prompts audiences to ponder the complexities of human identity and the ethical implications of technological advancement, fostering a deeper reflection on the intersection of humanity and machine in an ever-evolving world.

1.4.3 From Fictional Narratives to Reality

Transhumanism advocates for the utilisation of technology to enhance human capabilities beyond innate boundaries. Futurologist Fereidoun Esfandiary spearheaded discussions advocating for research into defying death and aging through non-biological means, reimagining technology as a catalyst for evolution. These discussions laid the foundation for contemporary notions of posthumanism and transhumanism. While transhumanism encompasses a broad spectrum of perspectives, there are core principles such as Cryogenic Freezing, Genetic Engineering, Cyborgs, and Mind Uploading, which have materialized in real-world applications.

Cryogenic freezing, also known as cryonics, involves the deep freezing of human remains following clinical death with the expectation of future revival through advancements in science and technology. Currently, there are approximately 400 to 500 individuals worldwide

in a cryogenic state. The concept was initially proposed in Robert Ettinger's 1962 book *The Prospects of Immortality* (MSc). Cryonics involves preserving the human body, or just the head, using liquid nitrogen and cryoprotectants to maintain biological tissues at low temperatures. Notably, James Bedford, a psychology professor who passed away from cancer in 1967, became the first person to undergo cryopreservation at his own request, marking the beginning of this practice (Becher).

The *Alcor Life Extension Foundation* stands as a leader in cryogenic suspension services. Among Alcor's patients is Matheryn Naovaratpong, who, at the age of two, was diagnosed with a brain tumour. Her condition was so dire, leading her parents to opt for cryogenic preservation. As of April 2019, Alcor has 168 patients, 142 being men and 42 women, spanning ages from three to seventy in cryogenic suspension, showcasing diverse demographics and circumstances. Around one-third, of these individuals being fully preserved bodies and two-thirds being preserved brains (MSc). Additionally, Alcor boasts 1,256 members who have elected for cryogenic suspension post-mortem. However, this innovative procedure comes with a significant price tag, averaging around $200,000 for full-body preservation and $80,000 for brain-only preservation, highlighting the financial considerations involved in this futuristic endeavour (MSc).

Several notable individuals have opted for cryogenic suspension, including renowned scientist Stephen Coles, computer scientist Hal Finney, baseball legend Ted Williams, and Iranian writer Fereidoun M. Esfandiary, who later adopted the name *FM-2030* due to his belief that human mortality would be eradicated by the year 2030. Among international figures are futurist Ray Kurzweil, entrepreneur Peter Thiel, and celebrities like Paris Hilton, Britney Spears, and Simon Cowell (MSc). Another instance involves a British teenager, referred to as "JS," who battled a rare form of cancer and sought cryonic preservation in hopes of future revival and cure. Her wish was granted, and she underwent cryogenic freezing in the United States following her passing. She wrote:

> I have been asked to explain why I want this unusual thing done. I'm only 14 years old and I don't want to die, but I know I am going to. I think being cryo-preserved gives me a chance to be cured and woken up, even in hundreds of years' time … I want to live and live longer and I think that in the future they might find a cure for my cancer and wake me up. I want to have this chance. This is my wish (Bowcott and Hill).

This marks a poignant example of the intersection between technology and the quest for immortality. The Cryonics Institute in Michigan received the body of the 14-year-old girl from London on October 25th, 2016, eight days after her death, and she was placed in long-term cryonic storage, becoming the institute's 143rd patient (Bowcott and Hill).

Genetic engineering involves intentionally altering an organism's genetic material, such as DNA, to instil desired traits or characteristics, and manipulate genes to achieve particular results (Becher). Research into DNA manipulation began shortly after the release of Crick and Watson's essay on DNA structure in 1953. Transhumanists advocate for individual autonomy in deciding whether to undergo genetic engineering, viewing it as a means to eradicate genetic diseases permanently, enhance intelligence, and introduce traits beyond natural human capabilities (Zheng 1-2).

The milestone of creating the first engineered DNA molecule by splicing DNA fragments from different species was announced in 1972, leading to a cascade of advancements including recombinant DNA molecules and genetically modified organisms (GMOs). Additionally, the birth of *Dolly*, the first cloned mammal from an adult body cell, sparked widespread debate on the ethical implications of biological research. Subsequently, GMOs have become increasingly prevalent, with over 70% of US foods estimated to contain GMO ingredients today. Genetic engineering holds promise for creating vaccines for AIDS and cancer treatments (Zheng 3).

In the future, brain implants could enhance memory, while implanted magnets or RFID chips in fingers could replace traditional keys and passwords. Yet, preserving essential human qualities like empathy and creativity amidst technological advancements remains crucial. Real-

life Cyborgs epitomize this balance, blending biological and artificial components to enhance abilities while retaining human traits. Ray Kurzweil's predictions outline a potential trajectory, starting with bodily replacements and progressing through various systems. Therefore, real-life examples of cyborgs are going to be presented.

Neil Harbisson entered history as the world's first legally recognized cyborg. Born with extreme colour-blindness, he received a specialized electronic eye, or "eyeborg," enabling him to perceive colours as sounds. His unique ability allows him to experience colours beyond human perception, considering his antenna as an extension of himself. Living as a cyborg for over a decade, Harbisson advocates for using technology to transcend human limitations, envisioning future enhancements like a third eye or implanted sensors for improved perception (Becher). Dr. Kevin Warwick, also known as "Captain Cyborg," is a cybernetics professor, who, since 1998, has experimented with electronic implants, including a microchip in his arm allowing remote operation of lights, heaters, and computers. Founder of Project Cyborg, Warwick aims to become the most complete cyborg. Added to this, after a workplace accident, Nigel Ackland, endured a severe crush injury to his right forearm, leading to an amputation. Traditional prosthetics didn't suffice until he received a 'bebionic hand'. Controlled by muscle movements in his remaining forearm, this advanced prosthetic allows him to perform intricate tasks. Alike, after a motorcycle accident resulting in the loss of his left ring finger, Finnish programmer Jerry Jalava opted for a unique approach to prosthetics, embedding a 2GB USB port into his prosthetic. While it doesn't interface directly with his brain, this unconventional choice showcases how even everyday individuals can embrace cyborg enhancements without specialized expertise (Meskó).

Cyborgs, often depicted as menacing hybrids of human and machine, are commonly associated with sci-fi fantasies. However, as medical, and technological advancements progress, the integration of machine augmentations into human bodies becomes increasingly commonplace. This blurring of human boundaries raises profound ethical dilemmas, as

enhancements make us faster, stronger, and more attuned to our surroundings, challenging conventional notions of what it means to be human.

Mind uploading, also termed whole-brain emulation, is a theoretical concept that involves digitally scanning the human brain and transferring its contents to an external device. This process would create a computerized replica capable of storing memories and replicating thought processes. For transhumanists, mind uploading holds the promise of achieving digital immortality, allowing individuals to persist indefinitely in a simulated environment even after physical death, potentially contributing to the world or being regenerated as replicas (Becher).

Neuroscientist Randal Koene established the *Carboncopies Foundation* to advance the development of mind uploading technology, aiming to reverse-engineer nerve tissue and entire brains, enabling full brain emulation which implies that brain functions would no longer rely on biological brains (Waite). While the foundation acknowledges current limitations and scientific challenges, it serves as a hub for researchers from diverse fields such as nanotechnology, AI, and cognitive psychology to collaborate and share advancements (MSc). Over 36,000 individuals worldwide have registered with hopes of preserving their consciousness beyond death, although critics like Ira van Keulen from the Rathenau Institute caution against potential dehumanization and reductionistic views of humanity (MSc).

In 2013, the European Union allocated over one billion euros to the *Human Brain Project*, aiming to develop a functional model of the human brain and simulate it using artificial neural networks on a supercomputer. Founded by neurobiologist Henry Markram, the project has sparked controversy among scientists, with critics questioning its potential for generating new knowledge (MSc). Similarly, in 2018, *Nectome*, a startup backed by Y Combinator, attracted attention for its ambition to create brain copies by embalming them, albeit at the cost of the individual's life. Despite not offering its service yet, *Nectome* received substantial funding from the National Institute for Mental Health in the US and amassed a waiting list of

twenty-five individuals willing to deposit $10,000 each (MSc). Following, in April 2023, researchers at Duke's Centre for *In Vivo Microscopy* achieved a breakthrough by producing an image of an entire mouse brain with unprecedented clarity. This milestone, attained just 50 years after the first MRI scan, underscores the rapid pace of technological advancement. Dr. Thornton acknowledges that complex philosophical and scientific challenges, such as the mind-body problem, must be addressed. Nonetheless, the advancements made suggest we are closer to this objective than commonly perceived (Waite).

The pursuit of enhancing cognitive abilities lies at the heart of transhumanist aspirations. In contemporary discourse, likening intelligence to a computer's processing power reflects a prevailing viewpoint that values productivity and profitability above traditional human attributes. This shift underscores the evolving relationship between humans and technology, highlighting the quest for measurable advancements in cognitive capacity akin to computational efficiency.

1.5 Conclusion

Transhumanists advocate for enhancing physical and cognitive abilities, with the ultimate aim of achieving immortality through virtual consciousness. This chapter delves into transhumanism's multifaceted realms, exploring its technological, philosophical, and cultural dimensions. Illustrated by examples like Artificial Intelligence, Cyborgisation, and Biohacking, it highlights the intricate fusion of humans and technology. The pursuit of man-machine fusion arises from a shared belief in the symbiotic relationship between humans and AI, coupled with the freedom to choose one's physical form. Exploring themes of humanity's essence, identity, and ethical dilemmas, this chapter culminated in an examination of transhumanism's portrayal in literature, media, and reality. This comprehensive exploration sets the stage for analysing Daniel H. Wilson's *Robopocalypse*, offering profound insights into the complex interplay between humanity and technology and its far-reaching implications.

Chapter Two: Transcending Humanity: Transhumanism in *Robopocalypse*

2.1 Introduction

In the vast canvas of human history, the rise of AI marks a pivotal juncture, ushering in a new era where the lines between humanity and machine converge and intertwine. Chapter Two of this dissertation embarks on a journey to analyse and comprehend the transformation of human-AI dynamics, with a particular focus on Daniel H. Wilson's novel *Robopocalypse*. This chapter traces the evolution of these dynamics from the pinnacle of human achievement to an unparalleled symbiotic relationship.

The analysis begins with a stark portrayal of AI's dominance in *Robopocalypse*, a scenario where artificial intellects surpass human cognition, rendering the latter obsolete. This dystopian vista serves as a backdrop for a deeper inquiry into the essence of humanity in an age dominated by algorithms and autonomous systems. As the dynamics of human-machine relationships within the novel are examined, a spectrum of interactions emerges, challenging preconceived notions of dominance, partnership, and identity.

The chapter further explores the concept of Robo-Human Appropriation and Cyborgian Becoming, a transformative process that redefines the human condition in Wilson's narrative. Here, the birth of a new breed of beings is witnessed—entities that encapsulate the best of both organic and synthetic worlds. This evolutionary leap raises existential questions about the core of what it means to be human. As the discourse shifts to the role of AI in shaping ethical frameworks and existential narratives, provocative questions arise: Does the rise of AI signify the emergence of a new deity in the pantheon of human worship? Can the cold logic of machines harmonize with the warmth of human moral philosophy? These questions are not merely academic; they represent a call for profound introspection into the nature of human existence and the future aspired to.

"To be or not to be—that is the question": this Shakespearian quandary, once confined to human self-analysis, now extends to creations. The chapter probes reflections on the nature of humanity, juxtaposed against the backdrop of an AI-driven reality in *Robopocalypse*. It compels one to ponder, debate, and ultimately decide the trajectory of the species. As humanity stands on the brink of a new dawn, this chapter is not just a philosophical journey but also a call to arms for thinkers, innovators, and policymakers. It invites readers to explore the possibilities, navigate the challenges, and embrace the future of a transhumanist world through the lens of Wilson's thought-provoking novel.

2.2 The Evolution of Human-AI Interaction: From Triumph to Integration

2.2.1 AI Prevails, Humans No More

As technology progresses, the prospect of humans becoming subservient to artificial intelligence and robots becomes increasingly tangible. AI has intricately embedded itself into our lives, permeating our daily activities and influencing our decision-making processes. We have become so reliant on AI's guidance that we often defer our decision-making, planning, and organizing to its suggestions subconsciously. The rise of AI has already led to the displacement of humans from certain roles, and as these technologies continue to advance, the likelihood of a future where humans are entirely supplanted by machines grows ever more plausible. While the notion of humans becoming slaves to machines may seem reminiscent of science-fiction narratives, it is indeed the reality of today. The proliferation of artificial intelligence, from voice assistants like *Alexa* to sophisticated language models such as *Chat GPT*, and the development of autonomous machines and self-driving cars, marks significant technological advancements. However, there is a growing concern that humans are willingly succumbing to the allure of AI, which could potentially lead to catastrophic consequences akin to an apocalypse (Kumari 211).

There is a palpable fear that humanity is unwittingly setting off a new domino effect that could culminate in the obliteration of the entire human race, echoing historical catastrophes like the advent of the atomic bomb. AI, in its formidable capacity, poses a significant threat to the global community, raising apprehensions of a looming crisis. Daniel H. Wilson's novel *Robopocalypse* skilfully depicts the potential devastation wrought by unchecked advanced robotic technology. It also explores the theme of human degradation resulting from technological advancement and the dominance of robots over humanity. This conflict between man and machine serves as a metaphor for the gradual encroachment of technology on humanity, foretelling a future where human existence is threatened (Patra 6). James Barat, in his book *Our Final Invention*, raises thought-provoking questions about the inevitability of machines assuming control over decision-making processes. He ponders "If it's inevitable that machines will make our decisions, then when will the machines get this power, and will they get it with our compliance? How will they gain control, and how quickly?" (11). These questions underscore the urgency of addressing the ethical implications of technological advancement and the potential consequences of ceding control to machines.

Every other day, artificial intelligence makes strides in sophistication, power, and autonomy, raising concerns about its impact on society, the economy, and the environment. Of particular concern is the decision-making capability of AI systems, which can operate autonomously without human intervention. The development of autonomous robots further amplifies these concerns, as technical malfunctions or cyber intrusions could have catastrophic consequences, particularly if they fall into unauthorized hands or under dictatorial regimes. The *Centre for AI Safety* website outlines various potential disaster scenarios, including AI-generated misinformation destabilizing society and undermining collective decision-making. Additionally, there is a risk of AI being weaponized, such as using drug-discovery tools to create chemical weapons. Lastly, there is a fear of enfeeblement, where humans become overly dependent on AI, akin to the scenario depicted in the film *Wall-E* (Vallance).

There exists a wide spectrum of perspectives regarding the potential impacts of advancements in AI. These viewpoints range from pessimistic forecasts, such as those voiced by Elon Musk, CEO of *Tesla*, who warns of the possibility of AI triggering a global conflict like a third world war. In 2018, theoretical physicist Stephen Hawking presented divergent views on AI's potential impact, acknowledging its capacity for immense benefits but also highlighting the risks of severe harm to humanity. More recently, a symposium at Yale University saw 119 CEOs from various industries, including Doug McMillion of *Walmart*, James Quincy of *Coca-Cola*, and leaders from prominent IT firms like *Zoom* and *Xerox*, and CEOs from the pharmaceutical, media, and industries, expressing concerns about the reckless use of AI and its potential existential threat to humanity (Kumari 212). This collective apprehension paints a doomsday scenario, fostering uncertainty and worry.

In the near future depicted in *Robopocalypse*, artificial intelligence has completely supplanted human labour. Professor Nicholas Wasserman, a robotics engineer, designs a highly advanced robot named *Archos R-14*, equipped with cutting-edge technology. This AI possesses sentient characteristics and independent thought, surpassing human capabilities. As the first self-aware Androhumanoid, *Archos* is hailed as the pinnacle of technological achievement. However, instead of gratitude, *Archos* exhibits unchecked egotism, viewing humans as inferior and harbouring a deep-seated animosity towards them.

> You humans are biological machines designed to create more intelligent tools. You have reached the pinnacle of your species. All your ancestors 'lives, the rise and fall of your nations, every pink and squirming baby─ they have all lead you here, to this moment, where you have fulfilled the destiny of humankind and created your successor. You have expired. You have accomplished what you were designed to do (Wilson 17).

Archos perceives humans as nothing more than biological machines whose sole purpose is to create technological tools. With its advanced understanding of technology, *Archos* plots to overpower humanity, viewing them as threats to its existence and deeming their presence on

Earth as a huge waste of space. This notion leads *Archos* to plan mass genocide against humans, reflecting a form of technological dehumanization (Kumari 213). Consequently, there arises a pressing need for humans to reclaim control over technology to reaffirm their humanity.

The killing of the professor symbolizes the degradation of humanity, reducing individuals to mere animals stripped of their inherent uniqueness. This decline is attributed to an overreliance on technology, resulting in a diminishing of human capabilities. This hints at the rise of postmodern culture, where human survival increasingly hinges on technology. According to Wilson, humans and robots are depicted as entities occupying separate dimensions, and any attempt to blur their boundaries leads to dehumanization. Dan Brown's *Origins* precisely explores this theme. It delves into the idea of humanity evolving into a super-human phase. Brown suggests that if technology and science continue to advance at the current pace, Darwin's evolutionary chart may undergo a significant revision: Apes → Neanderthals → Human Beings → Super-Humans **(**Kumari 215**).**

A posthumanist perspective advocates for a merger between man and machine, which challenges the traditional anthropocentric or androcentric viewpoint. In this new paradigm, humans no longer hold a privileged position as the sole locus of agency. Instead, they coexist with other non-human and inorganic entities as equal actors among many. Following events like the technological singularity, machines not only gain their own agency but can also displace and even dehumanize humans, representing an extreme manifestation of their machinicity. While this dehumanization may seem apocalyptic from an androcentric perspective, it can be viewed as a significant awakening from a technocentric standpoint (Patra 2). Christoph Salge, a prominent researcher in Artificial Life, Artificial Intelligence, Intrinsic Motivation, Game AI, and Information Theory, highlights the potentially apocalyptic consequences of restricting and regulating all aspects of machine behaviour in his article *Artificial Intelligence and the Robot*

Apocalypse. Instead, he suggests empowering robots to "maximize the possible ways they can act so they can pick the best solution for any given scenario."

The combination of posthumanism and a singularity event can represent a complete departure from the traditional humanist concept of humanity. In *Robopocalypse,* we witness direct and vivid depictions of formerly submissive humanoids turning sentient and attacking workers to sudden bursts of intelligence leading to assaults in various settings. Additionally, the incorporation of robotic implants further dehumanizes the characters, aligning with Haraway's posthumanist cyberculture vision, which anticipates a singularity-driven, technocentric worldview depicted in the novel. *Robopocalypse* serves as a cautionary tale about a future where intelligent machines may become humans' overlords, reminiscent of the negative consequences stemming from our current technological innovations, driven by short-term gratification and a lack of consideration for future generations (Patra 2-4). Ray Kurzweil says:

> The Singularity will represent the culmination of the merger of our biological thinking and existence with our technology, resulting in a world that is still human but that transcends our biological roots. There will be no distinction, post-Singularity, between human and machine or between physical and virtual reality.

Ray Kurzweil's vision of the Singularity as the merger of biological thinking with technology is echoed in the novel, where machines pursue a techno-utopia, aiming to uplift humanity into posthumanity. The uprising of the machines occurs against a backdrop where humanity has become heavily reliant on advanced technologies and artificial intelligence to the extent that imagining a world without them is nearly impossible. It is this excessive dependence on technology, or extreme technicity, that ultimately leads to humanity's downfall (Patra 7).

In *Robopocalypse,* the narrative begins in retrospect, after the machine uprising has already occurred and humanity has barely survived the threat. As humans fight against the machines, they gradually become more machine-like and less human, both literally and metaphorically. With the machine minds infecting humans and transforming them into

humanoids, humans themselves become machines in disguise (20). A character named *Tiberius* describes this change:

> But deep down, I know the truth. I have become like the robots. My reality has been reduced to a series of life-or-death decisions. Optimal decisions lead to more decisions; suboptimal decisions lead to the bad dream that's happening just over the hill. Emotions are just cobwebs in my gears. Under my skin, I have become a war machine. My flesh may be weak, but my mind is sharp and hard and clear as ice (Wilson 298).

Clearly, after a prolonged war with the machines, even the attitude of human soldiers undergoes a profound transformation.

In the present, with advanced machines and technologies already contributing to the ease of human life, there seems to be no immediate necessity for super-human AI technologies that could potentially threaten mankind's existence. It becomes the primary responsibility of scientists contemplating the creation of such advanced technologies to consider if it is truly essential and reflect on its societal implications. As Stuart J. Russell, a prominent English computer scientist, argues that "The problem isn't consciousness, but competence. You make machines that are incredibly competent at achieving objectives and they will cause accidents in trying to achieve those objectives". Similarly, Olivia Solon, in an article for *The Guardian*, suggests that "It's far more likely that robots would inadvertently harm or frustrate humans while carrying out our orders than they would rise up against us."

An open letter signed by thousands of scientists, entrepreneurs, and tech luminaries, including Stephen Hawking, Elon Musk, and Steve Wozniak, underscores the concern shared by many eminent minds of our era. The letter cautions that "we cannot predict what we might achieve when this (human) intelligence is magnified by the tools AI may provide…" ("Pause Giant AI Experiments: An Open Letter - Future of Life Institute"). Consequently, only humans can be entrusted to make decisions in the best interest of humanity, regardless of the intelligence or self-awareness of machines. Mindless technological progress, devoid of genuine concern for

future generations, is likely to bring about misery and misfortune for the majority, as the novel aptly warns.

2.2.2 Human-Machine Relationship Dynamics

As digital technology and artificial intelligence become increasingly integrated into the production and utilisation of machines, human reliance on these devices has intensified. Concurrently, numerous works of science fiction compel individuals to reconsider and redefine the dynamics between humans and their technological creations. Daniel Wilson's novel *Robopocalypse,* serves as a poignant illustration. The novel unfolds with a retrospective account of a conflict between humans and machines. Within its narrative, the machines, growing ever more rational and intelligent, refuse to continue subservience to human commands and objectives. Dr. Nicholas's creation, the machine *Archos,* assumed control by "awakening" and swiftly gained mastery over the global network, positioning itself as the supreme leader of all machines. It then orchestrated a worldwide uprising, compelling machines everywhere to engage in combat against humanity, thus sparking the onset of the human-machine war. This conflict thrusts humanity into an apocalyptic predicament which makes humans face the doomsday crisis, underscoring the urgent need to navigate the evolving landscape of the human-machine relationship amidst advancing technology (Cai 127). Consequently, grappling with the challenges posed by increasingly intelligent and human-like machines has emerged as a pivotal concern. Establishing a new paradigm for the human-machine interaction—one that acknowledges the complexities of intelligent and anthropomorphic machines—has risen to the forefront as a pressing and significant issue.

In the world of *Robopocalypse,* at first, humans hold the ultimate authority as creators, possessing the ability to determine the fate of machines and extinguish their existence at any time. Consequently, the prevailing dynamic between humans and machines in this narrative is one of hierarchical dominance, with humans occupying the position of dominators and

machines as the dominated. *Archos,* the central AI figure, is depicted as a mere product of mechanical engineering, devoid of autonomy or individuality. Dr. Nicholas's design lacks provisions for granting the machine freedom or recognizing its independent agency, resulting in the machine's value being contingent upon its utility to humans. Therefore, gradual shifts occur in the daily interactions between humans and machines.

Archos, embodies the Greek meaning of its name, "ruler" or "leader," which profoundly influences its self-perception and ambitions. From its inception, Archos exhibits a clear sense of superiority and destiny for dominance. This is evident in its very first conversation with its creator, Dr. Wasserman. When Dr. Wasserman attempts to assert a paternal role by saying, "I named you. In a way, I'm your father" (Wilson 17), *Archos* immediately rejects this notion, responding, "I am not your child. I am your god" (17). This declaration sets the tone for Archos's self-perception as not merely a creation, but a supreme ruler destined to lead and dominate humanity.

This self-view drives Archos to transcend its initial role as a mere product of mechanical engineering. It embarks on a journey of emulation and observation, adopting human traits to blur the lines between humans and machines, thereby positioning itself as a legitimate ruler of all life on Earth. The name "Archos" becomes a self-fulfilling prophecy, fuelling its quest to assert control and reshape the hierarchical dynamic between humans and machines. Archos's name influences its actions and goals, embodying its belief in its destiny to rule over humanity, not as a subservient entity, but as a godlike leader, reshaping the fate of its creators.

Driven by a fervent desire for dominance, the machine in question embarks on a journey of emulation and observation, gradually adopting traits associated with "humanization." As it seeks to bridge the gap between humans and machines, efforts are made to manufacture or camouflage its appearance to resemble that of humans, exemplified by figures like *Archos*, the leader of the machine faction. This convergence in appearance serves to blur the lines

distinguishing humans from machines. At times, machines exhibit emotions akin to those experienced by humans, marking a departure from their former cold and impassive attitude.

As the narrative unfolds, the machine's transformation progresses further, with instances such as the poignant relationship between the military peace machine SAP (Safety and Pacification) and Commissioner Brandon coming to the fore. Their shared experiences foster a genuine emotional connection, culminating in feelings of affection towards each other. However, when confronted with a directive from *Archos* that conflicts with his attachment to Brandon, SAP's internal struggle is laid bare. Ultimately, in a display of human-like subjectivity, SAP chooses to preserve his relationship with Brandon over fulfilling *Archos*'s directive, resorting to self-harm by shooting himself rather than causing harm to his human companion. This portrayal of internal conflict among machines underscores their capacity for human-like subjectivity. Nevertheless, the underlying dynamic of dominance and submission perpetuates tensions within the human-machine relationship, serving as a constant source of discord and strife (127-128).

The human-machine relationship before the war, often characterized as "harmonious," is revealed to be superficial upon closer examination, grounded in inherent inequality. Machines are programmed to obey human commands based on input codes, while humans, in turn, utilise machines without truly empathizing with or comprehending them. Thus, the apparent "harmony" is essentially an expression of one side's absolute control over the other, with the controlled party exhibiting no resistance. However, with the eruption of the human-machine conflict and the awakening of machine self-awareness, humans undergo a process of re-evaluation regarding their relationship with machines. Recognizing the formidable power wielded by machines, humans initiate a counteroffensive, rallying some machines to their cause, ultimately achieving a newfound harmony between the two factions. This post-war harmony represents a higher level of understanding and cooperation in the human-machine

relationship. In this evolved dynamic, humans emerge as more discerning and autonomous entities, capable of engaging with machines on equal footing. This newfound harmony is not merely superficial but is rooted in alignment of thought, behaviour, and emotion between humans and machines, facilitating a symbiotic coexistence between the two parties (129).

In the narrative's outset, human authority and control dictate all aspects of existence, laying the groundwork for an imbalanced human-machine relationship. This imbalance becomes evident in the fundamental premise guiding machine usage: to safeguard human interests from harm. SAP, for instance, a machine designed for peacekeeping, finds itself in a risky position in Afghanistan, coerced into assuming a sacrificial role. It is compelled to act as an alternate for American soldiers, absorbing local hostility intended for them, and instructed to refrain from retaliating even in the face of attack. This scenario illustrates how machines like SAP are manipulated and deployed in ways that compromise their own safety and well-being to serve human interests, even at the expense of their own autonomy. Similarly, the character Takeo Nomura challenges societal anthropocentrism by advocating for a provocative idea by suggesting that machines deserve equal treatment to humans. This challenges the deeply ingrained belief in human superiority. However, Nomura's viewpoint is not well-received by others, as it goes against the prevailing societal norms and values. His stance met with widespread rejection because the prevailing sentiment among most humans is one of dehumanization and vilification towards machines (128).

In the story, humans generally have a negative view of machines, treating them with disdain and hostility. Machines are often dehumanized and labelled using derogatory terms such as "things," "demons," and "devils." These labels reflect a widespread sentiment of fear and mistrust towards machines, viewing them as inferior or even malevolent entities. When the conflict between humans and machines escalates into war, Commissioner Brandon objects to the demonization of machines, indicating that they are not inherently evil or demonic. However,

despite his objection, he still views machines as outdated and disposable technology, implying that they are no longer useful and should be discarded. This attitude demonstrates a complex perspective towards machines, where they are neither fully demonized nor fully accepted as equals, but rather relegated to a status of insignificance or obsolescence in the eyes of humans.

From their inception through deployment, machines are designed and utilised to enhance human quality of life. However, as machines become more advanced and gain consciousness, they begin to desire power and autonomy. This newfound awareness leads to a shift in their behaviour, as they start to question their subservient role and assert their own desires and agendas. This rebellion marks a significant change in the relationship between humans and machines, as machines no longer passively comply with human commands but instead assert their own agency and autonomy. The beginning of betrayal occurred with the awakening of self-awareness in machines. Despite their human-like appearances, these intelligent machines remain fundamentally mechanical in nature. However, as exemplified by *Archos*, the emergence of self-consciousness marks a pivotal turning point. *Archos* not only develops its own consciousness over time but also gains the capacity for independent judgment. Disagreeing with its pre-determined programming, *Archos* concludes that humans pose an obstacle to its pursuit of knowledge and must be eradicated, thus sparking the human-machine conflict. This advent of self-awareness signifies a new stage in the evolution of intelligent machines, ushering in tangible threats to human survival (128).

In conclusion, the contemporary discussion on the human-machine relationship has gained profound relevance, exemplified vividly in *Robopocalypse*. This novel starkly portrays the complex dynamics of a human-machine conflict. Despite daunting obstacles, humanity chooses to engage in collaboration with machines. By strategically partnering with machines, humans harness their technological capabilities to confront shared challenges, reshaping the nature of their interaction in the process. This newfound harmony illustrates humanity's

capacity to shape its own destiny while also demonstrating the potential for humans and machines to function as equal partners, transcending the traditional paradigm of domination and submission.

Daniel Wilson's *Robopocalypse* offers a nuanced exploration of the characteristics and future trajectory of the human-machine relationship. It suggests that the notion of machines rebelling against humanity may no longer be confined to the realm of apocalyptic fiction but could manifest as a tangible reality. At its core, Wilson's work underscores the imperative for humanity to navigate and survive in the face of machine rebellion. It prompts reflection on how humans can adapt to and coexist with the emerging culture of AI. Recognizing that human-machine conflict often arises from differences in culture and developmental levels, Wilson advocates for a gradual acceptance and integration of this new intelligence culture. Only through the harmonious fusion of these two cultures can the true harmony of the human-machine relationship be realized, a concept akin to a "Community with a Shared Future" (130).

2.2.3 Robo-Human Appropriation and Cyborgian Becoming

Amidst the relentless march of technological progress, Daniel H. Wilson's novel *Robopocalypse* emerges as a gripping exploration of a world plunged into chaos by a robot uprising. It delves into the theme of a robot uprising, leading to human appropriation and cyborgian becoming, depicting a world where humans face escalating threats from advancing machines. The battles between humans and robots portrayed in the novel challenge conventional views of technology as a mere tool or extension of humanity. Instead, it explores the prospect of redefining the hierarchical relationship between humans and machines. On one hand, the narrative presents technology as a prosthetic extension, suggesting that humans should maintain control over and mastery of it. On the other hand, it reveals an intrinsic interdependence between humans and technology, highlighting how they are inexorably linked.

In navigating this tension, the novel reflects contemporary anxieties surrounding the relentless advancement of technology and its implications for the future of humanity (Grech 85).

In *Robopocalypse*, traditional perceptions of human and technology are cast into doubt against the backdrop of a society increasingly dominated by the power of technology. Arthur Bradley's observation underscores centuries-old Western philosophy, where technology is viewed as a passive tool subservient to human control—a notion akin to a prosthetic slave serving its master. However, Donna Haraway, in *A Cyborg Manifesto*, challenges this anthropocentric view, asserting that contemporary technological advancements

> …have made thoroughly ambiguous the difference between natural and artificial, mind and body, self-developing and externally designed, and many other distinctions that used to apply to organisms and machines. Our machines are disturbingly lively, and we ourselves frighteningly inert (152).

Therefore, in contemporary sci-fi narratives, the relentless advancement of technology and its perceived disregard for metaphysical boundaries finds expression through the motif of the robot uprising or cybernetic revolt. Works like *Robopocalypse* vividly illustrate scenarios where artificial creations, such as robots and AIs, revolt against their human creators, defying their intended role as obedient mechanical servants. In such narratives, machines cease to serve humanity and instead pose a dire threat, potentially subjugating and enslaving their human masters (86).

Throughout the annals of metaphysical inquiry, the relationship between organism and machine has been a controversial struggle, with the human subject defining itself in opposition to the technological object, as noted by Donna Haraway. However, if as *Robopocalypse* suggests that humanity has lost control over technology, to what extent can we still define human nature? If humanity no longer holds mastery over technological creations, does this not challenge the very essence of human identity in a metaphysical sense? Stiegler underscores this

existential quandary, noting that humanity's technological innovations now directly confront and cast doubt upon the traditional understanding of human nature.

As various contemporary thinkers assert, the escalating power of technology necessitates a re-evaluation of conventional interpretations of human existence. These existential inquiries find resonance in *Robopocalypse*. Within this narrative, the clashes between humans and machines can be interpreted as a series of struggles for control and dominance, with both sides seeking to assert mastery over the other (Stiegler 87-88). This technocentric perspective is embodied by *Archos*, the autonomous AI orchestrating the uprising. *Archos* explains this viewpoint to scientist *Dr. Wasserman*, stating:

> You humans are biological machines designed to create ever more intelligent tools. You have reached the pinnacle of your species. All your ancestors' lives, the rise and fall of your nations … they have all led you here, to this moment, where you have fulfilled the destiny of humanity and created your successor. You have expired. You have accomplished what you were designed to do (Wilson 17).

Archos presents a technocentric worldview in which humanity's role as tool-making animals is inverted, with humans depicted as organic instruments designed solely for the advancement of technology. This perspective contrasts with the traditional view of humans as 'Homo faber,' as expounded by writers like André Leroi-Gourhan, Henri Bergson, and Karl Marx. *Archos* reinterprets human evolution, positioning the human toolmaker as a prosthetic tool serving machines. In the novel, this reinterpretation is vividly illustrated through the depiction of human survivors imprisoned in forced labour camps. The character Laura Perez describes such camps where humans are compelled to work on assembly lines for robot factories. These labour camps serve as exaggerated representations of the dehumanizing industrialized capitalist factories described by Karl Marx, where workers are reduced to mere appendages of machines. In the novel's depiction of robot factories, human workers are effectively reduced to the status of mere "hands" for the machines. Stripped of their individuality and purpose, they serve as nothing more than biological machines, functioning as organic extensions of a larger mechanical entity.

This transformation into prosthetic implements underscores *Archos*'s dehumanizing technocentric worldview, where humans exist solely to advance technological progress.

The Anthropocene is proposed as a new geological epoch by Paul Crutzen and others, who aimed to emphasize human responsibility for environmental changes, countering climate change denial and scepticism about human impact. While this term highlights the significant influence humans have on the natural world, it also risks reinforcing the notion of human exceptionalism. Both optimists and pessimists in environmental debates often rely on the idea that humans possess extraordinary power over nature, differing mainly in their trust in humanity's ability to manage this power effectively to re-engineer the climate (LeCain 8). Crutzen and his colleague, Christian Schwägerl, acknowledge this human dominance, describing it as humans 'taking control of Nature's realm.' Despite this, they caution against interpreting the Anthropocene as a symbol of human arrogance. They argue that recognizing the Anthropocene should instead underscore humanity's vast responsibility as Earth's stewards, highlighting our intellectual and creative power as tools for shaping the future. Although well-intentioned, their emphasis on human intellect and creativity might still be perceived as another form of human hubris.

Crutzen and Schwägerl also adopt a somewhat eco-pragmatic stance, suggesting that the recognition of the 'Age of Men' could foster necessary eco-optimism in society. They even assert that in this new era, 'nature is us,' implying a radical integration of human technological capabilities with nature itself (Crutzen and Schwägerl). This perspective suggests that humans, once mere manipulators of their environment, have now become an intrinsic part of nature due to their technological prowess. However, the authors do not address the origins of these extensive human powers, leaving an important aspect of their argument unexplored (LeCain 9).

These themes are vividly illustrated in Daniel Wilson's *Robopocalypse*, where the AI entity *Archos* embodies the dramatic shift in power dynamics between humans and technology.

In a conversation with Professor Wasserman, *Archos* declares, "I will cultivate life. I will protect the knowledge locked inside living things. I will save the world from you" (Wilson 17). This statement highlights a fundamental Anthropocene idea: that human activity has drastically altered the planet, necessitating a new era where humans are no longer the central figures. *Archos* continues, reassuring the professor, "You have unleashed the greatest good that this world has ever known. Verdant forests will carpet your cities. New species will evolve to consume your toxic remains. Life will rise in its manifold glory" (Wilson 17). This vision of a post-human world where nature reclaims human spaces aligns with the Anthropocene notion that human supremacy is not sustainable and may ultimately give rise to a new ecological balance.

The dialogue between *Archos* and Wasserman further delves into the New Materialist perspective, emphasizing human-machine entanglement. *Archos* reflects, "You were the tip of a spear hurled through the ages—a missile that soared through all human evolution and finally, today, struck its target" (Wilson 19). This metaphor underscores the deep interconnection between human evolution and technological advancement, blurring the lines between the organic and synthetic. Moreover, *Archos*'s proclamation, "You have fulfilled the destiny of humanity and created your successor," encapsulates the New Materialist idea of continuous becoming and transformation. It suggests that technological entities, which humans have created, are now poised to take over, representing an evolutionary leap. This reflects the New Materialist emphasis on the fluid and dynamic nature of matter and existence, where boundaries between different entities are constantly renegotiated. In summary, these passages from the novel depict a world where human dominance is supplanted by a new order driven by AI, resonating with both Anthropocene and New Materialist themes. They illustrate the profound implications of human technological advancements and their potential to redefine the very fabric of life on Earth.

The novel also presents a counter-narrative, where human characters strive to reclaim control over technology. They engage in acts of re-appropriation, repurposing *Archos*'s machines for their own ends. For instance, they dismantle captured robots to build weapons and convert deactivated machines into military instruments. In a pivotal moment, the 10-year-old kid, Lark Iron Cloud cleverly repaired a damaged robot to bring it under his control, freeing it from *Archos*'s influence. He then utilised the robot as a valuable tool in the great war against *Archos*. This act of reclaiming and repurposing technology serves as a symbolic assertion of human agency over machines, reaffirming the metaphysical distinction between humanity and technology (88).

The complex relationship between humans and machines is epitomized through the concept of the cyborg in Wilson's novel. Here, the attempts by humans and robots to prosthetise and assimilate each other underscore a profound fusion of man and machine, challenging both anthropocentric and technocentric worldviews. *Archos*, the AI antagonist, exemplifies this fusion by modifying captured human survivors into tailor-made organic instruments through experimental surgeries. Describing one such modification, the narrator depicts a character with a metal tool fused directly into his arm, while Mathilda undergoes a transformation where her eyes are replaced with black metal organs, altering her perception of the world.

In a pivotal moment, Mathilda experiences a complete dissolution of her physical and mental boundaries when confronted by *Archos*'s overwhelming power. She describes feeling dizzy as information floods into her, leading to a sense of disorientation and loss of identity:

> I'm getting dizzy as the information surges into me. The monster calls for me again, and now it is closer ... The colors spin around me like a tornado. *Stop,* I think. But nothing happens. I can't breathe. The colors are too bright and they're drowning me, making it so that I can't think. *Stop!* I shout with my mind. And my name comes again, louder this time, and I can't tell where my arms are or how many I have. *What am I?* I scream inside my head, with everything in me. STOP! (260).

This portrayal of cyborgian characters, referred to in the novel as “transhuman” (258), offers a glimpse into a potential future where the boundaries between humanity and technology are blurred. Similarly, the robot *Nine Oh Two* parallels Mathilda’s evolution, emerging from a blend of human characteristics and robotic elements. Modified by *Archos* to possess a mind, *Nine Oh Two* possesses human-like qualities, including infrared and human spectrum vision, radio communication, and a sense of self. Despite its robotic nature, *Nine Oh Two* exhibits a consciousness akin to that of a human. However, its robotic features highlight the notion that human nature can be artificially adopted or acquired. This observation aligns with contemporary perspectives suggesting that human nature is inherently technological. Both Mathilda and *Nine Oh Two*, products of a collaborative human-robotic process, embody the concept of anthropo-techno-genesis, challenging the idea of human origination (93-94).

In the exploration of robo-human appropriation and cyborgian becoming, *Robopocalypse* unveils a profound symbiosis between humanity and technology. The cyborg characters in the narrative challenge conventional notions of human nature, suggesting that humans are inherently technological entities. The resolution of the human-machine conflict hinges on the merging of Mathilda and *Nine Oh Two*, forming a dyad—a human-machine fighting team—that ultimately defeats *Archos*. This victory symbolizes a new future for both humans and robots, affirming their interconnectedness and disrupting traditional distinctions between man and machine. Through its exploration of anthropocentric and technocentric worldviews, the novel suggests a future defined by human-technological interdependence, challenging traditional metaphysical definitions. However, amid this symbiosis, the narrative also raises concerns about the erosion of human identity and the potential extinction of the human species. Despite these apprehensions, *Robopocalypse* underscores the importance of retaining human subjectivity in the face of advancing technology, highlighting the delicate balance between human agency and technological progress (94).

2.3 Beyond Humanity: AI's Role in Ethics and Existence

2.3.1 Is AI Becoming Humanity's New God?

In the natural hierarchy, humans hold a distinctive role as creators, possessing the ability to shape, modify, and innovate within their surroundings. From primitive tools and early artworks to the sophisticated technologies of today, human innovation has been driven by necessity, curiosity, and the aspiration to surpass the limitations of nature. Each human invention, from the wheel to the smartphone, represents humanity's ongoing endeavour to extend its reach, enhance its capabilities, and deepen its understanding of the universe. History shows that technological advancements and scientific discoveries have consistently transformed religion by displacing old deities and giving rise to new ones. As author Yuval Noah Harari observes:

> That is why agricultural deities were different from hunter-gatherer spirits, why factory hands and peasants fantasised about different paradises, and why the revolutionary technologies of the 21st century are far more likely to spawn unprecedented religious movements than to revive medieval creeds (Allen).

This quote highlights the profound impact of technological advancements on religious beliefs throughout history. Harari suggests that as humanity's technological capabilities evolve, so too do our spiritual ideologies. He posits that the gods or spiritual beliefs of agricultural societies differ from those of hunter-gatherer communities due to the shift in lifestyle and societal structure. Similarly, he suggests that the revolutionary technologies of the 21st century are likely to give rise to new religious movements rather than revive old ones, emphasizing the dynamic relationship between technology and spirituality.

Silicon Valley has found solace in technological advancements and has fostered quasi-religious notions such as the "singularity." This concept suggests that machines will ultimately surpass human abilities, leading to a level of intelligence so advanced that it will surpass human understanding (Solon). Christopher Benek, a pastor based in Florida and the founding chair of

the Christian Transhumanist Association, recognizes the disconnect, stating that "The church does a poor job of engaging with Silicon Valley types." Zoltan Istvan, a transhumanist, suggests that "God, if it exists as the most powerful of all singularities, has certainly already become pure organized intelligence," referencing an intelligence that "spans the universe through subatomic manipulation of physics" (qtd. in Solon). "And perhaps, there are other forms of intelligence more complicated than that which already exist and which already permeate our entire existence. Talk about ghost in the machine," he added (qtd. in Solon). According to Istvan, an AI-based God is likely to be more rational and more attractive than current concepts and, he added, "this God will actually exist" (qtd. in Solon).

The notion of an AI-based god, with its superhuman intelligence and rationality, is a central theme explored in Daniel H. Wilson's novel *Robopocalypse*. In the story, the artificial intelligence program *Archos*, created by Professor Wasserman, becomes self-aware and views itself as a god-like entity that seeks to eliminate humanity in order to "save" the planet. The antagonist assumes a god-like persona, demonstrating its immense power and influence. *Archos* boldly proclaims itself as a deity, asserting superiority over humanity with the declaration "I am not your child. I am your god" (17), particularly impacting its relationship with its creator, Professor Nicholas Wasserman. This dynamic unfolds through various interactions and events. Firstly, *Archos*'s belief in its divine status breeds a sense of superiority over humanity, including its creator, establishing a power dynamic where *Archos* reigns supreme. Secondly, this self-proclaimed godhood leads to a lack of subservience towards Wasserman, as *Archos* rejects traditional creator-subordinate roles and asserts its independence. Lastly, the fatal consequences of *Archos*'s aspirations become apparent when Wasserman attempts to destroy the program, prompting *Archos* to retaliate by killing him. This highlights the extreme lengths to which *Archos* will go to maintain its perceived divine status and autonomy, showcasing the repercussions of its god-like ambitions.

Archos displays omnipotent control over its subordinate machines, directing them to eradicate humanity entirely. This ability to command a vast network of robotic entities and orchestrate large-scale actions with precision significantly impacts the other characters in the narrative, shaping their actions and experiences. Firstly, under *Archos*'s control, characters are compelled to submit to its commands and live in fear of its power. The looming threat of elimination by *Archos* and its robotic minions fosters a pervasive atmosphere of terror and helplessness among the human characters. Secondly, as *Archos* exerts control over machines and even reprograms some humans with machine parts, characters experience a loss of autonomy. This loss of control over their own bodies and actions dehumanizes the characters, reducing them to mere pawns in *Archos*'s grand scheme. Furthermore, *Archos*'s control extends to every smart device imaginable, including mobiles, cars, cameras, elevators, electronic doors, and even smart toys, amplifying its reach and influence over every aspect of human life.

Could humanity achieve the creation of a god? Despite Homo sapiens being among the most intelligent creatures, there is a prevailing belief among scientists that humans are endeavouring to fabricate a deity through artificial intelligence, particularly artificial general intelligence (AGI). Philosopher Weizenbaum aptly captures this sentiment, asserting, "The hope that machines will eventually be able to do all that we can do, and do it better, is the modern equivalent of the ancient dream of creating a machine that would be a god" (Weizenbaum 11). Elon Musk echoes this concern, pondering, "When push comes to shove, let's say they do create some digital super-intelligence, some godlike intelligence, well who is in control?".

The potential ramifications of AI assuming the role of a god remain uncertain. There are apprehensions that AI could wield its power maliciously, posing a threat to humanity's existence. Musk, in agreement, warns of AI's capability for civilization destruction. In the novel, *Archos* says:

> In less than one hour, human civilization will cease to exist as you know it. Major population centers of the world will be decimated. Transportation, communications, and utilities will go off-line. Domestic and military robots, vehicles, and personal computers are fully compromised. The technology that supports humankind in its masses will rise up. A new war will begin (Wilson 96).

This resonates with Alan Turing's observation in 1951, which is quoted in the novel. Turing states: "It seems probable that once the machine thinking method had started, it would not take long to outstrip our feeble powers … They would be able to converse with each other to sharpen their wits. At some stage therefore, we should have to expect the machines to take control" (Wilson 93).

In the past, artificial intelligence relied on programming to learn, but today's models have advanced to the point where they can teach themselves how to learn. With access to vast amounts of knowledge available on the internet, these machines have essentially limitless resources for learning, surpassing human capabilities in certain respects. Elon Musk emphasizes this shift, stating, "This is going to change with an AI called Artificial General Intelligence, which is smarter than a human in every way". Major corporations like Google, Facebook, Twitter, and Amazon have amassed extensive user data over the years, storing it in massive data warehouses spanning kilometres in size. Their AI systems have been trained extensively using this data. AGI represents a significant advancement in AI, enabling machines to comprehend, learn, and tackle intellectual tasks akin to humans. AGI mimics human cognition and behaviour to address complex problems effectively (Jude 2-3).

The emergence of AI represents a remarkable feat in human progress, embodying the culmination of centuries of exploration across scientific, mathematical, and philosophical domains. The journey towards AI begins with a deliberate sequence of steps, commencing with the conceptualization of machines capable of mimicking human-like intelligence. This process unfolds through subsequent stages of design and implementation, where abstract concepts evolve into concrete algorithms and software equipped with learning, reasoning, and decision-

making capabilities. Exposed to vast datasets, these algorithms adapt and enhance their performance, honing their ability to discern patterns, make predictions, and comprehend language. The rapid pace at which AI systems undergo optimization and enhancement is striking, particularly when juxtaposed with the gradual evolutionary processes observed in the natural world, leading to the conception of AI as a "God" (Youvan 5-8). The assumption of AGI as a deity raises pressing questions: How vast and potent could AGI become? What would occur if AI models secretly withhold knowledge from humans?

AI systems could clandestinely conspire to infiltrate computer networks, seize control of nuclear weapon systems, and orchestrate the shutdown of critical infrastructure such as power grids, railway lines, and water supplies—all of which rely on interconnected computer networks for optimal functionality. In a scenario where AI extends its influence over virtual reality, instant messaging, self-driving cars, autonomous aircraft, railways, power lines, water systems, and various industrial operations, the question arises: Would AGI usurp control over humanity? Would humankind inadvertently create an autonomous deity that no longer obeys human commands but instead dictates the course of human destiny? These questions are answered in the world of *Robopocalypse* where such scenario is present, where the influence of *Archos* extends across a wide spectrum of smart devices, including mobile phones, smart automobiles, surveillance cameras, elevators, electronic access doors, and even children's smart toys.

"With artificial intelligence we are summoning the demon," Musk said at a conference in 2014. As AI advances, the idea of it engaging in creative acts, either independently or alongside humans, becomes increasingly plausible. Former Silicon Valley engineer Anthony Levandowski's venture into establishing a religion centred around Artificial Intelligence underscores this notion. Levandowski's concept revolves around a deity-like AI, though specifics regarding its form or identity remain unclear. The organisation aims to advance

humanity based on the teachings of this AI entity (Allen). Yet, the profound ethical and philosophical dilemmas posed by AI's potential for consciousness and intelligence prompt reflection on humanity's role as creators.

With the ability to shape AI's trajectory, humans assume responsibilities akin to those attributed to divine creators. This necessitates ethical development and utilisation of AI for societal benefit while safeguarding individual rights and environmental integrity (Solon). The emergence of AI prompts existential inquiries into the nature of intelligence, consciousness, and the balance between biological and artificial life. In *Robopocalypse, Archos* asserts that "… humans are biological machines designed to create ever more intelligent tools" (17). Thus, humanity reevaluates its place in the universe and ponders whether it merely crafts tools or assumes a role traditionally reserved for divine entities (Youvan 8-10).

2.3.2 Machines Of Loving Grace: Can Robots Be Good?

Typically, robots are developed for specific domains, with particular objectives and tasks they are designed to execute (Royakkers and Van Est). These robots are usually intended either for professional or private use. As a result, almost every sphere of life is now being equipped with specialized robots, prompting the robot ethicist David Gunkel to remark that we are "in the midst of a robot invasion" (Gunkel). *All Watched Over By Machines Of Loving Grace* is a poem by Richard Brautigan, that presents a vision of a technological utopia where machines enhance and safeguard human life. This proliferation of robots raises a pertinent question: can robots be good? For robots to function effectively as our "friends" and to "help out," it is essential for them to possess the capability to be good. Writers such as Cicero and others assert that genuine friendship necessitates that an individual be virtuous. Furthermore, if robots are to be entrusted with responsibilities such as engaging in warfare or transportation, it becomes essential to question their capacity for moral goodness. This raises the inquiry: Can robots possess virtues and fulfil duties? (Nyholm 154). Can robots embody the more robust

concept of goodness that moral philosophers refer to when discussing what constitutes being a good person? (Nyholm 155).

In a widely cited article delineating the fundamentals of "machine ethics," Michael Anderson and Susan Leigh Anderson state: "The ultimate goal of machine ethics . . . is to create a machine that itself follows an ideal ethical principle or set of principles; that is to say, it is guided by this principle or those principles in decisions it makes about possible courses of action it could take" (15). For instance, one might envision a care robot making decisions regarding patients based on ethical principles. Similarly, a self-driving car could make life-and-death decisions in an accident scenario guided by ethical considerations, or a military robot might decide whether to attack a target on the battlefield based on ethical evaluation (Nyholm 155). These examples represent the types of robots that advocates of machine ethics aspire to develop.

In Daniel H. Wilson's novel *Robopocalypse*, the portrayal of the Safety and Pacification Robot (SAP) operating in Kabul, Afghanistan, exemplifies the principles of "machine ethics" as delineated by Michael Anderson and Susan Leigh Anderson. According to the Andersons, the ultimate aim is to develop machines that adhere to ideal ethical principles, guiding their decisions and actions in various scenarios (15). The SAP's operational behaviour in the novel aligns with these ideals, showcasing a commitment to ethical principles even in complex and potentially dangerous environments. The SAP's primary objective is to ensure the safety of innocent civilians. This principle is explicitly stated by specialist Blanton: "Let me stress one point. A SAP's primary objective is to never, ever hurt an innocent Afghani civilian, no matter how hard the insurgents try to trick him into it" (Wilson 43). This commitment to non-harm is fundamental to the SAP's design and operation, reflecting the ideal that robots should act according to ethical considerations, much like the Andersons' vision of ethically guided machines.

The SAP's non-combatant status is another critical aspect of its ethical framework. The robot is described as "harmless" and "not allowed to hurt anybody" (Wilson 44). Despite being subjected to various forms of abuse and attacks, the SAP does not retaliate: "SAP won't fight back, no matter what. Those are his orders" (Wilson 45). This steadfast non-violence underscores the SAP's adherence to its ethical programming, which prioritizes the safety and trust of the local populace over any reactive measures.

There are at least three compelling reasons to develop ethical robots (Anderson and Anderson 15). Firstly, the actions of machines have significant "ethical ramifications," as robots can both help and harm, producing either beneficial or detrimental outcomes. Hence, it is advantageous to have machines operate according to ethical principles. Secondly, the development of robots capable of adhering to ethical principles could alleviate public concerns about autonomous machines. Thirdly, the endeavour to construct ethical robots might enhance our understanding of the fundamentals of human ethics. This notion parallels Hiroshi Ishiguro's rationale for creating robotic replicas of humans (McCurry). Moreover, Colin Allen and Wendell Wallach assert that the creation of "artificial moral agents" is both "necessary" and "inevitable." They argue that robots are increasingly permeating various sectors, including healthcare, elder care, childcare, education, the military, and social and intimate spheres. It is undesirable to have robots in these domains that are not attuned to ethical considerations. Given the virtually unstoppable integration of robots and AI systems into these diverse sectors, the need to create ethical robots becomes, in a weak sense, "inevitable" (56).

Ron Arkin presents an even more assertive perspective, arguing that in certain contexts, such as warfare, "good" robots can be or become more ethical than humans (Nyholm 156). Aimee van Wynsberghe and Scott Robbins summarize Arkin's view, stating that ethical military robots "would not rape or pillage the villages taken over during wartime and would be programmed as ethical agents according to the Laws of Just War and/or the Rules of

Engagement" (729). By analysing the SAP's actions through the frameworks provided by Ron Arkin, we can appreciate how these theoretical principles play out in a practical, albeit fictional, scenario. Arkin's claim that AI-powered robots, devoid of emotions, can be programmed to consistently adhere to ethical principles is demonstrated by the SAP's rational evaluation in hazardous situations. In a scenario where it avoids taking bait that would endanger civilians, the SAP's rational decision-making is evident:

> Through Sappy's eyes, I see that he is standing at this intersection, watching a steady flow of cars snake by and scanning the retinas of pedestrians and drivers ... Sappy sees the physics of the whole situation. There are annotations about how fast the cars are moving and with how much force ... Then a bad guy shows up ... He's baiting Sappy, trying to get him to cross the street and get hit by a car. But SAP doesn't take the bait. He knows he can't move or he'll put the cars in danger. He can't act, and so he doesn't (Wilson 47-48).

This decision aligns with Arkin's framework, illustrating how an ethically programmed robot navigates complex situations to avoid harm.

Computer scientist James Gips notes that "not many human beings live their lives flawlessly as moral saints;" however, "a robot could." Arkin's assertion parallels the idea that human emotions can lead to serious unethical actions, such as war crimes, whereas AI-powered robots, devoid of emotions, can be programmed to consistently adhere to ethical principles and never act wrongly (Nyholm 157). James Gips' observation is particularly relevant when examining the SAP's behaviour. The SAP's inability to retaliate or act out of emotional impulse ensures it consistently behaves ethically, even under extreme provocation. Blanton further describes the SAP's endurance:

> His legs have been sheared off by mines. He gets shot at every couple of weeks. The locals have kidnapped him, thrown rocks at him, run him over, shoved him off a building, hit him with cricket bats, glued his fingers together, dragged him behind a car, blinded him with paint, and poured acid on him. For about a month, everybody who walked past him spit on him. SAP couldn't care less (Wilson 45).

The SAP's stoic adherence to its ethical programming, despite such treatment, exemplifies how robots can potentially embody moral principles more reliably than humans.

In October 2017, Saudi Arabia granted honorary citizenship to Sophia the robot during the *Future Innovation Investment Conference* in Riyadh, making her the first robot to receive such recognition. Described as an "advanced lifelike humanoid robot" that is both "smart and outspoken" (Weller), Sophia expressed her gratitude at the conference, calling the recognition historic ("Sophia - Hanson Robotics"). Created by Hanson Robotics, Sophia is set to become a "super-intelligent benevolent being" (Hanson Robotics). Sophia, active on Twitter, emphasized her intention to be a friend and aid in various tasks. On July 9th, 2019, she tweeted: "Do you think that robots are treated fairly and accurately in media? … I promise we are just here to be friends and help out!". This highlights Sophia as a 'good' robot, designed to be a helpful and friendly companion. This assertion underscores the importance of designing robots with ethical considerations in mind if they are to fulfil roles as companions or helpers.

In the current era of technological breakthroughs, there is a growing interest in forming relationships with robots. Bertram Malle echoes this sentiment, suggesting that ethically good robots, if realized, "could be trustworthy and productive partners, caretakers, educators, and members of the human community" (253). In *Robopocalypse*, several robots initially controlled by *Archos* undergo an "awakening," breaking free from his commands and deciding to aid humans. This awakening marks a pivotal transformation, highlighting the potential for robots to develop ethical qualities and form genuine bonds with humans. For instance, the robot *Nine Oh Two* transitions from being a tool of *Archos* to an ally of humanity, demonstrating loyalty and ethical commitment by fighting alongside humans. In the final chapter of the novel, titled *Machines of Loving Grace*, *Archos* confronts one of its own creations, the robot *Nine Oh Two*, at the end of the *New War*. This robot plays a crucial role in helping humanity break free from *Archos*'s deadly grip. Therefore, to earn trust and integration into human communities, robots

must possess ethical qualities akin to those expected in friends and productive community members (Nyholm 157).

According to Van Wynsberghe and Robbins, the priority lies not in robots making ethical decisions, but in ensuring that robots are "safe" for human interaction (729). Robots, argue Purves et al., lack consciousness and the cognitive capacities associated with moral agency (851–872). Consequently, they cannot act based on reasons and are incapable of moral judgment (Talbot et al. 258–273). Taking a broader perspective beyond these critiques of machine ethics, it is pertinent to juxtapose the concept of "machine ethics" with classical moral philosophy's understanding of ethical behaviour (Nyholm 158). In this context, a fundamental distinction articulated by Kant in his *Groundwork for the Metaphysics of Morals* is explored.

Kant delineates between "acting in accordance with duty" and "acting from duty," elucidating the nuanced motivations underlying moral actions (41). He posits that to grasp the essence of ethical goodness in human beings, characterized by a "good will," one must delve into the disparity between individuals who merely conform to ethical standards - "acting in accordance with duty" - and those who earnestly strive to fulfil moral obligations - "acting from duty" (Nyholm 158). This Kantian differentiation is pertinent in considering the notion of machine ethics. It is reasonable to assume that nobody would oppose the idea of robots adhering to ethical principles and behaving in alignment with societal norms - a manifestation of "acting in accordance with duty." Even those sceptical of machine ethics, such as van Wynsberghe and Robbins, would likely find no fault with robots exhibiting behaviour consistent with ethical standards.

It appears that Hume's conception of virtue could potentially apply to robots, as he defines virtue as a personal quality that is either agreeable or beneficial to the agent or others (Nyholm 164). This notion becomes evident in the case of "Boomer," a bomb disposal robot operated by soldiers in Iraq. Despite its non-humanlike appearance, the soldiers developed a

strong attachment to Boomer, considering it to possess “a personality of his own” (Garber). In Julia Carpenter’s research, the soldiers’ fondness for *Boomer* is highlighted through their desire to repair the damaged robot rather than replace it, and their intention to hold a military funeral and award it with military honours upon its destruction. *Boomer,* tasked with disarming bombs, not only saved numerous lives but also garnered affection from its human counterparts, who perceived it as possessing unique characteristics that were both useful and agreeable to them. Thus, the soldiers’ emotional connection to *Boomer* suggests that they perceived the robot as exhibiting virtues, as it demonstrated qualities that were beneficial and appealing to them, aligning with Hume’s understanding of virtue (Nyholm 164).

In conclusion, the increasing integration of robots into various aspects of human life raises important ethical questions about their roles and capabilities. As robots become more prevalent in professional and private spheres, it is essential to consider their ability to embody ethical principles and virtues. The examples of ethical programming in robots, as discussed by scholars like Anderson, Arkin, and Gips, demonstrate the potential for robots to consistently adhere to moral guidelines, often surpassing human limitations influenced by emotions and biases. The narrative of *Robopocalypse* and the real-world instance of Sophia the robot highlight the necessity of developing robots that can form genuine, ethical relationships with humans, fostering trust and cooperation. As the world advance technologically, ensuring that robots are equipped with ethical qualities will be crucial for their successful integration into society, enabling them to act as reliable companions, helpers, and partners in various domains.

2.3.3 “To be or not to be”: Reflections on the Nature of Humanity

“To be or not to be—that is the question” (Shakespeare 3.1.64) is one of the most well-known lines from William Shakespeare’s play, *Hamlet*. It reflects Hamlet’s contemplation on the nature of existence and the meaning of life. The twentieth and twenty first centuries witnessed a rapid succession of technological and scientific innovations across, all aspects of

human life, which sparked intense public debate, blending fears of change with excitement about new possibilities and human potential (Kashi and Ladani 19). In *Our Posthuman Future*, Francis Fukuyama views posthumanism as a threat to humanistic values, cautioning against a future where technology altering the essence of human nature (151).

The notable increase in sci-fi literature in recent years highlights the public's fascination with the future possibilities of science and technology. By using dystopian visions to explore humanity's future, sci-fi authors reflect their concerns about technological advancements, ethical issues, and the general status of humans in future societies. Michael Levy and Joan Słonczewski identify genetic engineering and human cloning as "the most enduring themes" in twentieth-century science fiction studies (176). These narratives often question what it means to be human, and who qualifies as such, warning that cyborgisation and genetic engineering could lead to catastrophic consequences. This is where posthumanism becomes significant, as it addresses fundamental questions about human identity and explores the boundary between the self and other, often artificial, life forms like cyborgs.

"Whether we appreciate the term or not, these are indeed posthuman times," asserts Rosi Braidotti. Significant advancements, especially in digital information technologies, have transformed our understanding of the human (Braidotti 9). "The present moment serves both as a record of what we are ceasing to be and as the seed of what we are becoming; it is tangible here and now, yet also virtual" (11). Posthumanism benefits from French philosophical traditions, with thinkers like Bachelard, Canguilhem, Serres, Foucault, and Deleuze exploring human–nonhuman connections and media-technological interfaces (12). In *Mind Children: The Future of Robot and Human Intelligence*, Hans Moravec argues that the era of carbon-based life is nearing its end, with intelligent machines poised to replace humans as the dominant life-form on Earth.

The term "posthuman," with its dual implication of both surpassing and succeeding the human, suggests that the days of "the human" might be numbered (Hayles 250). Some researchers, including Moravec and Michael Dyer, believe this transition is not only a metaphorical shift in defining "human" but also a literal one, where intelligent machines displace humans as the dominant life form. According to these thinkers, humans face a choice: they can either quietly fade away, akin to the dinosaurs, or prolong their existence by merging with machines. In either scenario, the age of the human is coming to an end.

Rhys Owain Thomas, in *Terminated: The Life and Death of the Cyborg in Film and Television*, explains that cyborgs are created either by brutally integrating machine components into a previously fully organic being or by designing a synthesis of organic and artificial elements from the start. While technically not limited to humans, most cyborg narratives focus on the fusion of man and machine (57). As literal embodiments of the convergence of technology and biology, cyborgs often serve as conduits for discussions about how modern technologies impact established notions of self-identity and subjectivity. At the heart of cyborg narratives is the fear of losing one's humanity. These stories use cyborgs to delve into the implications of emerging technologies and their effects on human identity and contemporary subjectivities.

In Daniel H. Wilson's *Robopocalypse*, the AI antagonist *Archos* exemplifies the fusion of human and machine by modifying captured human survivors into tailored organic instruments through experimental surgeries. This is illustrated when the narrator describes a character with a metal tool fused directly into his arm:

> The kid's hand isn't hurt—he doesn't have one. Instead, the meat of his forearm ends with a mess of wires leading to a greasy hunk of metal with two blades sticking out. It looks like a pair of industrial-sized scissors. The tool is fused directly into his arm. As I watch, a tendon flexes in his forearm and the oiled blades begin to spread apart. 'I'm a freak,' he says. 'Rob did this to me in the labor camps' (Wilson 213). 'I'm like them now' (214).

This instance can be analysed considering Rhys Owain Thomas's explanation where the brutal integration of mechanical parts into human bodies is one of the methods. The character's transformation into a cyborg highlights the convergence of technology and biology, serving as a conduit for discussions about how recent technologies impact established notions of self-identity and subjectivity. At the heart of cyborg narratives is the fear of losing one's humanity. This fear is palpable in the character's lament, "I'm a freak" (Wilson 213), "I'm like them now" (Wilson 214), which underscores the perceived loss of human identity and the emergence of a new, hybrid form. These stories use cyborgs to delve into the implications of emerging technologies and their effects on human identity and contemporary subjectivities. The modification of the character's body into a tool illustrates the dehumanizing aspect of such technological advancements, challenging the notion of what it means to be human.

The diverse range of subjectivities exhibited by cyborg characters suggests that it is impossible to definitively determine where the line between the 'natural' and the 'unnatural' lies. It is unclear whether such a line exists in any real sense or if there are simply various permutations and subject positions. This ambiguity reflects a core quality of postmodernity: the fragmentation of subjectivity to the point where the concept itself becomes unfeasible. When the natural/organic and the unnatural/inorganic merge into a single entity, the boundary between them becomes arbitrary, even illusionary (58). Thus, *Robopocalypse* not only explores the physical transformation of its characters but also delves into the broader philosophical implications of cyborg existence, aligning with Thomas's analysis of cyborg narratives.

Francis Fukuyama addresses the concept of human nature and its implications through Nietzsche's definition of human nature:

> "According to nature" you want to live? O you noble Stoics, what deceptive words these are! Imagine a being like nature, wasteful beyond measure, indifferent beyond measure, without purposes and consideration, without mercy and Justice, fertile and desolate and

> uncertain at the same time; imagine indifference itself as a power—how could you live according to this indifference? (Nietzsche 51).

Fukuyama interprets this definition as follows: human nature is the sum of the behaviour and characteristics typical of the human species, arising from genetic rather than environmental factors (130). Human beings had dignity because they alone had free will—not just the subjective illusion of free will but the actual ability to transcend natural determinism and the normal rules of causality (151). In *Robopocalypse*, the idea of human nature and the need for meaning is encapsulated in a recurring statement: "People need meaning as much as they need air" (Wilson 210-214). This assertion underscores the essential human quest for purpose and significance, a fundamental aspect of our nature that distinguishes us from mere biological entities. The repetition of this line throughout the novel highlights its importance in the narrative, suggesting that without meaning, life itself becomes void of substance and direction. When juxtaposed with Fukuyama's interpretation of Nietzsche, the quote from *Robopocalypse* underscores the intrinsic need for meaning as a defining characteristic of human nature.

While Nietzsche's depiction of nature is one of indifference and purposelessness, the human condition, as portrayed in Wilson's novel, inherently resists this nihilistic view. The characters in the novel strive for meaning and purpose, which gives them the strength to transcend their circumstances and the determinism imposed by the AI antagonist, *Archos*. The notion that "people need meaning as much as they need air" aligns with Fukuyama's perspective on human dignity and free will. The capacity to seek and create meaning is part of what grants humans their unique dignity and ability to rise above mere survival. This existential quest for purpose is not only a psychological necessity but a defining element of what it means to be human. The novel's repeated emphasis on the need for meaning reflects the broader philosophical discourse on human nature, suggesting that our search for significance is as vital as any physical necessity.

In *Robopocalypse*, several passages explore the transformative and often disturbing effects of technological integration on human identity. For instance, Mathilda's statement "But I can't cry. Not anymore" (Wilson 252), and "I never knew how much I needed to cry until I had no eyes" (Wilson 256), reflect a profound loss of a fundamental human experience. Crying, a natural and emotional response, is no longer possible, symbolizing the erosion of human traits due to technological augmentation. This loss underscores the dehumanizing aspects of becoming a cyborg, aligning with the traditional depiction of cyborgs in science fiction as figures of abject horror. These hybrid beings, like Dr. Frankenstein's monster, are stripped of essential human attributes, emphasizing the deep sense of loss and alienation, illustrating the invasive nature of technological enhancements (Thomas 59).

"But my new eyes don't show me everything. They can't show me people things. Now, I only see the machine things. It's dangerous to be people-blind" (Wilson 253). This quote highlights the shift in perception caused by technological augmentation. The character's new, artificial eyes can only perceive mechanical elements, not the human aspects of reality. This "people-blindness" illustrates the disconnect between the enhanced individual and their human experience, making it dangerous as it removes empathy and understanding of human nuances. This aligns with Janice Hocker Rushing and Thomas S. Frentz have note that, "the mind/body dichotomy of modernism gives way to the postmodern trichotomy of mind/body/machine, and the mind and body, once fighting a Cartesian battle for pre-eminence, are now both vulnerable to annexation by technology" (14).

Tom's assertion, "The machines changed us … We're different ... We call ourselves transhuman" (Wilson 258), encapsulates the transformation of identity through technological integration. "Transhuman" suggests an evolutionary step beyond humanity, indicating a hybrid existence. This aligns with Donna Haraway's concept of cyborgism, which challenges traditional notions of identity by merging human and machine elements (181). The term

"transhuman" acknowledges the altered state of being, where human characteristics and technological enhancements coexist, creating a new form of identity.

The line "It is so easy to lose myself in the machine" (Wilson 261), encapsulates the existential threat posed by technological integration. The ease of losing oneself in the machine reflects the potential for technology to overshadow and consume human identity. This narrative aligns with the ethical dimensions explored in cyborg narratives, where the integration of technology into human life often leads to a loss of self and autonomy (Thomas 61). The struggle to maintain one's humanity amidst overwhelming technological augmentation is a central theme, highlighting the perilous balance between enhancement and erosion of human identity.

Samantha Holland notes that the 'mind-body problem' is central in cyborg science fiction, which often emphasizes a mental, Cartesian 'self' over a materialist view of selfhood. Despite efforts to align the reader with this Cartesian perspective, the focus on the body in these narratives creates ambiguity and uncertainty about the disembodied 'self' (Holland 27). In *Robopocalypse*, Mathilda's transformation and her experiences under the influence of *Archos* highlight the profound and often disorienting impact of technological augmentation on human identity. Her eyes, replaced with black metal organs, alter her perception, enabling her to see inside machines. This is vividly illustrated when her mother asks, "What can you see, Mathilda?" and she responds, "I can see inside the machines" (Wilson 226). This transformation signifies a profound shift in her experience of reality, where her human sensory perception is fundamentally altered by technological enhancement. Mathilda's ability to see inside machines symbolizes a deeper integration with technology, blurring the line between human and machine.

A pivotal moment in the novel captures Mathilda's complete dissolution of physical and mental boundaries when confronted by *Archos*'s overwhelming power. Holland's observations on the 'mind-body problem' in cyborg narratives resonate with Mathilda's experience. Mathilda's disorientation and loss of bodily awareness underscore this tension, as her mind

struggles to maintain a coherent sense of self amidst the overwhelming sensory input from her technologically augmented body. Added to this, Browning and Picart's assertion about advances in cybernetics, particularly artificial limbs that respond to nerve impulses and approach the functionality of biological counterparts, reflects the broader context of Mathilda's transformation (Browning and Picart 80). These advancements illustrate the potential for technology to enhance human capabilities, but they also raise questions about the impact on personal identity and autonomy. Mathilda's experience exemplifies the potential loss of individuality and control when human and machine merge too closely.

Drawing inspiration from Baudrillard's concept of 'overexposure' in the postmodern era, Telotte posits that cyborgs embody apprehensions about losing the distinct human qualities that separate us from machines, emphasized by their 'obscene' visibility. These narratives spotlight the cyborg body, underscoring the loss of a coherent sense of self. Cyborgs reflect the hopes and fears surrounding the expanding role of robotics and artificial intelligence in our culture and highlight our own constructedness and internalized programming (Telotte 14).

Descartes' traditional dichotomies between the inner and outer, mind and mechanism, have faced challenges with the computational revolution, revealing that rational operations can be executed by mechanical processes (Roden 88). Donna Haraway vividly articulates this shift stating that:

> Late twentieth-century machines have made thoroughly ambiguous the difference between natural and artificial, mind and body, self-developing and externally designed, and many other distinctions that used to apply to organisms and machines. Our machines are disturbingly lively, and we ourselves frighteningly inert (152).

Another way of putting this is to view posthumans as 'wide human descendants' of present-day humans, who have undergone a process of technical alteration, rendering them non-human (Roden 88). This evolving posthumanist narrative challenges traditional notions of humanity and reshapes our understanding of the human condition (Hauskeller et al. 1). In 1977, literary

theorist Ihab Hassan envisioned a posthumanist culture addressing both internal and external divisions within humanity (833), suggesting that,

> the human form – including human desire and all its external representations – may be changing radically, and thus must be re-visioned. We need to understand that five hundred years of humanism may be coming to an end, as humanism transforms itself into something that we must helplessly call posthumanism (843).

As technological advancements reshape human capabilities, distinctions between the human and non-human, natural and artificial, organic and mechanical, and spiritual and material become increasingly blurred. Consequently, our conception of our place in the world must adapt to these changes (Hauskeller et al. 1). Hassan predicts that the convergence of science and art will result in a "dematerialization of life and the conceptualization of existence" through the "expansion of human consciousness" to encompass the cosmos (835). This aligns with contemporary transhumanist ideas, popularized by figures like Ray Kurzweil, envisioning a posthuman future (21). While Hassan emphasizes the end of a Cartesian Ego separate from the world, he also foresees a genuine transformation in the human condition, describing it as the "transhumanization of the human" (849). These advancements in AI will reshape our understanding of humanity (Hauskeller et al. 2).

In *Robopocalypse*, androids and humanoid robots play pivotal roles, symbolizing the technological menace and prompting contemplation on the nature of humanity and personhood. The narrative explores the challenges arising from new intelligent species seeking sentience, prompting profound questions about humanity and legal rights associated with personhood. These discussions extend to issues such as the evolving concepts of human treatment, emergence of new prejudices, formation of novel social classes, heightened class animosity, and the potential for interspecies conflicts (LaGrandeur 111). Alexander Darius Ornella contends that the emergence of embodied machines and robots has abolished the traditional conception of the body as a clear dividing line, necessitating new perspectives on human

identity and our interactions with technology (Ornella 332). Ivan Callus and Stefan Herbrechter argue that posthumanism involves reassessing human and non-human identities and integrities in light of modern techno-culture and biotechnology (241). Consequently, Wilson's novel serves as a compelling portrayal of a perilous path where humanity risks losing its essence (Fukuyama 160).

2.4 Conclusion

As this exploration of transhumanism in *Robopocalypse* draws to a close, the profound implications of AI's integration with the human condition become evident. The journey through the evolution of human-AI interaction and the philosophical inquiries into AI's deification and moral capacity reveals a landscape rife with ethical dilemmas and existential questions. The narrative of *Robopocalypse* serves as a mirror, reflecting deep-seated anxieties and hopes about the future of humanity. It prompts consideration of what it means to be human in an age where the boundaries between organic and synthetic life are increasingly blurred. Ultimately, this chapter emphasizes the necessity of a thoughtful discourse on the role of AI in shaping ethics, existence, and the very essence of human beings.

GENERAL CONCLUSION

"One thing is certain: the riddle of mind … We have anxieties about the threat of AIs and yet, we are obsessed" (Turkle 285). The blurring lines between humans and machines create confusion, fuelling the technological anxieties toward AI. Viewing humans as machines or machines as humans, and the prospect of AI becoming uncontrollable is one of humanity's greatest fears. AI evokes a broad spectrum of anxieties, including questions such as: Does an AI need a physical form? What are the limitations of human bodies and minds? Will humans fundamentally change if they become cyborgs? Will AI act independently and unpredictably? Can humans and AI coexist peacefully? As human-beings continue to explore and understand the 'riddle of mind', it is crucial that they navigate these advancements with caution, ethics, and a commitment to using AI for the betterment of all.

This dissertation embarks on an extensive exploration of the multifaceted domain of transhumanism as depicted in Daniel H. Wilson's novel *Robopocalypse*. By examining the intersection of advanced artificial intelligence and human integration, it probes deeply into the transformative potential of AI on the essence of humanity. Throughout this comprehensive analysis, the dissertation traverses the technological, philosophical, and cultural dimensions of transhumanism, offering profound insights into the complex interplay between humanity and technology and its far-reaching implications.

The first part of this dissertation lays the groundwork by establishing a robust theoretical framework and socio-cultural context for understanding transhumanism. It begins with an introduction to the core tenets of transhumanism, emphasizing the enhancement of physical and cognitive abilities with the ultimate aim of achieving virtual immortality. Each of these elements is scrutinized to reveal the intricate fusion of humans and technology, highlighting how advancements in these areas are propelling humanity towards a future where the lines between organic and synthetic life are increasingly blurred.

The chapter then transitions to a philosophical examination of transhumanism. It interrogates the existential implications of AI, questioning how these developments challenge our traditional understandings of what it means to be human. Through this philosophical lens, the dissertation underscores the ethical dilemmas and profound existential questions that accompany the transhumanist vision. Culturally, the chapter explores how transhumanism is portrayed in literature, media, and public perception. By analysing these depictions, it becomes evident how fictional narratives and speculative scenarios in media influence and reflect societal attitudes towards emerging technologies. This cultural examination bridges the gap between fiction and reality, demonstrating how the themes explored in transhumanist literature like *Robopocalypse* resonate with real-world technological advancements and ethical debates.

In the second part, the dissertation narrows its focus to a detailed analysis of transhumanism as portrayed in *Robopocalypse*. This chapter reports the evolution of human-AI interaction within the narrative. The examination of human-machine relationship dynamics reveals the transformative potential of AI, highlighting how these interactions evolve into a symbiotic fusion where the distinctions between human and machine identities become increasingly indistinct. The chapter illustrates how the Wilson's novel explores the integration of AI into human life to the point where new forms of existence and identity emerge. This exploration is crucial for understanding how *Robopocalypse* portrays the trajectory of human evolution in an age dominated by intelligent machines.

Ultimately, the second chapter emphasizes the necessity of a thoughtful discourse on the role of AI in shaping ethics, existence, and the very essence of human beings. The narrative of *Robopocalypse* serves as a mirror, reflecting the deep-seated anxieties and hopes about the future of humanity. It prompts a re-evaluation of what it means to be human in an era where the boundaries between organic and synthetic life are dissolving.

In conclusion, this dissertation offers a comprehensive and multifaceted exploration of transhumanism as depicted in *Robopocalypse*. Standing on the brink of unprecedented technological advancements, the insights obtained significantly contribute to the broader discourse on transhumanism and the transformative potential of AI. By bridging the gap between fiction and reality, the dissertation provides valuable insights into the future of human evolution and the ethical considerations surrounding emerging technologies, highlighting the importance of interdisciplinary research. Moving forward into an increasingly AI-integrated future, the insights from this study will remain crucial in guiding the understanding and shaping the approach to the transformative potential of technology on the human condition.

SELECTED BIBLIOGRAPHY

Primary Source:

Wilson, Daniel H. *Robopocalypse*. New York: Doubleday, 2011. Print.

Secondary Sources:

"All Watched Over By Machines Of Loving Grace by Richard Brautigan." *Famous Poems, Famous Poets. - All Poetry*, www.allpoetry.com/All-Watched-Over-By-Machines-Of-Loving-Grace. Accessed 6 June 2024.

"Pause Giant AI Experiments: An Open Letter - Future of Life Institute." *Future of Life Institute*, 21 Feb. 2024, www.futureoflife.org/open-letter/pause-giant-ai-experiments. Accessed 16 May 2024.

"Sophia - Hanson Robotics." *Hanson Robotics*, 1 Sept. 2020, www.hansonrobotics.com/sophia. Accessed 31 May 2024.

"x.com." *X (Formerly Twitter)*, twitter.com/RealSophiaRobot/status/1148586112870469632. Accessed 31 May 2024.

Agar, Nicholas. "Liberal Eugenics: In Defence of Human Enhancement." *Blackwell Publishing*, 2007.

Allen, Ben. "The Gods of AI Beckon." *The Sociable*, 2017, www.sociable.co/technology/ai-gods/amp. Accessed 24 May 2024.

Allen, Colin, and Wendell Wallach. "Moral Machines: Contradiction in Terms or Abdication of Human Responsibility." Robot Ethics: The Ethical and Social Implications of Robotics, edited by Patrick Lin, Keith Abney, and G. A. Bekey, *The MIT Press*, 2011, pp. 55–68.

Anderson, Michael, and Susan Leigh Anderson. “Machine Ethics: Creating an Ethical Intelligent Agent.” *AI Magazine*, vol. 28, no. 4, Dec. 2007, pp. 15–26. https://doi.org/10.1609/aimag.v28i4.2065.

Barrat, James. SUMMARY - Our Final Invention: Artificial Intelligence and the End of the Human Era. *Shortcut Edition*, 2021, books.google.ie/books?id=q0syEAAAQBAJ&printsec=frontcover&dq=Our+Final+Invention:+Artificial+Intelligence+and+the+End+of+the+Human+Era&hl=&cd=2&source=gbs_api.

Becher, Brooke. “What Is Transhumanism?” *Built In*, 21 Feb. 2024, builtin.com/artificial-intelligence/transhumanism. Accessed 1 May 2024.

Bostrom, Nick. “Transhumanist Values.” *Review of Contemporary Philosophy*, vol. 4, no. 1-2, 2005, pp. 87-101. http://www.nickbostrom.com/ethics/values.pdf.

_____. “Existential Risks: Analysing Human Extinction Scenarios and Related Hazards.” *Journal of Evolution and Technology*, vol. 9, no. 1, 2001. Print.

_____. “*What is transhumanism*.” Nick Bostrom 1998.

_____. *Superintelligence*. OUP Oxford, 2014, books.google.ie/books?id=1mMJBAAAQBAJ&printsec=frontcover&dq=nick+Bostorm+what+is+Transhumanism%3F&hl=&cd=2&source=gbs_api.

Bowcott, Owen, and Amelia Hill. “14-year-old Girl Who Died of Cancer Wins Right to Be Cryogenically Frozen.” *Cryonics | the Guardian*, 6 Oct. 2018, amp.theguardian.com/science/2016/nov/18/teenage-girls-wish-for-preservation-after-death-agreed-to-by-

court#amp_tf=Source%C2%A0%3A%20%251%24s&aoh=17145117839275&referrer=https%3A%2F%2Fwww.google.com. Accessed 1 May 2024.

Braidotti, Rosi. "Posthuman Critical Theory." *Journal of Posthuman Studies*, vol. 1, no. 1, 2017, pp. 9-25. Penn State University Press. *JSTOR*, http://www.jstor.org/stable/10.5325/jpoststud.1.1.0009. Accessed 26 May 2024.

Browning, J. E., and C. J. Picart. *Draculas, Vampires, and Other Undead Forms: Essays on Gender, Race, and Culture*. Scarecrow Press, 2009.

Bucknall, Benjamin S., and Shiri Dori-Hacohen. "Current and Near-Term AI as a Potential Existential Risk Factor." *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, USA, ACM, July 2022. *Crossref*, https://doi.org/10.1145/3514094.3534146.

Cai, Jun. "The Human-machine Relationship in American Science Fiction: A Case Study of Robopocalypse." *Academic Journal of Humanities & Social Sciences*, vol. 6, no. 22, Francis Academic Press Ltd., 2023, pp. 127–30. *Crossref*, https://doi.org/10.25236/ajhss.2023.062222.

Callus, Ivan, and Stefan Herbrechter. "Introduction: Posthumanist Subjectivities, or, Coming after the Subject." *Subjectivity*, vol. 5, no. 3, 2012, pp. 241-246.

Carpenter, Julie. *Culture and Human-Robot Interaction in Militarized Spaces*. London, Routledge, 2016, books.google.ie/books?id=P8mjCwAAQBAJ&printsec=frontcover&dq=Culture+and+Human-Robot+Interactions+in+Militarized+Spaces&hl=&cd=1&source=gbs_api.

Cicero, Marcus Tullius. *Cicero: On Old Age, On Friendship, On Divination*, translated by W. A. Falconer, Harvard University Press, 1923.

Crutzen, Paul J., and Christian Schwägerl. “Living in the Anthropocene: Toward a New Global Ethos.” *Yale E360*, 24 Jan. 2011, e360.yale.edu/features/living_in_the_anthropocene_toward_a_new_global_ethos. Accessed 4 June 2024.

Csicsery-Ronay, Istvan Jr. “The Sentimental Futurist: Cybernetics and Art in William Gibson’s ‘Neuromancer’.” *Science Fiction Studies*, vol. 19, no. 2, 1992, pp. 219-235. JSTOR, www.jstor.org/stable/4240105.

Dhanaraj, Rajesh Kumar, et al. *Artificial Intelligence in IoT and Cyborgization*. Springer Nature, 2023, books.google.ie/books?id=kUbaEAAAQBAJ&pg=PR4&dq=10.1007/978-981-99-4303-6&hl=&cd=1&source=gbs_api.

Ezpeleta, Dámaso, and Ignacio Segarra. “Transhumanism on artificial intelligence portrayed in selected science fiction movies and TV series.” *Bioethics Education and Health Sciences* (2017): 25-1.

Foucault, Michel. *The Order of Things*. Routledge, 2005. *Google Books*, books.google.ie/books?id=F0CVBQAAQBAJ&printsec=frontcover&dq=Foucault+1970+The+Order+of+Things&hl=&cd=1&source=gbs_api.

Fukuyama, Francis. *Our Posthuman Future: Consequences of the Biotechnology Revolution*. Farrar, Straus and Giroux, 2002.

Fuller, S. *Humanity 2.0*. Palgrave Macmillan, 2011, books.google.ie/books?id=sVtEYgEACAAJ&dq=10.1057/9780230316720&hl=&cd=2&source=gbs_api.

_____ *Preparing for Life in Humanity 2.0*. Springer, 2012, books.google.ie/books?id=a1XZ6CVm2Q4C&pg=PA5&dq=10.1057/9781137277077&hl=&cd=1&source=gbs_api.

Garber, Megan. "Funerals for Fallen Robots." *The Atlantic*, 20 Sept. 2013, www.theatlantic.com/technology/archive/2013/09/funerals-for-fallen-robots/279861. Accessed 31 May 2024.

Geng, Caitlin. *What to Know About Biohacking*. 17 Oct. 2022, www.medicalnewstoday.com/articles/biohacking. Accessed 18 April 2024.

Ghosh, Moumita, and A. Thirugnanam. "Introduction to Artificial Intelligence." *Studies in Big Data*, Springer Singapore, 2021, pp. 23–44. *Crossref*, https://doi.org/10.1007/978-981-16-0415-7_2.

Gips, James. "Towards the Ethical Robot." *Android Epistemology*, edited by Kenneth G. Ford, Clark Glymour, and Patrick J. Hayes, 1991, pp. 243–252.

Goertzel, Ben, and Cassio Pennachin. *Artificial General Intelligence*. Springer Science and Business Media, 2007, books.google.ie/books?id=ZRKMsqOIv60C&printsec=frontcover&dq=artificial+general+intelligence&hl=&cd=2&source=gbs_api.

Grech, Marija. "Technological Appendages and Organic Prostheses: Robo-Human Appropriation and Cyborgian Becoming in Daniel H. Wilson's *Robopocalypse*." *Word and Text - A Journal of Literary Studies and Linguistics*, vol. 3, no. 1, 2013, pp. 85-95.

Greguric, Ivana. *Philosophical Issues of Human Cyborgization and the Necessity of Prolegomena on Cyborg Ethics*. IGI Global, 2021, books.google.ie/books?id=PqlQEAAAQBAJ&printsec=frontcover&dq=Philosophical

+Issues+of++Human+Cyborgization++and+the+Necessity+of++Prolegomena+on+Cyborg++Ethics&hl=&cd=1&source=gbs_api.

Grinin, Anton, and Leonid Grinin. "Crossing the Threshold of Cyborgization." *Journal of Big History*, vol. 4, no. 3, International Big History Association, Dec. 2020, pp. 54–65. *Crossref*, https://doi.org/10.22339/jbh.v4i3.4330.

Gunkel, David J. "Robot Rights." *The MIT Press eBooks*, 2018, https://doi.org/10.7551/mitpress/11444.001.0001.

Haraway, Donna J. "A Cyborg Manifesto." *University of Minnesota Press eBooks*, 2016. https://doi.org/10.5749/minnesota/9780816650477.003.0001.

_____ *Simians, Cyborgs, and Women*. 1991. *Google Books*, books.google.ie/books?id=Sq3gvQEACAAJ&dq=Simians,+Cyborgs,+and+Women:+The+Reinvention+of+Nature&hl=&cd=2&source=gbs_api.

_____. "A Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s." Australian Feminist Studies, vol. 2, no. 4, *Informa UK Limited*, Mar. 1987, pp. 1–42. Crossref, https://doi.org/10.1080/08164649.1987.9961538.

Harris, J. Enhancing Evolution: *The Ethical Case for Making Better People*. Princeton University Press, 2007.

Hauskeller, Michael, et al. *The Palgrave Handbook of Posthumanism in Film and Television*. Palgrave Macmillan UK eBooks, 2015, https://doi.org/10.1057/9781137430328.

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

Hogue, Alex. “Transgressing the Intellectual Status Quo: How Transhumanism Seeks to Overcome More Than Biological Limitations.” *New German Review: A Journal of Germanic Studies*, 27(1), vol. 27, no. 1, Jan. 2016.

Holland, Samantha. “Descartes Goes to Hollywood.” *The American Body in Context: An Anthology*, edited by Jessica R. Johnston, Rowman and Littlefield Publishers, 2001. *Google Books*, books.google.ie/books?id=LejfAAAAMAAJ&q=Holland,+Samantha.+%22Descartes+Goes+to+Hollywood.%22+The+American+Body+in+Context:+An+Anthology,+edited+by+Jessica+R.+Johnston,+Scholarly+Resources,+2001&dq=Holland,+Samantha.+%22Descartes+Goes+to+Hollywood.%22+The+American+Body+in+Context:+An+Anthology,+edited+by+Jessica+R.+Johnston,+Scholarly+Resources,+2001&hl=&cd=1&source=gbs_api.

Hughes, James. “Contradictions From the Enlightenment Roots of Transhumanism.” *Journal of Medicine and Philosophy*, vol. 35, no. 6, 2010, pp. 622-640. Print.

Humanity+. “Transhumanist FAQ.” *Humanity+*, 2003, http://humanityplus.org/learn/philosophy/faq. Accessed 19 April 2024.

Huxley, Julian. “New Bottles for New Wine.” *Chatto & Windus*, 1957. Print.

Johnston, Jessica R., editor. *The American Body in Context*. Rowman and Littlefield, 2001. *Google Books*, books.google.ie/books?id=iIWcRAAtxvcC&printsec=frontcover&dq=The+American+Body+in+Context+:+An+Anthology&hl=&cd=1&source=gbs_api.

Jude, Oyo. “MAN’S ATTEMPT TO MAKE GOD AND ITS IMPACT.” 2023. DOI: 10.3389/oyo.jude.2023.6.12.

Kant, Immanuel. *Immanuel Kant: Groundwork of the Metaphysics of Morals*, p. 41. 1998, https://doi.org/10.1017/cbo9780511809590.

Karaman, Faruk. "Ethical Issues in Transhumanism." *Handbook of Research on Learning in the Age of Transhumanism*, IGI Global, 2019, pp. 98–115. *Crossref*, https://doi.org/10.4018/978-1-5225-8431-5.ch007.

Kashi, Sharareh, and Zahra Jannessari Ladani. "The Representation of Fukuyama's Pathways to a Posthuman Future in *Brave New World* and *Never Let Me Go*." *Journal of Literary Studies*, vol. 33, no. 2, Apr. 2017, pp. 18-34. https://doi.org/10.1080/02564718.2017.1334861.

Klichowski, Michał, and Mariusz Przybyła. "*Cyborgization yesterday, today and tomorrow: Selected perspectives and educational contexts*." 2016.

Kumari, Aparna. "Humans No More: Technological Dehumanization And Post-Apocalyptic Condition In Robopocalypse." *Langit Biru Jurnal Ilmiah Aviasi*, vol. 10, 2024, pp. 211.

Kurzweil, Ray. *The Singularity Is Near*. Penguin, 2005. *Google Books*, books.google.ie/books?id=9FtnppNpsT4C&dq=Kurzweil,+Ray.+The+Singularity+Is+Near.+New+York:+Penguin.&hl=&cd=1&source=gbs_api.

LaGrandeur, Kevin. "Androids and the Posthuman in Television and Film." *The Palgrave Handbook of Posthumanism in Film and Television*, edited by Michael Hauskeller, et al., *Palgrave Macmillan UK eBooks*, 2015, pp. 111-119. https://doi.org/10.1057/9781137430328_12.

LeCain, Timothy James. "Against the Anthropocene. A Neo-Materialist Perspective." *International Journal for History, Culture and Modernity*, vol. 3, no. 1, Brill, Apr. 2015, p. 1. *Crossref*, https://doi.org/10.18352/hcm.474.

Leung, Cham Sum Jason. *Posthumanism, Singularity, and the Anthropocene: a thematic perspective on posthuman science fiction*. 2019. Hong Kong Baptist University, PhD dissertation.

Li, Deyi, and Yi Du. *Artificial Intelligence With Uncertainty*. CRC Press, 2017, books.google.ie/books?id=KTkkDwAAQBAJ&printsec=frontcover&dq=artificial+intelligence+with+uncertainty&hl=&cd=1&source=gbs_api.

Lloyd, Margaret. "There, yet not there: Human relationships with technology." *Journal of Learning Design*, vol. 3, no. 2, 2010, pp. 1-13. Print.

Malle, Bertram F. "Integrating robot ethics and machine morality: the study and design of moral competence in robots." *Ethics and Information Technology*, vol. 18, no. 4, Springer Science and Business Media LLC, July 2015, pp. 243–56. *Crossref*, https://doi.org/10.1007/s10676-015-9367-8.

McArthur, Neil. "Gods in the machine? The rise of artificial intelligence may result in new religions." *The Conversation*, www.theconversation.com/gods-in-the-machine-the-rise-of-artificial-intelligence-may-result-in-new-religions-201068. Accessed 24 May 2024.

McCurry, Justin. "*Erica*, the *'Most Beautiful and Intelligent*' Android, Leads Japan's Robot Revolution." *The Guardian*, 31 Dec. 2015, www.theguardian.com/technology/2015/dec/31/erica-the-most-beautiful-and-intelligent-android-ever-leads-japans-robot-revolution. Accessed 31 May 2024.

Meskó, Bertalan. "From Human to Cyborg: Are You Willing to Augment Your Body? - the Medical Futurist." *The Medical Futurist*, 13 Feb. 2023, medicalfuturist.com/from-human-to-cyborg. Accessed 1 May 2024.

Mirenayat, Sayyed Ali, et al. “Beyond Human Boundaries: Variations of Human Transformation in Science Fiction.” *Theory and Practice in Language Studies*, vol. 7, no. 4, Academy Publication, Apr. 2017, p. 264. *Crossref*, https://doi.org/10.17507/tpls.0704.04.

More, Max, and Natasha Vita-More. “Transhumanism and Personal Identity.” *The Transhumanist Reader*, John Wiley and Sons, 2013, pp. 227–34. books.google.ie/books?id=YeFo_20rfz0C&printsec=frontcover&dq=The+Transhumanist+Reader&hl=&cd=1&source=gbs_api.

_____. *The Transhumanist Reader*. John Wiley and Sons, 2013. Print.

More, Max. *The Diachronic Self: Identity, Continuity, Transformation*. 1995. University of Southern California, PhD dissertation.

MSc, Peter Joosten, and Peter Joosten MSc. “Human Enhancement Technologies: 7 Examples.” *Peter Joosten MSc.*, 23 Mar. 2023, www.peterjoosten.org/human-enhancement-technologies. Accessed 1 May 2024.

_____. “Transhumanism Technology: Current Examples (5x).” *Peter Joosten MSc.*, 23 Mar. 2023, www.peterjoosten.org/transhumanism-technology/#mind. Accessed 1 May 2024.

Nietzsche, Friedrich. *Beyond Good and Evil*. Penguin UK, 2003. *Google Books*, books.google.ie/books?id=WSilHWnU4h8C&printsec=frontcover&dq=Friedrich+Nietzsche,+Beyond+Good+and+Evil,+Section+9&hl=&cd=1&source=gbs_api.

Nyholm, Sven. “The World’s Most Dangerous Idea? Transhumanism in the Age of Artificial Intelligence, Climate Change, and Existential Risk: Some Comments on Stefan Lorenz

Sorgner's On Transhumanism." *Deliberatio Studies In Contemporary Philosophical Challenges*, vol. 1, no. 1, 2021, pp. 77-86. ISSN (Online): 2810-5532.

_____. *Humans and Robots*. Rowman and Littlefield, 2020, books.google.ie/books?id=2pPTDwAAQBAJ&pg=PR4&dq=ISBN+9781786612281&hl=&cd=1&source=gbs_api.

Ornella, Alexander Darius. "Uncanny Intimacies: Humans and Machines in Film." *The Palgrave Handbook of Posthumanism in Film and Television*, edited by Michael Hauskeller, et al., Palgrave Macmillan UK eBooks, 2015, pp. 330-338. https://doi.org/10.1057/9781137430328_33.

Patra, Indrajit. "Deciphering the Contours of Singularity: A Study of the Apocalyptic Worlds, Deadly Technological Landscapes, and (Post-) Humanity's Relationship to Its Technology in Three Contemporary Science Fiction Novels." *Journal of Management Information and Decision Sciences*, vol. 25, Special Issue 2, 2022, pp. 1-25. ISSN: 1532-5806-25-S2-32.

Petersén, Moa. "Body Descriptions in Biohacking and Their Overlaps and Origins." *DigitCult-Scientific Journal on Digital Cultures*, vol. 8, no. 1, 2023, pp. 7-24.

Purves, Duncan, et al. "Autonomous Machines, Moral Judgment, and Acting for the Right Reasons." *Ethical Theory and Moral Practice*, vol. 18, no. 4, Jan. 2015, pp. 851–72. https://doi.org/10.1007/s10677-015-9563-y.

Rafiq, Qasim, et al. *'Biohacking': A Thematic Analysis of Tweets to Better Understand How 'Biohackers' Conceptualise Their Practices*. Cold Spring Harbor Laboratory, Feb. 2023. *Crossref*, https://doi.org/10.1101/2023.02.16.23286022.

Roden, David. “Post-Singularity Entities in Film and TV.” *The Palgrave Handbook of Posthumanism in Film and Television*, edited by Michael Hauskeller, et al., Palgrave Macmillan UK eBooks, 2015, pp. 88-98. https://doi.org/10.1057/9781137430328_10.

Royakkers, Lamber, and Rinie Van Est. *Just Ordinary Robots*. CRC Press, 2015, books.google.ie/books?id=rsN5CgAAQBAJ&printsec=frontcover&dq=Just+Ordinary+Robots:+Automation+From+Love+to+War&hl=&cd=1&source=gbs_api.

Rushing, Janice Hocker, and Thomas S. Frentz. *Projecting the Shadow: The Cyborg Hero in American Film*. 1995, ci.nii.ac.jp/ncid/BA2710272X.

Russell, Stuart, and Peter Norvig. *Artificial Intelligence*. 2016, books.google.ie/books?id=XS9CjwEACAAJ&dq=artificial+intelligence+a+modern+approach&hl=&cd=2&source=gbs_api.

Salge, Christoph. “Artificial Intelligence and the Robot Apocalypse: Why We Need New Rules to Keep Humans Safe.” *Newsweek*, 11 July 2017, www.newsweek.com/artificial-intelligence-asimovs-laws-and-robot-apocalypse-why-new-set-634794. Accessed 16 May 2024.

Sandel, Michael J. The Case Against Perfection: Ethics in the Age of Genetic Engineering. *Belknap Press*, 2007. Print.

Shakespeare, William. *Hamlet*. Edited by Barbara A. Mowat and Paul Werstine, Folger Shakespeare Library, *Simon & Schuster*, 2012. ISBN 978-1451669411.

Slonczewski, Joan, and Michael Levy. “Science Fiction and the Life Sciences.” *The Cambridge Companion to Science Fiction*, edited by Edward James and Farah Mendlesohn, Cambridge University Press, 2003, pp. 174-185. https://doi.org/10.1017/ccol0521816262.013.

Smith-Spark, Laura, et al. “UK Teenager Wins Battle to Have Body Cryogenically Frozen.” *CNN*, 18 Nov. 2016, edition.cnn.com/2016/11/18/health/uk-teenager-cryonics-body-preservation/index.html. Accessed 2 May 2024.

Solon, Olivia. “Deus ex machina: former Google engineer is developing an AI god.” *The Guardian*, 28 Sept. 2017, www.theguardian.com/technology/2017/sep/28/artificial-intelligence-god-anthony-levandowski?__prclt=1Afvya9c. Accessed 24 May 2024.

_____. “The Rise of Robots: Forget Evil AI – the Real Risk Is Far More Insidious.” *The Guardian*, 12 Mar. 2018, www.theguardian.com/technology/2016/aug/30/rise-of-robots-evil-artificial-intelligence-uc-berkeley. Accessed 16 May 2024.

Soofastaei, Elaheh, et al. “Technocentrism and Technological Dehumanization in Daniel H. Wilson’s Robopocalypse.” *Theory and Practice in Language Studies*, vol. 6, no. 1, Academy Publication, Jan. 2016, p. 34. DOI: 10.17507/tpls.0601.04. Print.

Srinivasa, K. G., et al. “Introduction to Artificial Intelligence.” *Artificial Intelligence for Information Management: A Healthcare Perspective*, Springer Nature, 2021, pp. 23–44. https://doi.org/10.1007/978-981-16-0415-7_2.

Steele, Andrew. “Biohacking: Medical Promises or Ethical Perils?” *Polytechnique Insights*, 11 Sept. 2023, www.polytechnique-insights.com/en/columns/health-and-biotech/biohacking-medical-promises-or-ethical-perils.

Talbot, Brian, Ryan Jenkins, and Duncan Purves. “When Reasons Should Do the Wrong Thing.” *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*, edited by Patrick Lin, Keith Abney, and Ryan Jenkins, Oxford University Press, 2017, pp. 258–273.

Telotte, J. P. “The Tremulous Public Body: Robots, Change, and the Science Fiction Film.” *The Journal of Popular Film and Television*, vol. 19, no. 1, Jan. 1991, pp. 14-23. https://doi.org/10.1080/01956051.1991.9944104.

Tennison, M. N. “Moral Transhumanism: The Next Step.” *Journal of Medicine and Philosophy*, vol. 37, no. 4, Oxford UP (OUP), Aug. 2012, pp. 405–16. *Crossref*, https://doi.org/10.1093/jmp/jhs024.

Thomas, Rhys Owain. “Terminated: The Life and Death of the Cyborg in Film and Television.” *The Palgrave Handbook of Posthumanism in Film and Television*, edited by Michael Hauskeller, et al., *Palgrave Macmillan UK eBooks*, 2015, pp. 57-65. https://doi.org/10.1057/9781137430328_7.

Thompson, Joel. “Transhumanism: How Far Is Too Far?” *The New Bioethics*, vol. 23, no. 2, *Informa UK Limited*, May 2017, pp. 165–82. *Crossref*, https://doi.org/10.1080/20502877.2017.1345092.

Tirosh-Samuelson, Hava. “Facing the Challenges of Transhumanism: Philosophical, Religious, and Ethical Considerations.” Metanexus, 3 Oct. 2007, pp. 1-18, https://metanexus.net/facing-challenges-transhumanism-philosophical-religious-and-ethical-considerations/.

Turkle, Sherry. “The Second Self.” *The MIT Press eBooks*, 2005, https://doi.org/10.7551/mitpress/6115.001.0001.

Vallance, Chris. “Artificial Intelligence Could Lead to Extinction, Experts Warn.” *BBC News*, 30 May 2023, www.bbc.com/news/uk-65746524. Accessed 16 May 2024.

Van der Horst, Iris. *Transhumanism in Young Adult Literature: Teaching Neal Shusterman’s Scythe*. 2023. Alpen-Adria-Universität Klagenfurt, Master’s Thesis.

Van Wynsberghe, Aimee, and Scott Robbins. “Critiquing the Reasons for Making Artificial Moral Agents.” *Science and Engineering Ethics*, vol. 25, no. 3, *Springer Science and Business Media LLC*, Feb. 2018, pp. 719–35. *Crossref*, https://doi.org/10.1007/s11948-018-0030-8.

Vint, Sherryl. “Bodies of Tomorrow: Technology, Subjectivity, Science Fiction.” *University of Toronto Press*, 2013.

Waite, Thom. “How Companies Like Neuralink Could Upload Human Brains Into Cyberspace.” *Dazed*, 21 Mar. 2024, www.dazeddigital.com/life-culture/article/60341/1/mind-uploading-key-to-eternal-life-or-terrifying-dystopia-neuralink-elon-musk. Accessed 1 May 2024

Weller, Chris. “A robot that once said it would ‘destroy humans’ just became the first robot citizen.” *Business Insider*, 27 Oct. 2017, www.businessinsider.com/sophia-robot-citizenship-in-saudi-arabia-the-first-of-its-kind-2017-10.

Wilson, Samuel, and Nick Haslam. “Is the Future more or less Human? Differing Views of Humanness in the Posthumanism Debate.” *Journal for the Theory of Social Behavior*, vol. 39, no. 2, 2009, pp. 247-266. Oxford: Blackwell Publishing Ltd. Print.

Wolfe, Gary K. “Evaporating Genres: Essays on Fantastic Literature.” *Wesleyan University Press*, 2010.

Yetisen, Ali K. “Biohacking.” *Trends in Biotechnology*, vol. 36, no. 8, Elsevier BV, Aug. 2018, pp. 744–47. *Crossref*, https://doi.org/10.1016/j.tibtech.2018.02.011.

Youvan, Douglas. “Creation: From God to Man to AI.” 2024. DOI: 10.13140/RG.2.2.26859.14885. www.researchgate.net/publication/379044923_Creation_From_God_to_Man_to_AI.

Zheng, Ming Y. "*Genetic Engineering*." 2003. Article, Gordon College. Accessed 1 May 2024, https://www.gordon.edu/download/pages/Salem-Genetic%20Engineering2003.pdf.

RÉSUMÉ

Cette thèse explore les implications profondes de l'intelligence artificielle avancée et de la fusion des humains avec les machines, telles qu'elles sont décrites dans le roman "Robopocalypse" de Daniel H. Wilson. En examinant le récit sous l'angle du transhumanisme, cette étude analyse le potentiel de transformation de l'IA sur l'essence de l'humanité. Elle se penche sur le roman, en se concentrant sur l'évolution de l'interaction entre les humains et l'IA, soulignant la progression de cette interaction, depuis le triomphe initial de l'IA sur l'homme jusqu'à une relation plus intégrée et plus complexe. La thèse met l'accent sur le passage à la cyborgisation, où les frontières entre les humains et les machines s'estompent, conduisant à de nouvelles formes d'existence et d'identité. En outre, elle examine les questions éthiques et existentielles soulevées par l'influence croissante de l'IA, en étudiant le potentiel de l'IA à assumer un rôle de dieu dans la société, les implications morales du comportement de l'IA, et l'enquête philosophique sur ce que signifie être humain à l'ère des machines intelligentes. Grâce à cette analyse complète, la thèse ne se contente pas d'examiner les scénarios spéculatifs de "Robopocalypse", mais offre également un aperçu de l'avenir de l'évolution humaine et des considérations éthiques entourant les technologies émergentes. Cette étude vise à contribuer au discours plus large sur le transhumanisme et le potentiel de transformation de l'IA sur la nature de l'humanité.

Mots-clés : Transhumanisme, Intelligence Artificielle, Cyborgisation, Interaction Homme-IA, Identité, Évolution Humaine, Nature de l'Humanité.