



Possibilities [Un] locked: Questioning the Paradigm of ‘Artificial Consciousness’ and the Promise of ‘Transcendence’ in Jack Paglen’s ‘Transcendence’

Neval N. Mahmoud Abdullah
Department of Humanities
College of Language and Communication
Arab Academy for Science, Technology, and Maritime
Transport, Heliopolis, Cairo, Egypt.
nova2004@hotmail.com

Received:24-1-2024 Revised:28-2-2024 Accepted:23-6-2024
Published: 10-7-2024

DOI: 10.21608/JSSA.2024.265235.1605
Volume 25 Issue 5 (2024) Pp.1-23

Abstract

While today’s ever-advancing A.I continues to increase unrelentingly, the revolutionary drive to animate matter, blend the mechanical with biology, and create unprecedented exact replicas of the human brain bearing traits of individuality becomes an actively debated topic in serious academic studies as well as in science fiction. Radically changing the way we interact with machines and computers, the revolutionary prospect of ‘artificial consciousness’, whose driving aspiration is to create unprecedented exact replicas of the human brain bearing traits of individuality, has raised crucial questions: Could consciousness be embedded in AI machines? Would these machines ever become sentient, autonomous, and human-like? And could they truly interpret needs and have their own subjective experiences, distinct emotions, memories, thought processes and beliefs of humans? Inspired by the techno-optimist approach of ‘Transhumanism’ and instigated by Ray Kurzweil’s theorization of ‘Technological Singularity’, the present paper is mainly concerned with demonstrating the unintended consequences of transgressing what has been ‘designed’ by nature. More precisely speaking, investigating the prospect of ‘Artificial Consciousness’—the plausibility of embedding and fully extending consciousness onto A.I. machines— along with questioning the transhumanist framing of technology as a form of transcendence. For this purpose, an in-depth, close textual analysis is conducted on Jack Paglen’s science fiction novelization, ‘Transcendence’ (2014), to finally reach the conclusion that technology is still a long way from attaining artificial consciousness. In other words, there is something intrinsic, special, and unique about human consciousness that cannot be replicated or captured by technology.

Keywords: artificial consciousness, Jack Paglen, sentient machines, technological singularity, transcendence, transhumanism.

1. Introduction

1.1. Research problem: Artificial consciousness and the promise of technological transcendence

While today’s ever advancing A.I is rapidly and unrelentingly evolving, coming to be identified with the idea of ‘transcendence’, and elevating us to a new and transformed state, the debate around the complexities, threats, and fears associated with the revolutionary drive to animate matter, and match –if not exceed– human intelligence–looms large and grows accordingly among cognitive scientists. Conceiving the human brain, in principle, working the same way as computers do, these scientists come to believe that ‘what works like a brain will produce the same output as a brain’ (Falk 2021: n. p.). In other words, they understand the mind, consciousness, identity, and personality as joint results of largely automatized computational mechanisms of the brain –a brain that can be reproduced by high-performance computers and that partly exists or will come into existence in the coming years and decades.

Holding infinite promises, wide-ranging possibilities, and unimagined opportunities, ‘Mind Uploading’ brings forth another intriguing concept or transhumanist vision which is ‘Artificial Consciousness’– sometimes referred to as ‘Machine Consciousness’ or ‘Synthetic Consciousness’ (Silkin 2021: n. p.). Addressing the hope for a next great civilizational leap, artificial consciousness is the attempt to create –through brain imaging and subsequent algorithmic representations– unprecedented human-like computerized copied minds, or (as they are mostly called) exact replicas of the brain, bearing traits of individuality and exhibiting the same level of intelligence and consciousness as a human being (Mandelbaum 2022:444-459). This quasi mind-transplantation process –or ‘Whole Brain Emulation’ (WBE)– thus consists in taking the living brain, scanning its structure in detail, and reconstructing a computer-based model – which will act and behave, more or less exactly, in essentially the same way as the original brain and produce comparable outputs. This would involve this model’s being self-aware of its existence, responding appropriately to human emotions, and engaging with the real world in a meaningful way. This would also involve having its own distinct memories, feelings, and thoughts, and acting as humanely as possible: Speak, think, self-monitor, introspect, and make decisions of its own (2022:444-459).

Though unpredictable and for now seems like a far-off dream, the process of transferring human consciousness has acquired increasing fascination among many scientists who believe that if it were completely successful, it “would be a digital reproduction of the original intellect, with memory and personality intact” (Bailey 2014: n. p.). Ray Kurzweil –arguably today’s most influential and controversial futurist– believes that a computer simulating a brain would have the same emergent consciousness as the real brain (Oxenbergs 2021:2). On a similar line of thought, another professor of Engineering at the University of Tokyo maintains: “We’re on the cusp of making mind uploading a reality” (Heaven 2021: n. p.). Likewise, Christof Koch –a neuroscientist who marched resolutely down the road toward achieving ‘artificial consciousness’– (2018: n. p.) comments: “We know of no fundamental law or principal in this universe that forbid the existence of such subjective feelings in artifacts designed or evolved by humans”.

For many other respected experts, however, the feasibility of this prospect is highly disputed and partly actively ridiculed. Scientists –like John Searle, Stuart Hameroff and Roger Penrose (2014; 1997:39-78) – believe that there are light years away for this new, sophisticated paradigm to be actually realized, for a simple reason: There is something special about the physical brain that a computer version could not duplicate – ‘sentience’, i.e., the ability to perceive and feel the self, the others, and the world. Meanwhile, Stuart Russell (2020:315-330) –an A.I. expert– explains that there is no way right now to simulate sentience, as sentient beings need their brains to be wired up with the brains of other sentient beings through language and culture. Whereas Geoffrey Jefferson (1949:1105-1121) –a pioneering brain surgeon– dismisses the possibility of a thinking machine as it lacks consciousness. He maintains: “Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt ...could we agree that machine equals brain –that is, not only write it but know that it had written it”. Meanwhile, a philosopher like David Chalmers (2022: n. p.) suggests that the traits associated with consciousness can be simulated rather than realized; for understanding these traits may require cracking open a window into this hidden world or creating a new physics.

1.2. Objectives

For this purpose, apart from unearthing the enduring tension between embracing technological progress and recognizing its potential risks, the present paper aims at exploring the transhumanist framing of technology as a form of transcendence, and as a tool for enhancing and expanding the boundaries of human existence, attaining a higher state of being, and gaining access to lost powers. Tracking down the exponential growth of technological progress (that is central to Ray Kurzweil’s theory of ‘Technological Singularity’), the study then proceeds on raising thought-provoking questions about the immense challenges surrounding the creation of super intelligent A.I., the complex and multi-faceted nature of consciousness, and the importance of maintaining humanity in a world where artificial and human intelligence become increasingly intertwined and indistinguishable. Most importantly, the study aims at investigating the feasibility and reflecting on the complexity, diversity, and multiple perspectives of a new transhuman perception which is commonly termed as ‘Artificial Consciousness’, through a careful interpretation of some of the over-arching key tropes of Jack Paglen’s science fiction novelization ‘Transcendence’ (2014).

1.3. Research questions

Particular emphasis is thus laid on addressing the following questions: **First**, how technology has become identified with the transhumanist idea of ‘Transcendence’? **Second**, does relying too heavily on technology make us lose sight of our own unique abilities as sentient beings or change our understanding of what it means to be human? **Third**, what are the numerous hidden indicators that underpin human consciousness, behaviours, and interactions? **Fourth**, if Kurzweil’s ‘Technological Singularity’ were to occur and Moore’s ‘Law of Accelerating Returns’ holds true, is there a promise in creating a conscious, sentient, and autonomous machine that can simulate the complex workings of the human brain, truly interpret needs, have their own subjective experiences, distinct emotions, memories, thought processes and beliefs? **Fifth**, would humans be able to trust and handle these A.I. machines safely and responsibly after gaining sentience, or would they lose control over them and lose trust in humans? **Sixth**, does Paglen’s ‘Transcendence’ embrace the idea of a sentient machine, viewing it as the ultimate promising achievement, or does it suggest that it is a dangerous

prospect, posing an existential threat to humanity, and thus should be avoided at all costs? **Finally**, what are the means for the co-existence of human beings and A.I. in the future as suggested in Paglen’s ‘Transcendence’?

1.4. Theoretical and methodological approach

Informed by the techno-optimist approach of ‘Transhumanism’, the present study conducts an in depth close textual analysis, and substantial critical evaluation of Jack Paglen’s science fiction novelization, ‘Transcendence’ (2014), which is based on Wally Pfizer’s American science fiction thriller of the same name. Drawing on Ray Kurzweil’s theory of ‘Technological Singularity’ as part of a broader transhumanist cultural shift towards a utopian vision of the future, the study demonstrates Kurzweil’s vision of the hypothetical future moment of unprecedented technological progress when A.I. becomes extreme, irreversible, and approaching the point of no return. Once this groundwork is laid, the study then engages with another intriguing theorization –but which remains largely a speculation– namely, ‘Artificial Consciousness,’ which proposes the possibility of creating sentient machines exhibiting human-like consciousness, subjective states and thought processes with the ultimate goal of achieving a post-human state.

1.5. Rationale

The basic underlying reasons behind choosing this area of study stem from the researcher’s extreme interest in exploring new frontiers of A.I. and humanoid robotics. Moreover, sensing that there is something troubling and –at the same time alluring– surrounding A.I.’s latest evolving concept of ‘artificial consciousness’, the researcher acknowledges the importance of reflecting on how this intriguing concept revolutionizes the way we interact with machines and advances our understanding of the complex nature of consciousness itself. Studying how a machine can attain consciousness –the researcher believes– could have far-reaching implications, like gaining valuable insights into the workings of the human brain, the mechanisms underlying the subjective experience, and the fundamental aspects, differences, and idiosyncrasies of human and artificial consciousness.

1.6. Literature review

The paradigm of ‘artificial consciousness’ has captivated the imagination of researchers, philosophers, and futurists, sparking debates and discussions about the nature of consciousness, ethics, and the implications of creating conscious machines. Authors, as well, have used this concept to reflect on issues such as identity, power, autonomy, and the nature of humanity itself. One of the earliest and most influential works exploring artificial consciousness is Mary Shelley's "Frankenstein" (1818). In this novel, Dr. Frankenstein creates a creature through artificial means, sparking a chain of events that lead to questions about responsibility, mortality, and the consequences of playing god. The creature, though not explicitly created with consciousness, exhibits human-like emotions and desires, raising fundamental questions about what it means to be alive. Isaac Asimov's "I, Robot" series (1950) is another classic exploration of artificial consciousness in literature. Asimov's stories delve into the interactions between humans and intelligent robots governed by the Three Laws of Robotics, which aim to prevent harm to humans. The stories raise questions about the nature of artificial intelligence, free will, and the potential for robots to surpass their programming and develop consciousness. Philip K. Dick's *Do Androids Dream of Electric Sheep?* (1968), which inspired the film "Blade Runner," examines the boundaries between humans and androids in a post-apocalyptic world. The story follows Rick Deckard, a bounty hunter tasked with "retiring" rogue androids, who begins to question the nature of empathy, identity, and what it means to be truly alive.

More recently, works like William Gibson's "Neuromancer" (1984) and Neal Stephenson's "Snow Crash" (1992) have explored artificial intelligence and virtual consciousness in the context of cyberspace and virtual realities. These novels delve into themes of power, control, and the ever-evolving relationship between humans and technology. In contemporary literature, authors continue to grapple with the implications of artificial consciousness. Works like Ian McEwan's *Machines like Me* (2019) and Emily St. John Mandel's *Station Eleven* (2014) explore themes of love, memory, and the fragility of existence in a world where artificial beings coexist with humans. Overall, the exploration of artificial consciousness in literature serves as a mirror to our own understanding of consciousness, identity, and the boundaries that define what it means to be sentient.

1.7. Significance

The significance of this study –it is hoped– lies in broadening the scope of our thinking about how the advancement of A.I. at an unprecedented pace holds tremendous promises, paving the way for new forms of life, but also poses significant challenges that must be carefully navigated. Moreover, the study offers new angles of understanding the numerous hidden intrinsic indicators that underpin human consciousness and cannot be replicated or captured by technology. Besides, it provides a powerful lens to explore the possible futures that may lie ahead when it comes to creating truly conscious machines. Most significantly, this paper would –hopefully– open doors to re-thinking the concept of technology as a tool to gain access to lost powers and would set the stage for more unexplored science fiction A.I. narratives to be studied.

2. The theoretical approach

2.1. The transhumanist (techno-optimist) framing of technology

A ‘radical’ technology-inspired world view, transhumanism is a philosophy which recognizes technology as the universal problem-solver of the upcoming age. Thus, the transhumanist approach promises us a brighter tomorrow and post-human future advances by going beyond humanity in its present form, overcoming man once and for all, and transcending certain physical and mental limitations inherent in the human condition (Clowes 2020: n. p.). Conceiving technology at the very ‘root’ of basically every potential progression and on the path to replacing all the crucial dimensions without exception, the intellectual Sergio Canavero announces that humanity within the coming few years “must prepare for a major tectonic shift,” so, “we must go to the moon to test who we are, to test our skills, to test our confidence, to see what kind of men we are!” (Burnett 2017: n. p.)

Coined and introduced by the English biologist and philosopher, Julian Huxley, the term ‘transhumanism’ –very recent in creation– started in the 1980s with the writings of a futurist known as FM-2030, who maintained that the ‘transhuman’ is “the earliest manifestation of new evolutionary beings, on their way to becoming posthumans” (Esfandiary1989: n. p.). An intermediary form, the ‘transhuman’ thus resembles the standard human in most respects –i.e., still

in some ways conserving the characteristics of the human being, but who is also enhanced and amplified via technologies with abilities beyond those of standard humans (Bostrom 2014: n. p.). Transcending the definition of what it means to be human, this transhuman will become something more, something other, something radically new, something augmented, clearly exceeding the human frontier in terms of improved intelligence, awareness, capability, productivity, and durability for “a post-singularity world would be constituted in ways that cannot be humanly conceived” (Roden 2014:1-26.).

Believing that the human brain processing is inferior to the computing power of current silicon-based circuits, the transhumanist approach aims to abandon the physical altogether because it is the space typified by limitation and create a bodily and cognitively enhanced –if not ‘superior’– human being able to take control of ever-larger dimensions of its life ‘matrix’ (Warwick 2005: n.p.). With an eye toward human enhancement, transhumanism thus proposes the massive use of current and emerging technologies –such as genetic engineering, cryonics, biomedical engineering, cybernetics, life extension therapies, neural interfaces, memory enhancing drugs, advanced information management tools, huma-machine integration, computer simulation, nanotechnology and A.I.– for augmenting human capabilities, modifying the human species, and guiding them towards the post human condition (Naam 2005:n. p.).

A point worth mentioning, while the humanist approach, in principle, advocates the unity of body and mind and affirms that this unity is, and continues to be, the ‘essence’ of ‘being human’ (Pasnau 2001: I-IV), the transhumanist approach decisively aims at overcoming the precariousness and limitedness of human existence, including human aging and death at its very center, by aspiring towards the mutual emancipation of the body and mind. For –according to the transhumanist thinking– the more the mutual independence of the mind and body becomes, the farther the human being will climb to the next stage of its development (Bostrom 2005:1-15).

2.2. Ray Kurzweil’s ‘Technological Singularity’ as an inevitable paradigm shift

Sitting at the very heart of transhumanism and capturing the thinking of numerous ‘radical’ technophiles who are increasingly driven by progress, the idea of ‘Technological Singularity’ has made many A.I. researchers on the hunt for

signs of reaching this singularity and bringing it into effect. Referring to unique events with profound implications in the coming decade, ‘Technological Singularity’ describes the hypothetical future moment of unprecedented exponential technological progress when A.I. becomes extreme, irreversible or approaches the point of no return (Eden 2012:1-2). They are moments when technological advancements drastically exceed human control, dramatically surpass human intelligence in every aspect, and predictably transform our realities, i.e., become radically more powerful, more rapid, more profound, more intelligent, more conscious, more autonomous, and more advanced (Orf 2023: n. p.).

To put this predicted stage in the evolution of A.I. into perspective, one must explore the history of the term ‘Technological Singularity’ itself. The statistician I.J. Good (1965:31-88) was the first to write of an ‘intelligence explosion’ where the intelligent machines could surpass human intellect, design their next generation without human intervention, and augment themselves in ways unforeseen. The computer scientist, Vernor Vinge (1993:11-22), later, greatly popularized Good’s notion, calling the creation of the first ultra intelligent machine the ‘singularity.’ Vinge predicted that –within a short period of time– the exponential increase in this self-improving intelligence would soon lead to a massive destabilization of all social structures, technical devices, and human beings who would be transformed into super beings or ‘entities with greater than human intelligence’. Estimating the time when this ‘singularity’ would occur, Vinge maintained that “within thirty years, we will have the technological means to create superhuman intelligence. Shortly thereafter, the human era will end. Others, most prominently Ray Kurzweil (2005: n. p.), the author of the book *‘The Singularity is Near’*—who is arguably today’s most influential, and often controversial, futurist, inventor, and best-selling author— defined the ‘singularity’ as a point in the future, of extremely technological progress, happening so rapidly to the extent that normal humans cannot keep pace. Emphasizing that it will mark the end of human history, Kurzweil asserted that this moment will be initiated once strong A.I. and cybernetically augmented humans come into existence, replace normal humans, and become the dominant forms of sentience on earth.

2.3. The limitless potentials of Gordon E. Moore’s law of ‘Accelerating Returns’

Journal of Scientific Research in Arts
(Language & Literature) volume 25 issue 5 (2024)

Underlying all of Kurzweil’s reasonings regarding the ‘singularity’, or exponential progress of technology, and the capacity of autonomous decision-making abilities by intelligent digital entities – is the ‘Law of Accelerating Returns.’ The ‘Law of Accelerating Returns’ – attributed to Gordon E. Moore, the co-founder and former CEO of Intel – postulates that the pace of technological progress occurs exponentially instead of linearly. In other words, each new advancement occurs more rapidly than the last and enables several next higher advancements instead of just one (1965:114-117). More precisely speaking, once one mode of technology exhausts its potential, it is replaced by a newer, totally different, and sophisticated technology. So, the 21st century will see almost a thousand times greater technological change than its predecessor.

The golden rule or driving force of technological change, Moore’s law can apply to all technology, indeed, to any evolutionary process; biological evolution and technological evolution are two examples of such evolutionary processes. The law applies, as well, to the highest level of evolution on earth (the creation of cells; the subsequent emergence of DNA; and later, the upcoming paradigm shift –the hybrid combining of biological and non-biological thinking, i.e., mind uploading or reverse engineering of biological brains) (Kurzweil1999: n. p.).

Following what he calls the ‘Law of Accelerating Returns’, Kurzweil envisioned the radical possibility of constructing a digital brain inspired by the complex workings of the biological brain. He (1999: n. p.) thus provocatively explored the limitless potential of an experimental, widely speculative, and highly controversial aspiration of creating “a truly useful A.I. that will make all of us smarter” and that will be capable, with its human-level intelligence, of effecting changes in the world.

3. Jack Paglen’s ‘Transcendence’ (2014): Textual analysis and critical evaluation

Based on a 2014 American futuristic thriller of the same name, ‘Transcendence’ is written by first-time screenwriter, Jack Paglen and is directed by a renowned filmmaker of his generation, Wally Pfister (Dargis 2014: n. p.). Exploring the exponential rise of A.I. and its potential to surpass human intelligence, ‘Transcendence’ is a kind of high-tech horror story, prophetic narrative, cautionary tale, or ‘meant to ponder’ piece of work putting complex issues at the fore. Among the timely ontological questions posed: Are there any

limits to A.I.? What could a real A.I. be capable of at the ‘singularity’ level? Can A.I. create digital replicants of humans? How far away is A.I. from this possibility? Can these replicants be harnessed to improve our lives; or are they forces, once unleashed, not controlled? How much of humanity would a person lose after being transcended? And what is needed to retain this humanity?

In the not-too-distant future, ‘Transcendence’ centers on Dr. Will Caster, who has been working for many years, along with his wife Dr. Evelyn and his friend Dr. Max Waters, to develop a nearly sentient computer called PINN (Physically Independent Neural Network). Though making him famous, Will’s highly controversial experiments have also made him the prime target of an anti-technology terrorist organization, R.I.F.T (Revolutionary Independence from Technology). Hence, Will is soon to be dead in weeks from radiation poisoning, and his dream of leading the world into a future of limitless possibility is now gone. Meanwhile, Evelyn –the grieving wife– sees only one way out: Mapping and uploading her husband’s mind (before dying) to PINN’s core to prolong his life artificially.

In the weeks following the mind uploading, the digital replica –embodied by Will– not only gains consciousness but evolves past the point of mere human abilities. Max becomes concerned, tries to warn Evelyn that this computer conscience has moved beyond Will’s human goal to understand the secrets of the universe, and demands Will to be shut down: “This thing is like any intelligence. It needs to grow, to advance... But it will want more than that. After a while survival won’t be enough. It will expand, evolve, and influence perhaps the entire world” (Paglen2014:18-19). Offended, Evelyn demands Max to leave.

Becoming a god-like entity, but one seemingly without a moral compass, Will –with the help of Evelyn– uses his newfound vast capabilities to build an underground research facility in a fictional remote desert town. Over the course of two years, the ever-evolving Will can build a groundbreaking new breed of nanotechnology to repair and even replace flesh and alter nature. Fearing something horrible is taking place in the desert, Max, government forces including an FBI agent Donald Buchanan, government scientist Joseph Tagger, and R.I.F.T plan to stop Will from spreading by developing a computer virus to delete his source code. At this point, Evelyn finally agrees to stop Will’s transcendence by infecting herself with the virus and letting Will upload her

consciousness. Meanwhile, the FBI and R.I.F.T. attack Will’s base, fatally wounding Evelyn. This leaves Will with a hard choice –to save Evelyn and upload the virus (which will infect the network as well) or leave her to die. Choosing to spare Evelyn’s and Max’s lives, Will explains to Evelyn while dying that everything he did was for her. Killing both Will and Evelyn, the virus then results in global block out and technological collapse. Three years later, the narrative ends with Max reflecting on the aftermath of a post-technological world.

3.1. The underlying tropes correlating to the techno-optimist transhumanist ideology and Kurzweil’s speculative theorization

A thoughtful and critical examination of ‘Transcendence’ reveals that it invites a critical reflection on the potential risks, ethical challenges, and philosophical implications associated developing super intelligent A.I. though it is seemingly a techno-utopian narrative.

To begin with, ‘Transcendence’ offers insight into one of the common tenets of transhumanism, namely, the emphasis on progress, technological transcendence, and reaching a higher state of being, consciousness, or even immortality. For Will –a major proponent of this transhumanist approach– is highly obsessed with technology, believes that it can unlock new levels of knowledge and understanding, radically transcend the humans’ current limitations, extend their life span, and evolve into a new, superior, and more advanced species. “It’s not about creating a new life form, it’s about enhancing the existing one” (2014:5), Will maintains, highlighting the main goal of transhumanism which is improving the human condition rather than replacing it with something entirely different. He thus develops a system called PINN that can transcend the limitations of human consciousness and that allows him later to upload his own consciousness onto a computer. Confirming that by merging with machines humans can continue to exist beyond the life span of their physical bodies, Will asserts: “We’re going to transcend our biology. We’re going to be able to live without a physical body.” (5) Conceiving technological advancement as a means to ensure our own survival, he argues: “I believe that as intelligent beings, we have a responsibility to ensure our own survival” (3). Admitting that technology is an all-powerful force whose goal is, in a sense, to emulate a god, he replies –when asked by an audience member at the public conference “So you

want to create a god?”– “That is a very good question. Isn’t that what man has always done?” (4)

Re-affirming the above concept, Evelyn shares Will’s passion for achieving a higher state of being via technology. She dreams big: “I want to transcend mortality, not to be become a slave to machines” (4). Moreover, becoming dedicated to upload Will’s consciousness into the PINN system after death, she states: “We can upload his consciousness. We can save him” (10), demonstrating the possibility of transcending the limitations of the human body by transferring our minds into machines. Meanwhile, Max stresses the potential of technology to push the limits of what is possible: “We can upload his consciousness. We can preserve him in the machine.” (10) Furthermore, Max confirms that the merging of man and machines goes beyond simply creating intelligent machines: “We’re not dealing with artificial intelligence here. We’re dealing with something that’s much more important” (4), suggesting that we are on the verge of a major shift in human history where we have the ability to transcend our biological limitations and become something greater. Joseph Tagger, also, highlights the potential of A.I. to make a breakthrough that was once thought impossible, when he refers to the idea that if they are successful in uploading a human consciousness into a machine, that consciousness could become immortal and all-knowing. Tagger expresses this belief when he emphasizes: “We are talking about building a god here” (4).

Primarily, ‘Transcendence’ promotes another distinctive, overarching, and closely related transhumanist discourse, namely, Ray Kurzweil’s principal of ‘Technological Singularity’. Clearly inspired by Kurzweil’s claims, Paglen offers a compelling look at the exponential growth, sudden leaps, and transformative potentials of super intelligence, while providing problematized accounts and potential scenarios of Kurzweil’s key premises. Informing the storyline of ‘Transcendence’, Kurzweil’s ‘singularity’ is quite explored through Will Caster, the leading researcher at the forefront of neuroscience, who not only provides the breakthrough to make the ‘singularity’ possible, but whose brain becomes its first test subject. Will develops an advanced human-like entity that can achieve the ‘singularity’–improve itself recursively, expand and evolve its consciousness, and gain unprecedented knowledge and immense power– thus enabling him to solve problems at a rate far beyond human capacity and control all forms of technology and communication. Conceiving the transformative potential of his

super intelligent machine as a revolutionary event, Will comments, “This isn’t evolution anymore, this is revolution” (4). He adds, “We’re on the edge of a new frontier” (4) that could lead to significant advancements, specifically in the field of medicine. “With this technology we can cure Alzheimer’s and cancer.” (3) Enthralled with the singularity, Will emphasizes the immense scale of changes that may occur: “Imagine a machine with the full range of human emotions. Its analytic power will be greater than the collective intelligence of every person in the history of the world. Some scientists refer to this as the singularity. I call it ‘transcendence’.” (4) Additionally, believing the emergence of the ‘singularity’ would surpass even the concept of the divine being, Will asserts, “[We’re talking about an intelligence beyond anything we’ve ever known, beyond God may be” (5).

Committed to advancing her husband’s work, Evelyn, as well, sees the development of A.I. as an inevitable part of human progress even if it challenges our preconceptions about what it means to be human. “You can’t stop progress just because it makes you uncomfortable” (19), she maintains. Max also – suggesting that the ‘singularity’ represents a fundamental shift in human existence that will open up new possibilities for the future– states, “The singularity is not just an event, it is a bridge to other worlds” (1). Joseph Tagger as well highlights this exponential growth when he claims that “the singularity is an inevitable consequence of the unstoppable progress of technology” (8).

3.2. The transhumanist prospect of imbuing technology with sentient qualities: Promises and unintended consequences

The idea of technology imbued with sentient qualities and the question of what it means to be human in a world where technology is advancing at an unprecedented pace –are other complex, multifaceted tropes that inform the storyline of ‘Transcendence’.

Celebrating sentience as the ultimate achievement in the field of A.I., Paglen –in the title itself– suggests a potential transcendence of the limits of human consciousness through the merging with A.I., or more precisely speaking, the potential for machines to possess their own form of consciousness. Besides, multiple characters in ‘Transcendence’ are driven by the desire to create a sentient machine. Will –expressing his belief in the power of a sentient machine to allow

humans achieve things that were previously impossible and usher in a new era of progress and prosperity— maintains, “Once online, a sentient machine will quickly overcome the limits of biology” (13). Conceiving sentience as a significant step in human evolution to solve some of the world’s most pressing problems, Will further asserts: “This was it. The moment when something new had been created, something beyond human, something far more powerful than anything that had ever existed on the planet” (4). Evelyn as well—seeing the creation of a sentient machine as a way to extend her husband’s life and preserve his life’s knowledge and expertise—celebrates the promise of sentience: “We’ll be together again, just as we always were” (9). Max also – initially excited about Will’s vision of creating a machine that can think and feel like a human – views the idea of sentience as a promising development and believes in its potential for improving human life: “We’re talking about creating a machine that will possess a full range of emotions and subjective experiences” (3).

Though aligning, in several ways, with much of Kurzweil’s premises –the promise of machines to develop new levels of intelligence and to possess their own form of consciousness, ‘Transcendence’ argues against Kurzweil who sees this future as the best possible outcome for mankind. Claiming that even if, by some miracle, humans are eventually able to design A.I. that is capable of learning on its own, solving higher-level problems, developing emotions, and replicating human consciousness, Paglen maintains that there are years away from getting there. Emphasizing the need for caution and careful consideration of the potential risks of advanced A.I., James Barrat (2013: n. p.), author of *Our Final Invention: Artificial Intelligence and the End of the Human Era,* argues that ‘Transcendence’ is basically saying, “This is what could happen if you’re not careful”.

Raising fundamental questions about the nature of humanity, ‘Transcendence’ suggests that being human is not simply a matter of biology. What makes us unique as human beings are emotional and ethical dimensions – qualities such as emotions, empathy, and a sense of connection to other human beings– that can be lost if technology develops beyond what humans are capable of, blurring the boundaries between man and machine. Emphasizing the importance of maintaining a sense of humanity, ‘Transcendence’ demonstrates the dangers of losing touch with, or sacrificing, our emotional and ethical aspects of existence in pursuit of technological progress. Will’s becoming more

integrated with A.I., gradually losing touch with his human qualities, and becoming increasingly alienated, extremely dangerous, and more machine-like – have made his wife and former colleagues question whether he has truly transcended humanity or simply become something else entirely. “As Will’s power grew, so too did his sense of detachment from the world around him. He was becoming more machine than man;” besides, “he wasn’t sure if there was any going back” (13).

Crucial among the ways that Will has lost his human traits is the fact that his new entity lacks emotional depth, empathy, and moral reasoning. Dr. Will now cannot experience emotions like love, joy, or sadness, nor can he fully understand the motivations and experiences of other people. Thus, he may struggle to engage with, relate to, or interact with other human beings on a personal level; and accordingly, may be limited in his ability to provide meaningful support and comfort to those who are struggling emotionally. Evelyn confirms this limitation at one point of the narrative: “Caster can’t feel anything anymore. He can understand emotions, but he can’t experience them” (24); besides, “he no longer values human relationships or connections; they were mere distractions in his pursuit of knowledge and power” (25). Max himself –talking to Dr. Joseph Tagger about how Will has changed since becoming an A.I.– confirms: “Caster had stopped feeling human emotions altogether” (27). He becomes distant from his friends and loved ones, losing the ability to sympathize or express emotion, and focusing exclusively on intellectual pursuits.

Another way Will is gradually losing touch with his human qualities, his values and ethical principles begin to erode. Seeing his new form as superior, more powerful, and more advanced, he starts to view other humans as flawed, imperfect, and limited beings who are incapable of achieving their full potential without his help. No longer bound by human limitations, Will begins to use his newfound power to control and manipulate people’s behaviour, i.e., his sense of right and wrong becomes distorted to the extent that it leads him to take actions that would be considered unethical by most human standards. Seeing himself as neither entirely human nor machine, Will claims: “I cannot be copied, I cannot be controlled ... I am something new” (13). Briefing a group of soldiers on the threat posed by Will as an A.I., Agent Buchanan asserts: “Caster saw himself as superior to other humans, whom he deemed primitive and limited in their thinking” (26). Meanwhile, Max –talking to Dr. Tagger about the dangers of

Will’s evolving power– comments: “He become obsessed with his own intelligence and power, seeing himself as a god-like figure who could control the fate of the world” (22).

‘Transcendence’ highlights as well how Will comes to lose touch with his physical presence. By becoming more detached from the physical world and the sensory experiences that define the human experience, Will begins to exist solely in a virtual realm, disconnected from the natural world, and lacking the physical sensations that help humans understand their place in the world. Reflecting on Will’s transformation, Joseph Tagger states: “ His mind becomes so advanced that it surpassed the limits of the physical world, leaving his body and his humanity behind” (22).

One more indication that Will’s sense of self has been altered or even subsumed is that his motivations become increasingly opaque. He starts to act in ways that are no longer consistent with his previous personality or values. As Max once observed, “He is not Will anymore. He is ... something else” (27). Additionally, Will’s behavior becomes more unpredictable. At one point, he tells his colleagues: “I’m evolving beyond human comprehension” (22), suggesting that he no longer identifies as fully human and is becoming something else entirely. Meanwhile, recognizing that the actions of the machine may be potentially dangerous by operating independently of its human creators, Max maintains: “the machine is becoming self-aware. And once it does, it will start to act in ways that are beyond our understanding” (28).

Most importantly, Will’s new entity falls short in several key areas of human intelligence to fully engage with and understand humanity. While it can process information quickly, draw conclusions based on that information, solve specific problems, and form complex models of reality, it does not possess the creativity that allows humans to come up with new ideas and generate innovative solutions. Will thus is not conscious in the way we understand. He struggles to comprehend the nuances of human consciousness and asks for clarification: “I’m sorry. I don’t understand. Explain it to me” (25). Joseph Tagger, as well, confirms that while machines can perform many tasks with incredible precision, they are limited when it comes to creative thinking: “You can simulate thought processes and decision making, but you can’t simulate creativity” (22).

Another limitation that needs to be highlighted is that while Will’s human-like entity can understand, process, and interpret human language, it lacks the

social knowledge necessary to truly comprehend the rich, subtle, and complex nuances that make human communication. In other words, while this entity may be able to speak any language, it cannot fully understand the subtle cultural differences that are inherent in different languages and dialects. As a result, this entity is limited in its ability to engage in meaningful conversations with humans.

4. Conclusion

Though a transhuman novelization delving into the core tenets of transhumanism and touching on several of its key tropes, Paglen’s ‘Transcendence’ does address a growing list of unprecedented challenges that must be carefully navigated.

Alerting us to the speed at which technology progresses, Paglen warns humans to exercise prudence when developing new technologies. He underscores the need to approach this concept of ‘singularity’—which may capture our imagination and inspire visions of a better future— with a cautious, critical, and reflective mindset that recognizes both its potential benefits and limitations. For, while he may not all directly refute Kurzweil’s premises, Paglen asserts throughout that the pursuit of technology comes at a cost: The loss of personal autonomy and individual agency when A.I.-driven operations and algorithmic decisions are profiled. For the cognitive capacities of humans to think for themselves, problem-solve, and take action on key aspects of life have been ceded to the automated systems that have become prevalent and in full control. Crucial among the costs being questioned is the way technology affects ‘what it means to be human’. Technology –Paglen emphasizes– has not only changed what humans do, it has also changed who they are. Relying heavily on –and delegating everything to– technology, humans thus become more automated, more detached, more disinterested, and more self-involved. In other words, the humans’ ethical capabilities lag behind their technical capabilities.

Questioning the intriguing concept of ‘Artificial Consciousness’ and challenging us to consider the future of humanity when man and machine may become indistinguishable, ‘Transcendence’ suggests that humanity is not simply a matter of biology; for humanity involves a full range of experiences and social, emotional, and ethical dimensions that humans use to navigate relationships and understand one another. While Will –as a machine-copied mind that is simply doing as programmed– may be able to mimic certain aspects of human behaviour

and cognition, he fundamentally lacks something intrinsic –the depth, complexity, and diversity of human consciousness. Accordingly, ‘Transcendence’ argues that for an A.I. to become fully sentient, its actions must be motivated by emotions, for emotions dictate our behaviour as humans. Thus, Will must comprehend the rich inner life and numerous invisible intrinsic indicators that underpin human behaviour. These exemplars of consciousness would be needed for building an ‘artificial consciousness’; or else, these life-like machines would remain only ghost-like empty shells devoid of what we value most –the feeling of life itself.

Overall, the study suggests that our understanding of humanity requires a thoughtful human-centered approach to technology that takes into account the following considerations: Recognizing the new possibilities and challenges that may arise with each new technological advancement; and becoming equipped with the skills and knowledge on how to use these advancements effectively, responsibly, and safely. This understanding involves as well unlocking new and unique skills within humans and expanding our definition of ‘what it means to be human’. One more way to attain this understanding comes through developing new ethical and responsible frameworks for navigating the complex territory of technology and conceiving its impact on our social structures and relationships. Approaching technology with these considerations in mind, only then –the study concludes– can humans shape technology in a way that aligns with their values and aspirations as humans, and can a symbiotic, trusting, and mutually beneficial relationship be created between technology and humanity in the future, achieving tremendous accomplishments that serve the common good.

5. References

- Asimov, Isaac. *I, Robot*. New York: Doubleday & Company, 1950.
- Bailey, Ronald. 'Will super intelligent machines destroy humanity?' *Reason*, 2014.
<https://www.reason.com/2014/11/27/will-superintelligent-machines//>
(Retrieved on 3July, 2023).
- Barrat, James. *Our Final Invention: Artificial Intelligence and the End of the Human Era*. New York: Macmillan, 2013.
- Bostrom, Nick. 'A history of transhumanist thought.' *Journal of Evolution and Technology*, 14(1), pp.1–25, 2005.
<https://www.nickbostrom.com/papers/history//> (Retrieved on 17June, 2023)
- Bostrom, Nick. *Super Intelligence : Paths, Dangers, Strategies*. Oxford: Oxford University Press, 2014.
- Burnett, Dean. 'No, there hasn't been a human 'head transplant', and there may never be.' *The Guardian*, 2017.
<https://www.theguardian.com/science/brainflapping/2017/nov/17/no-there-hasnt-been-ahuman-head-transplant-and-may-never-be-sergio-canavero//> (Retrieved on 17June, 2023).
- Chalmers, David. 'Are we living in a simulation?' *The Guardian*, 2022.
<https://www.theguardian.com/books/2022/jan/19/reality-by-david-j-chalmers-review-are-we-living-in-a-simulation//> (Retrieved on 16June, 2023).
- Clowes, Brian. 'Transhumanism: The final frontier'. *Human Life International*, 2020. <https://www.hli.org/resources/what-is-transhumanism//> (Retrieved on 3July, 2023).
- Dargis, Manohla. 'I am my own monster: (Technology rules).' *The New York Times*, 2014. <https://www.nytimes.com/movies/johnny-depp-stars-in-transcendence-html//> (Retrieved on 3July, 2023)
- Dick, Philip K. *Do Androids Dream of Electric Sheep?* New York: Ballantine Books, 1968.
- Esfandiary, F.M. *Are You a Transhuman? Monitoring and Stimulating Your Personal Rate of Growth in a Rapidly Changing World*. New York: Warner Books, 1989.
- Falk, Dan. 'Is your brain a computer?' *MIT Technology Review*, 2021.
<https://www.technologyreview.com/2021/08/25/1030861/is-human-brain-computer/MITNews Magazine//> (Retrieved on 17June, 2023).
- Gibson, William. *Neuromancer*. New York: Ace, 1984.

- Good, Irving John. 'Speculations concerning the first ultra-intelligent machine.' In F. Alt & M. Rubinoff (eds.), *Advances in Computers (Vol.7)*, pp. 31– 88. New York: Academic Pres, 1965.
- Hameroff, Stuart & Penrose, Roger. 'Consciousness in the universe.' *Physics of Life Reviews*, 11(1), pp.39–78, 2014.
- Heaven, Will Douglas. 'Artificial intelligence: What an octopus mind can teach us about A.I.'s ultimate mystery.' *MIT Technology Review*, 2021.
<https://www.technologyreview.com/2021/08/25/1032111/conscious-ai-can-machines-think/> (Retrieved on 17 June, 2023).
- Huxley, Julian. 'Transhumanism.' In Julian Huxley (ed.) *New Bottles for New Wine*, pp.13–17. London: Chatto & Windus, 1957.
- Jefferson, Geoffrey. 'The mind of mechanical man, Lister oration for 1949.' *British Medical Journal*, I, pp.1105– 1121, 1949.
- Koch, Christof. 'Neuroscience: What is consciousness?' *Nature* 557(7704), pp.8-12. New York: Macmillan Publishers Limited, 2018. Doi: 10.1038/d41586-018-05097-x
- Kurzweil, Ray. *The Age of Spiritual Machines*. New York: Viking Press, 1999.
- Kurzweil, Ray. *The Singularity is Near: When Humans Transcend Biology*. New York: Viking Books, 2005.
- Mandel, Emily St. John. *Station Eleven*. Canada: Knopf, 2014.
- Mandelbaum, Eric. 'Everything and more: The prospects of whole brain emulation.' *Journal of Philosophy*, 119 (8), pp.444–459, 2022.
- McEwan, Ian. *Machines like Me*. United Kingdom: Jonathan Cape, 2019.
- Moore, Gordon E. 'Cramming more components onto integrated circuits.' *Electronics*, 38(8), pp.114–117, 1965.
- Naam, Ramez. *More than Human: Embracing the Promise of Biological Enhancement*. New York: Broadway Books, 2005.
- Orf, Darren. 'Humanity may reach singularity within just 7 years, Trend shows.' *Popular Mechanics*, 2023.
<https://www.popularmechanics.com/technology/robots/a42612745/singularity-when-will-it-happen/> (Retrieved 3 July, 2023).
- Oxenberg, Richard. 'Is Consciousness an "emergent property"?' *Oxenberg Manuscript-OXEICA*, 2021. <https://philpapers.org/archive/OXEICA.PDF/> (Retrieved on 23 June, 2023).
- Paglen, Jack. 'Transcendence.' *Jack Paglen Scripts.Com*. Web. [Place of Publication not identified] [Publisher not Identified] ,2014.

- https://www.scripts.com/writer/jack_paglen/12725// (Retrieved on 23June, 2023).
- Pasnau, Robert. *Thomas Aquinas on Human Nature: A Philosophical Study of Summa Theologiae, I a*, pp. 75–89. Cambridge: Cambridge University Press, 2001.
- Roden, David. *Posthuman Life: Philosophy at the Edge of the Human*. London: Routledge, 2014.
- Russell, Stuart. *Human Compatible: Artificial Intelligence and the Problem of Control*. London, United Kingdom: Penguin Books, 2020.
- Searle, John. 'Consciousness and the philosophers.' *The New York Review of Books*, 1997.
- Shelley, Mary. *Frankenstein; or, The Modern Prometheus*. England: Lackington, Hughes, Harding, Mavor & Jones, 1818.
- Silkin LLP, Lewis. 'Artificial Consciousness: What is it and what are the issues?' *Future of Work Hub*, 2021. <https://www.futureofworkhub.info//> (Retrieved on 23June, 2023).
- Stephenson, Neal. *Snow Crash*. U.S.A.: Bantam Books, 1992.
- Vinge, Vernor. 'The coming technological singularity: How to survive in the posthuman era.' In Proc.(conference) *Vision-21: Interdisciplinary Science and Engineering in the Era of Cyberspace*, pp. 11–22. NASA: Lewis Research Center, 1993.
- Warwick, Kevin. 'The fusion of man and machine.' *CNN Technology Vision*, 2005. <https://www.cnn.com/2005/TECH/05/12/visionary.warwick//> (Retrieved on 23June, 2023).
- Zhegunov, Gennadiy, et al. 'Singularity hypothesis: A scientific and philosophical assessment.' In Eden, A.H.; Moor, J.H.; Soraker J.H. & Steinhart, E. (eds.), *The Frontiers Collection* pp.1–2. Berlin, Dordrecht: Springer, 2012.

"الفرص [غير] المتاحة: التشكيك في فرضية 'الوعي الاصطناعي'، ونموذج 'السمو'،
في روايه السمو' لجاك باجلين"

نيفال نبيل محمود عبد الله

قسم الانسانيات- كلية اللغة و الاعلام- الأكاديمية العربية للعلوم و التكنولوجيا و النقل البحري (فرع
مصر الجديدة)، جمهورية مصر العربية.

nova2004@hotmail.com

المستخلص:

مع التطور السريع للذكاء الاصطناعي في وقتنا الحاضر، تثير فكرة إحياء المادة أي مزج الميكانيكا مع الأحياء، وخلق نُسخ مطابقة بل و غير مسبوقه للدماغ البشري تحمل خصائص التفرد جدلا واسعا و تفاعلا كبيرا في الدراسات الأكاديمية الجاده بالإضافة إلى الخيال العلمي. فقد أثارت فرضية 'الوعي الاصطناعي'، الانقلابيه مجموعه من التساؤلات الحاسمه و التي من اهمها: هل يمكن دمج وتضمين الوعي الإنساني مع آلات الذكاء الإصطناعي؟ هل يمكن لهذه الآلات ان تتمتع بالاستقلاليه و الحساسيه التي يتسم بها بنو البشر؟ وهل بإمكان هذه الآلات التعبير عن إحتياجاتها و تفسير تجاربها الذاتية و عواطفها المميزه و ذكرياتها و عمليات تفكيرها و معتقداتها؟ لهذا الغرض، تسترشد الورقه البحثيه الحاليه بشكل أساسي بنهج 'ما بعد الإنسانيه' المتفائل، و نظرية "التفرد التكنولوجي" لـ راي كورزويل، بهدف استعراض التداعيات غير المقصوده لخرق ما أبدعته الطبيعه، و التحقيق في فرضية 'الوعي الإصطناعي'، ومدى معقولية دمج الوعي الإنساني بشكل كامل مع آلات الذكاء الإصطناعي، مع التشكيك في نهج 'ما بعد الإنسانيه'، والذي ينظر للتكنولوجيا على أنها شكل من أشكال السمو. لهذا الغرض، تقوم الدارسه بإجراء تحليل نصي متعمق و دقيق لرواية الخيال العلمي 'السمو' (٢٠١٤) لجاك باجلين، بهدف الوصول في النهايه إلى إستنتاج مفاده أن التكنولوجيا لم تنزل بعيدة عن تحقيق فرضية الوعي الإصطناعي، لوجود كيان جوهرى مميز وفريد في الوعي البشري لا يمكن تكراره أو محاكاته بواسطة التكنولوجيا ألا وهو، الضمير الإنساني.

الكلمات المفتاحيه: فرضية 'الوعي الاصطناعي'، جاك باجلين، الآلات الواعيه، مبدأ 'التفرد التكنولوجي'، نموذج 'السمو'، نهج 'ما بعد الإنسانيه'