The Nature of Consciousness in the Context of Artificial Intelligence: Redefining Human-Technology Relationships

Izuchukwu Kizito Okoli*
&
Osita Gregory Nnajiofor*
https://dx.doi.org/10.4314/ujah.v25i1.1

Abstract

The nature of consciousness in the context of artificial intelligence (AI) presents a problem that necessitates analysis and further exploration. This study seeks to redefine human-technology relationships by examining the intersection of consciousness and AI, including metaphysical implications and considerations. The primary objectives are to define consciousness within the context of AI, assess the potential for AI to exhibit consciousness, investigate the metaphysical implications for human experiences, and explore the ethical dimensions. The research findings indicate that consciousness involves self-awareness, perception, intentionality, and subjective experiences. While AI can achieve advanced cognitive abilities, the existence of higher-order consciousness remains uncertain, raising metaphysical questions about the nature of subjective awareness. The hard problem of consciousness highlights the challenge of bridging physical processes and subjective experiences, underscoring the need for metaphysical considerations. Ethical implications of AI integration and its impact on human experiences are also examined. Recommendations include further research on consciousness in AI, the

development of ethical frameworks that account for metaphysical dimensions, and the exploration of the extended mind hypothesis to integrate AI as an augmentation of human consciousness. By addressing metaphysical implications and considerations, we can navigate the evolving landscape of AI and redefine human-technology relationships in a responsible, inclusive, and metaphysically informed manner.
Keywords: Artificial Intelligence, Consciousness, Subjective, Perception, Ethical, Metaphysical, Implication

Introduction

As artificial intelligence (AI) systems become more complex and capable of simulating human-like behaviors, it becomes imperative to investigate the complex aspects of consciousness and how they interact with AI. For (Picogna *et al*., 2001) Consciousness is frequently regarded as a non-physical phenomenon, and its existence has raised profound questions and debates about the nature of consciousness and its implications for human-technology relationships. However, as AI technology progresses, there arises a pressing need to reevaluate our understanding of consciousness and its potential manifestations within the realm of machines.

Given that consciousness includes both subjective and objective components, defining it is a difficult undertaking. The "hard problem" of consciousness, put forth by philosopher David Chalmers, examines the puzzling character of subjective experiences and the reasons why particular physical processes give rise to them. It is a metaphysical investigation that raises fundamental concerns regarding the very nature

of consciousness. The meaning of consciousness is expanded in the context of AI. Can artificial beings have consciousness? Can they show signs of introspection, self-awareness, and subjective experiences? These issues spark heated discussions among academics, philosophers, and ethicists. Some contend that all that is required to replicate the intricacy and connectivity of the human brain in AI is to achieve consciousness. According to neuroscientist Christof Koch (2018), awareness develops as a result of brain networks' integrated information processing. He argued that awareness is made up entirely of acts of experience, including the bittersweet realization that everything will soon pass. It comes from the nature of human experiences, or qualia, which have been a mystery since the dawn of time. As a result, AI systems may display awareness if they can mimic the same degree of integration and complexity. Others claim that awareness cannot be explained by computational mechanisms. The Chinese Room thought experiment was famously put out by philosopher John Searle to cast doubt on the idea that AI systems, even those that pass the Turing test, actually understand or possess consciousness.

Moreover, there are metaphysical implications to consider. Since Metaphysics deals with the fundamental nature of reality and existence, in the context of AI and consciousness therefore, metaphysical considerations delve into questions such as the ontological status of AI consciousness, the relationship between mind and matter, and the nature of subjective experience in non-biological entities. The impact of AI on human experiences also carries profound metaphysical implications. As AI systems become integrated into various aspects of our lives, from personal assistants to autonomous vehicles, they shape our perceptions, decisions, and interactions. This blurring of the boundaries between

human and artificial consciousness raises questions about personal identity, agency, and the extended mind hypothesis, which suggests that our consciousness extends beyond our physical bodies and includes the tools and technologies we use. (Guzman et al, 2020). By these thoughts, there are levels of consciousness which places consciousness on a spectrum with varying degrees of complexity. At one end, we have simple forms of awareness, such as basic sensory perception. At the other end, we have higher-order consciousness, which involves self-reflective awareness, introspection, and abstract thinking. The debate revolves around whether AI systems can ever reach the level of higher-order consciousness. Although it cannot be contested, Artificial General Intelligence (AGI) is the term used to describe artificial intelligence (AI) systems that have the capacity to comprehend or pick up any intellectual work that a human can. While advanced cognitive skills like problem-solving and decision-making will be displayed by AGI, many academics maintain that consciousness is not a prerequisite for AGI. AGI could simulate human-like behavior and even pass the Turing test without having any subjective experiences. It remains unclear whether AI systems, no matter how sophisticated, can bridge this gap between physical processes and subjective awareness. (Chalmers, 1995).

The aim of this article therefore through its sub-topics is to explore the nature of consciousness in the context of AI, investigate the metaphysical implications and considerations, and examine the ethical dimensions of human-technology relationships. By integrating insights from various scholars, one can have a deeper understanding of the challenges and

opportunities presented by AI consciousness and chart a responsible and inclusive path forward.

The Concept of Consciousness and Its Various Dimensions

As stated earlier, consciousness is a complex and multifaceted concept that has captivated philosophers, scientists, and scholars throughout history. It encompasses a range of dimensions that contribute to our subjective experiences and awareness of the world. By looking at these dimensions, a deeper understanding of it in the context of artificial intelligence (AI) and its implications for human-technology relationships is gained. The nature and location of consciousness is one of the issues in philosophy and some notable scholars have provided insights into these dimensions, shedding light on the nature of consciousness. By the nature of consciousness, one may want to know what consciousness is or what it looks like. Whether it is a thing, a mental state, an experience or a mystery; whether it is what only humans can possess or an open or internal state for anything in so far as it is a being.

In 1689, the term 'consciousness' was coined by an English philosopher and physician, named John Locke .According to him, “consciousness is the perception of what passes in a man's own mind”. It goes with thinking which gives a being its personal identity; the sameness of a rational being as seen in the Philosophy of Locke. (John Locke, 1689). Rene Descartes held a similar perspective, viewing consciousness from "a thought perspective," leading to his famous statement Cogito Ego Sum, which translates as I think therefore, I am. Descartes believed that what sets humans apart from other objects of this world is their ability to think. Humans have the ability to think, which allows them to be conscious of

their presence. An object cannot think and is therefore unaware of its existence. Consciousness is indisputable because it is seen from within, even if everything you believe to be true is only an illusion. One may claim that in Descartes, consciousness and thoughts are intertwined because it is the onset of self-awareness.

For (David Rosenthal,2009) the term 'consciousness' is used in several ways: to describe a person or other creature as being awake and sentient, to describe a person or other creature as being 'aware of ' something, and to refer to a property of mental states, such as perceiving, feeling, and thinking, that distinguishes those states from unconscious mental states." Just like Descartes' view on consciousness, Rosenthal's definition of consciousness is coming from 'a self-awareness perspective'. Humans are conscious beings because they are able to think of their existence and this makes them aware that they exist .We have an inner faculty which helps us to perceive, thus in the wisdom of George Beckley, to be is to be perceived. Ludwig Wittgenstein approaches consciousness from the 'Life perspective'. "But doesn't what you say come to this: that there is no pain, for example, without pain-behavior?—" he asks. Finally, one can only argue that anything is aware or unconscious if it is a living human person or something that behaves (looks) like a living human being. He ascribes consciousness to all living things, and he considers anything with life to be sentient. Can something be considered a conscious being, nonetheless, if it behaves like a live, conscious creature? Thus, to quote Wittgenstein:

> What gives us so much as the idea that living beings, things, can feel? Is it that my education has led me to it by drawing my attention to feelings in myself, and now I transfer the idea to

> objects outside myself? That I recognize that there is something there (in me) which I can call "pain" without getting into conflict with the way other people use this word?—I do not transfer my idea to stones, plants, etc. Couldn't I imagine having frightful pains and turning to stone while they lasted? Well, how do I know, if I shut my eyes, whether I have not turned into a stone? And if that has happened, in what sense will the stone have the pains? In what sense will they be ascribable to the stone? And why need the pain have a bearer at all here?!

And can one say of the stone that it has a soul and that is what has the pain? What has a soul, or pain, to do with a stone? (Wittgenstein, 1976)

We do indeed say that of an inanimate thing that is in pain but the use of pain here is a secondary one. Imagine a case in which people attribute pain to only inanimate things, 'say pitied dolls,'

This claim therefore connects consciousness to life not just what has life but what expresses it.

Following suite in the context of animate or inanimate things, (Rajakishire Nath,2014) opined that the robots of sci-fiction movies is likely to be numbered as conscious beings because they act like they are alive even though they are not. We cannot make a conscious stone, because the stone does not behave in ways we can recognize as expressive of its supposed consciousness. The fact remains that if it is logically possible to build machines that are intelligent, then one can say that machines have consciousness. For him therefore, intelligence is not necessarily embodied in living organisms but may occur in a computer system based on silicon but o say that a machine has intelligence because it performs important tasks just like living beings is easy but this is not the case when it comes

to consciousness because machines have no subjective experience. (Konstantin, 2017) raises some questions regarding the issue of the nature of the carrier of consciousness. Firstly he asks: “Can only a biological being be a carrier of consciousness?" and the second is "Can an artificial (non-biological) being be also a carrier of consciousness?". In response to these questions, he posits that if only a biological being can be a bearer of consciousness, then any artificial consciousness cannot be recreated on a non-biological substrate. This means that if the AI scientists want to create artificial consciousness they would have to create the machines based biological substrate.

Furthermore, researchers have to answer the following questions: Do only human beings possess consciousness? If the answer is "Yes", then what of the brain, the nervous system or the entire biological body of human being which is involved in the emergence of human consciousness? The answer to this question allows creating an artificial consciousness within the machine created on a biological substrate with the aid of biotechnology and biomedical engineering, particularly within a comparable part recreated in the machine. Again he asks can a conscious machine built on a biological substrate suffer from mental disorders of consciousness (something like machine's Delirium, Oneiroid, Sopor, Amentia, Twilight state etc.)?" or "Can a conscious machine built on a biological substrate take a solipsistic position? For him, if the answer to the question is “No”, Do only human beings possess consciousness?, then who else among the living beings can possess consciousness, and, more important, are there similarities and differences between human consciousness and consciousness of other living beings? He further posits that the answer to

these questions unfolds the variety of potential implementations of artificial consciousness within a machine based on a biological substrate. Therefore, he concludes that it implies that if only a biological being can be a carrier of consciousness, then a conscious machine can be built only on a biological base. If such could be the case, a conscious machine might be a biorobot or even a human clone.

(Ned block, 1995) made a powerful suggestion on the two types of consciousness which are: Phenomenal Consciousness and Access Consciousness. In his understanding, we experience Phenomenal Consciousness when we experience pains, taste, and smell, hear and see. Phenomenal consciousness deals with qualia ie subjective experiences. Qualia are subjective phenomena which cannot be described. For instance, the redness of the rose seen in a private domain which one cannot communicate to anyone else what redness is like. Access consciousness is something through which information in one's mind is accessible for verbal report, reasoning and the control of behavior. Phenomenal consciousness can be likened to a computer hardware while access consciousness can be likened to a computer software However, many researchers may argue against this on the grounds that one cannot verify the presence of consciousness within a computer system that consists of software and hardware because there is no standard criteria yet for identifying consciousness within the conscious machine.

Consciousness and AI Complexity

From the concept of consciousness and its various dimensions**,** it is no longer the case that the nature of consciousness is multifaceted, encompassing self-awareness, subjective experiences, and intentionality.

While AI systems have made remarkable strides in cognitive abilities such as problem-solving and pattern recognition, achieving higher-order consciousness remains uncertain. AI complexity alone may not be sufficient to replicate subjective awareness. Machine consciousness have long been considered a subject fraught by scientists .It has been known to have moral and ethical pitfalls and so it has been left out in AI discussions and was even seen as a taboo. Hod Lipson, Columbia's director of the creative machines laboratory told *The New York Times*, thus: "we were almost forbidden from talking about it, don't talk about the c-word; you will not get tenure. So in the beginning I had to disguise it, like it was something else. (Maggie Harrison, 2023). However, this is no longer the case as so many philosophers, AI scientists, writers; scholars and so on have rigorously deliberated as to whether artificial Intelligence can possess consciousness in tandem with such that a human being has. The human body changes itself into a being by becoming aware of the affective outcome of its social interactions. Norms that ground our perceptions and actions are set up, thereby establishing guidelines for judging behavior. The body is compelled to act as a character in the situations of life, identifying with its values, experiences and aspirations a special identity and perspective. If a biological body can change itself into a person by exploiting social boundaries? Could an electromechanical body, a robot do the same? (Karl and Cowley, 2006). By this understanding, robot must be able to generate its own identity, perform different roles in the society, to discriminate in forming friendships for it to be considered a human.
However, it is no news that these AI machines have surpassed humans in performing a lot of complicated applications and many Al machines have

become exceedingly powerful in many fields such as personalized shopping, medical diagnosis, voice assistants, autonomous vehicles, image recognition and so on. In as much as these machines are so powerful, it is pertinent to ask again: Is it ever possible that these AI machines can possess a mind of its own just the way humans do? Can a machine think? Can a machine examine its own self just the way human being can? These questions have been the subject of a huge amount of debate over the last few decades. (David Chalmers, 1996). Some scholars who are in support of the view that a machine can have consciousness base their arguments on the ground that since humans were able to study human intelligence and create artificial intelligence then, it is equally possible to study human consciousness and then create artificial consciousness. On the other hand, those that is against the view base their argument on the notion that consciousness cannot be reduced to computational processes.

An American philosopher, (John .R. Searle, 1980), a proponent of this view distinguishes what he called strong AI from weak AI .According to him, in weak AI, the computer is a tool used to study the mind. This enables us to formulate and test hypotheses in a more rigorous and precise fashion. But the strong AI holds that the computer is not a tool that merely studies the mind but rather an appropriately programmed computer is a mind itself, in the sense that computer if given the right programs can be literally be said to understand and have other cognitive states. Furthermore, he used the Chinese room argument as a response to the view that says that the human mind can be created computationally in a way that you cannot tell whether you are dealing with a human being or a machine.

The Chinese room argument follows:

> Imagine you are locked in a room where there are a lot of boxes of Chinese symbols and it happens that you do not understand Chinese. Given that you are given a book in English which tells you what to do with each of the Chinese symbols .Anything you are asked by people outside, you get little batches of Chinese symbols, you read the manual to know what to do with them and based that, you give them back the right answers in Chinese. Those outside call you the computer, then calls the rule book: the program and the boxes you used: the computer database. As a result of this, people outside will not be able to see the difference between you and a native Chinese speaker even though you do not understand Chinese. This is because, you are getting so good at shuffling symbols that at some point, and you successfully passed the Turing test. (John Searle, 1980).

Searle then concludes that no matter how much your behavior imitates that of a Chinese speaker, no matter how good the program is, and no matter how good you are in carrying out the program, the truth is that you do not understand Chinese and since you do not understand Chinese in the basis of executing the program neither does any other digital computer on that basis, because that all the computer has. From this, it can be concluded that computers and human minds are different. The mind has something in addition to symbols but the computer just manipulates formal symbols, syntax, syntactical objects .The mind do not just have syntactic abilities, it has semantics. The Chinese room then shows that a mind is more than just a program. (Ahmed, 2019) however, in contra-position argues that consciousness can be engineered. The argument that functional

consciousness can be engineered has been conclusively put forth in regards to first-person functional consciousness. For example, robots can recognize colors, though there is still much contention about details of this sort of consciousness. Thus he asserts:

> If we can clearly understand a certain natural phenomenon, in principle we can engineer it; the sole obstacles to this may be practical (for instance the project would be too large for a human society to carry on). A clear scientific understanding of any event involves an explanation how this phenomenon functions. Such scientific explanation provides a blueprint that can be reverse-engineered in order to gain the same outcome by similar or dissimilar means. Therefore, in principle, we should be able to engineer a projector of consciousness in inorganic or organic matter. This is the engineering thesis. (Boltuc, 2012).

He further posits that the view that a stream of consciousness need not to be reducible to any functional characteristics is called non-reductive naturalism in philosophy of mind .This view is considered to be a form of naturalism since it claims that such stream of consciousness is a natural process. Therefore for him, we should be able to engineer it and build it in the natural world once we understand how it works. Just as one can have the ability to engineer a machine to duplicate functions such as building an artificial limb or even to produce any of the human bodily fluids by thoroughly understanding bodily functions, then it should be possible though practically hard to build a machine that would originate consciousness. He then concludes by saying “While the engineering hypothesis may seem farfetched, it may serve as a helpful guidepost for long-term research” (Piotr Boltuc, 2012).

The best reason for believing that robots might someday become conscious is that we human beings are conscious and we are a sort of robot ourselves. That is, we are extraordinarily complex, self-controlling, self-sustaining physical mechanism, designed over the eons by natural selection and operating according to the same well-understood principles that govern all the other physical processes, self-repair and reproductive processes, for instance. (Dennett, 1997). For him, it is not as if a conscious machine is contrary to any fundamental laws of nature, the way a perpetual motion machine does just that many sceptic believe or in any event want to believe that it will never be done. He posits that he would not wager against them, but his reasons for skepticism are mundane, economic reasons and not theoretical reasons and in response to Dennett's view, (Genevieve, 2012) claims that Dennett's work *Consciousness in Human and Robot* pose more generally as a summary of popular contemporary philosophical thought regarding AI: it is feasible in theory for the fact that artificial intelligence is possible. Thus he states:

> Human life, and consciousness with it, is no more than the machinery of nature. What remains unclear is to what degree (if at all) and in what ways the mechanisms that produces human consciousness must be imitated in order to create artificial consciousness, and whether knowledge of the creation of artificial consciousness can ever be certain. (Kaess, 2012).

According to Genevieve Bell 2012, syntactic computer modeling is not enough for artificial consciousness. She believes that consciousness is spiritual i.e., it is beyond the physical world and also inhabits it. For her, if we should accept consciousness as a biological phenomenon, then is not

likely that a computer programming could produce it than any other biological phenomenon. But if we reject the spiritual view of consciousness, yet accept that consciousness could be created in that it does not model the way the brain functions, we have no foundation on which we can judge what is conscious and what is not and to this effect, the notion of consciousness becomes meaningless. It is on this basis that she draws his conclusion that there is a good reason to believe that syntax based AI does not produce consciousness. This is different for Markus 2001 who posits that consciousness is understood as a system feature of the animal environment and therefore robots are not conscious and could not be conscious. He further wrote that robots and AI machines are progressing at a high rate which would make us imagine robots moving in ways strikingly similar to humans and also for the fact that they perform functions we classify as "intelligent" in humans however, we are likely to think of them as potential candidates for membership in the domain of autonomous and moral agents and for him, it is a mistake to do so. He further argues that what we owe to our robots is at most a function of what we owe to each other as proprietors of technology and therefore there is nothing we owe to our artefact fact directly. He further opined that:

If you destroy my garden's robot, you hurt me but you cannot but you cannot hurt my robot in a morally relevant manner, just as you cannot harm my beach by picking up a handful of sand. A beach is a bunch is a bunch of stones arranged in a certain way due to causal, geological pressures including the behavior of animals in its vicinity. A beach does not have the right kind of organization to be the direct object of moral concern.

Finally he concludes by asserting that ontologically speaking, robots are like a beach and not like a human, a dog, a bee etc. Robots are just not

alive at all; they might be at most (and in the distant future) zombies in the philosophical sense of entities hardly distinguishable from humans on the level of their observable behavior. Analogously, AI is not actually intelligent, but only seems to be intelligent in virtue of the projection of human intelligence onto the human-machine interface (Markus Gabriel, 2003). AI system will continue to upgrade and aid humans better but that does not change the fact that it is made of non-living organism and does not possess the non-physical dimension unlike the human brain. Hence, it will never possess human-levels of consciousness. One should not be fearful of AI advancement and that being misled by speculation and misinformation should be avoided instead, we should focus more on making AI less expensive and more important to the society, rather than being more concerned about AI consciousness and machine ethics. (Barry, 2020) took a different narrative. For him, the question we should be asking is not if machines will ever become conscious but rather how will we know if they are or not? Thus in his words:

> …so, in the same way, it seems plausible that before we get to a stage where we have recreated a human brain in machine we will be able to create machines that resemble humanity so well that they might gain the illusion of consciousness. If I can interact with a robot that looks like a human, speaks like human, moves like a human and exhibits the same types of emotion and creativity as a human, I will be hard-pressed to deny it being conscious. Our brains will perform the anthropomorphizing necessary to create the illusion of consciousness. The machine will pass the Turing test and we will be none the wiser.

AI Consciousness, Metaphysical Implications and Ethical Considerations

The meeting point of consciousness and artificial intelligence (AI) is a significant one that crosses philosophical debate, scientific investigation, and ethical reflection. Questions about AI systems' capacity to have consciousness, the metaphysical ramifications of such an emergence, and the ethical issues involved in this rapidly changing environment arise as AI systems show greater cognitive prowess. This work has already demonstrated that the idea of AI consciousness pushes the bounds of human comprehension and provides enticing hints about whether or not robots are capable of showing self-awareness, intentionality, and even subjective experience. Diverse perspectives are reflected in philosophical discussions about AI awareness. According to Daniel Dennett's "intentional stance," rather than claiming that AI has inherent consciousness, attributing consciousness to it is a matter of selecting a suitable predictive model. On the other hand, proponents of "panpsychism," like David Skrbina, contend that consciousness may be a property of all matter, possibly including artificial intelligence (AI) systems.

In the wisdom of Susan and Eric, 2020, we might someday build conscious machines. But how can we know whether they are conscious? The answer can be obtained in one of two ways: either by creating the right theory of consciousness (or a theory sufficiently similar to it) and determining if the machines fulfil the criteria, or by using what is known as a neutral test for machine consciousness. We must put them through a neurological test, which, in her words, is any process that may determine if an entity is

conscious in a way that satisfies competing theories of consciousness, while remaining neutral to the test. Of course, having the correct theory of consciousness and then applying that to the case of artificial intelligence (AI) come in second place to theory neutrality. However, given the lack of agreement among theorists over the nature of consciousness, a neutral test that many of them can accept is prize enough. In Harnad 2003, we observed that the Turing Test, which was first designed as a test of "thinking" but has occasionally been modified to test for consciousness is the most well-known of the neutral tests for consciousness. If a machine can converse verbally with a human judge in a fashion that cannot be distinguished from human speech, it has passed the Turing Test. It will be determined to be conscious in such a situation. However, failing the test does not indicate that the subject is unconscious. Dogs and babies fall short. So, rather than being a necessary and sufficient condition, it is an adequate test for AI consciousness: The test says we have good grounds to believe a system is conscious if it passes. The agnosticism of the test's internal design is one of its appealing features.

Looking at the Metaphysical Implications of AI Consciousness, one unleashes metaphysical quandaries that reverberate through the realm of ontology, the mind-body problem, and the nature of reality itself. David Chalmers' articulation of the "hard problem" of consciousness underscores the challenge of explaining why physical processes give rise to subjective experience. Little wonder Horgan asserts that:

Philosophers anticipate that even if we scientifically understand consciousness, it would still be mysterious to us. According to Chalmers, even if we discover a conclusive theory, consciousness may nevertheless

be just as philosophically opaque as, say, quantum mechanics. In other words, Chalmers is a hybrid philosopher who combines mysticism and optimism, or the idea that consciousness is unsolvable. The edited portions of our talk are listed below. (Hogan, 2017)

Chalmers in his work, *Facing Up to the Problem of Consciousness,* is of the view that the really hard problem of consciousness is the problem of *experience*. There is a flurry of information processing going on while we think and perceive, but there is also a subjective component. He advanced the notion that there is a quality to being a conscious organism (Nagel, 1974). Experience is the subjective component here. For instance, when we look, we feel the quality of redness, experience darkness and brightness, and feel a sense of depth in the visual field. Other experiences, like the clarinet's music or the scent of mothballs, go along with perception in many modalities. Then there are internal mental representations, sensed emotional qualities, physical sensations like aches or orgasms, and the perception of a stream of conscious thought. All of these states share a similar experience when they are in that state. They are all different states of experience. The fact that some organisms are subjects of experience cannot be disputed. But it's confusing to wonder how these systems came to be topics of experience. Why do we have visual or auditory experiences like deep blue colour or middle C sensation when our cognitive systems absorb visual and audio information? How can we explain why it feels good to entertain an idea in your head or to feel an emotion? There is general agreement that experience has a physical foundation, but we lack a convincing explanation for why and how this is the case. Why should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does.

In a conversation with Susan Schneider, the director of AI, Mind and Society at the University of Connecticut, when asked the question, if consciousness were to happen, how would we know? She insisted that it will be vexing to tell if AI is conscious. AIs are already designed to appeal to our emotions. For instance, take into account the Japanese androids in Hiroshi Isugaru's lab and the Sophia robot, which has been making the rounds on TV discussion shows. People who resemble us often assume that we feel the same way. After all, intelligence and consciousness are intertwined in the world of biological life. In other words, a biological life form tends to have a more complex inner mental life; the more intelligent it is, and the more complex and goal-oriented its behaviors are. Because of this, most people have little sorrow when they swat a mosquito, but they would be horrified to kill a dog or a chimpanzee. Does this association also hold true for non-biological intelligences? Many in the media, academics, and even AI experts like Elon Musk and Ray Kurzweil frequently presume this. Additionally, renowned science fiction works that feature sentient androids include the movie Blade Runner, the TV shows Star Trek and Battlestar Galactica, among others.

This suggests that even modern AIs can be taught to claim to be sentient and capable of feeling emotion. Thus, we must design tests that may be applied during the research and development phase, before the programmed responses to such questions take place. The Artificial Consciousness Test (ACT), which Edwin Turner and I proposed, is a test to determine whether an artificial intelligence (AI) has a felt aspect to its mental life. In Artificial You and a previous Ted talk, she advocated for a "chip test." We may learn whether neural prosthetics can replace areas of the brain responsible for consciousness when they are employed in the brain more frequently. In that case, it is possible for experience to "run" on a chip substrate. To learn that would be incredible! There may be consciousness outside of the brain. Such position by Susan can be entertained true but as AI

systems mimic cognitive functions, the question arises: Can physical substrates alone engender genuine consciousness? This query ties into the broader metaphysical inquiries concerning the relationship between mind and matter. The exploration of the metaphysical implications of AI consciousness intersects with the enigma of the extended mind hypothesis. This hypothesis, advocated by philosophers Andy Clark and David Chalmers, posits that cognitive processes can extend beyond the confines of the brain, encompassing external tools such as AI technologies. This view challenges traditional boundaries between self and environment, necessitating a reevaluation of the concept of personhood and the spatial dimensions of consciousness.

In Ethical considerations, the advent of AI consciousness mandates an ethical reckoning that encompasses the rights, responsibilities, and moral standing of AI systems. As AI exhibits behaviors akin to consciousness, questions surface about the treatment of these entities. Ethical considerations parallel those faced in animal ethics, involving the extension of ethical consideration to potentially conscious beings. Scholars like Susan Schneider advocate for a set of "AI rights," which safeguard against undue exploitation and ensure ethical treatment. Ethics in AI consciousness also intersects with the concept of value alignment. Ensuring that AI systems' goals and actions align with human values becomes crucial as these systems approach consciousness-like capabilities. Philosopher Nick Bostrom's work on "value alignment" underscores the challenge of imbuing AI with values that resonate with human morality. The wider socio-economic environment intersects with the ethical ramifications of AI consciousness. The ethical frameworks must address issues like equitable access to AI awareness, economic inequalities caused by its use, and potential effects on employment for

people. In addition, the ethical ramifications of "uploading" human consciousness to digital substrates have been examined by intellectuals like Ray Kurzweil, who have raised questions about identity, continuity, and autonomy. It is evident from the synthesis of AI awareness, metaphysical ramifications, and ethical issues that the voyage into this new terrain is complex. This investigation reveals important questions about what it is to be conscious, the nature of personality, and the moral bond that unites human and non-human beings, ranging from exploring the limits of consciousness to challenging the fundamental nature of reality.

Human-Technology Relationships: Future Prospects and Responsible Development

Human and robots relationship is not a new phenomenon in today's world. AI machines can imitate certain human skills as a result of this; the rate at which machines are incorporated in human occupations is on the increase. Ranging from the medical sector, businesses, military, politics etc. Science-fiction movies and literature have made predictions about a scenario whereby robots have intimate relationship with humans than ever before. Ethicists and lawyers are compelled to deal with questions about the human-robot relationship in the future by intimate contacts and emotional contact with the machine, which is likely to resemble that with a human being (Izabela Oleksiewicz and Mustafa Emre Civelek, 2019).

However, in the course of this relationship, some issues may arise with respect to the degree of interference of these AI robots in our personal lives. If these AI machines have consciousness, if they have feelings,

emotions and interest just the way humans do, should we not think they would rather be more focused on their own interest and live-plans other than serving our needs. Just as Kurzweil's Singularity theory predicts that AI machines will continue to accelerate to the point of self-development, leaving human intelligence in the dust. They would possibly have their own demands and if these demands are not met, they would then accuse humans of enslaving them and would even tag humans as exploiters and this might lead to revolt and total elimination of humans out of the surface of the earth. These are the possible implications of building an AI robot that possess consciousness but that is if it is feasible to build one because some scholars who wrote that it is possible to build a conscious machine end up saying its only possible in principle . There should be guidelines, policies in building these AI machines because unregulated expansion of AI machines can lead to massive unemployment and other harms to humanity. AI regulatory bodies should be formed and the government as well should step in to control the spread of AI. In work places, AI machines should lessen the burden of the workers and should not be allowed to become managers of the workers. Failure to do so, in the nearest future, AI machines will completely take over man's occupation, leaving a lot of people unemployed. In support of the view that AI machines should be built to add value and not to destroy human lives, Erik Brynjolfsson, director of MIT Initiative and Digital Economy says "AI and related technologies have already achieved superhuman performance in many areas and there is little doubt that their capabilities will improve, probably very significant, by 2030...he believes that it is more likely than not that we will use this power to make the world a better place." For instance, we can virtually get rid of global hunger, eradicate disease, and

provide better education to about every corner of the globe. AI and machine learning, on the other hand, can also be used to further concentrate wealth and power, leaving many people behind, and to produce even more terrifying weapons. Since neither outcome is inevitable, the correct question is not "what will happen?" but "what will humans choose to do?" we need to work hard to ensure that technology matches our values. At all levels, from government, to industry, to academia, and to individual choices, this can and must be done. (Janna Anderson *et al*., 2018).

For Sonia Katyal, therefore, how perceptions of AI and their application will influence the future trajectory of civil rights will be the most important set questions to ask in 2030. He further posits that questions about privacy, speech, the right of assembly and technological construction of personhood will all re-surface, putting into question, our deepest-held assumptions about equality and opportunity for all.

Conclusion

Deep insights and unanswerable issues have been discovered by the investigation of the complex interactions between consciousness and artificial intelligence (AI) in the context of redefining human-technology connections. grasp the consequences of AI awareness necessitates a sophisticated grasp of its numerous facets as AI technology develops and becomes more and more incorporated into our daily lives. Self-awareness, subjective experience, intentionality, and perceptual awareness are all aspects of consciousness, which is a fundamental aspect of human existence. While AI systems exhibit impressive cognitive abilities, it is

still difficult to make the transition from these abilities to the complex fabric of human consciousness. The line separating the objective world of consciousness and the physical operations of AI is highlighted by the illusive character of the "hard problem" of consciousness, as described by David Chalmers.

This discourse is supported by ethical considerations. The necessity for strong ethical frameworks becomes evident as AI technology gets closer to being sophisticated enough to mimic conscious behavior. Discussions about the intrinsic value of consciousness and the ethical treatment of potentially conscious creatures are sparked by questions about the rights, obligations, and moral agency of AI systems. The discourse's metaphysical undertones cut across academic boundaries. The ontological position of AI consciousness raises issues regarding the essence of reality. Philosophers such as Thomas Metzinger and Alva Noë underscore the embodied nature of consciousness, emphasizing the role of the body and environment in shaping our conscious experiences. The limits of cognitive extension have been pushed by the incorporation of AI into relationships between humans and technology. The extended mind theory, which raises the unsettling yet exciting possibility that artificial intelligence (AI), may be viewed as an enhancement of human consciousness, blurs the line between internal cognition and exterior instruments. There are significant repercussions for self-identity, autonomy, and our comprehension of personal limits. Responsible development and interdisciplinary cooperation serve as pillars in setting a direction for the future. The discussion is enriched by the ideas of philosophers like Daniel Dennett, who offers a naturalistic account of consciousness. The development of

AI technology must be guided by moral ideals based on knowledge of consciousness and metaphysical viewpoints.

The study of consciousness within the context of AI transcends disciplinary borders. The conversation spans philosophy, neurology, ethics, and metaphysics as technology pushes us closer to the frontier of potential AI consciousness. This voyage sheds light on the fundamentals of what it means to be a human in an age of technological wonders as well as the nature of awareness. Our understanding of consciousness serves as a compass to navigate the uncharted seas of a new technological frontier as we reimagine the ties between humans and technology.

Izuchukwu Kizito Okoli
Department of Philosophy
Nnamdi Azikiwe University, Awka
izk.okoli@unizik.edu.ng
&
Osita Gregory Nnajiofor
Department of Philosophy
Nnamdi Azikiwe University, Awka
og.nnajiofor@unizik.edu.ng

References

Anderson, Janna, Lee Rainie, and Alan Luchsinger. "Artificial Intelligence and the Future of Humans." Journal of Pew Research Center, vol. 5, 2018.

Block, Ned. "On a Confusion About a Function of Consciousness." *The Nature of Consciousness: Philosophical Debates*, MIT Press, 1995, p. 231.

Boltuc, Piotr. "The Engineering Thesis in Machine Consciousness." Journal of Research in Philosophy and Technology, vol. 2, 2012, pp. 187-201.

Chalmers, David J. "Facing Up to the Problem of Consciousness." *Journal of Consciousness Studies*, vol. 2, no. 3, 1995, pp. 200-219.

_____. "The Hard Problem of Consciousness." *Blackwell Publishing Ltd.*, 2007, p. 225.

_____. "The Conscious Mind: In Search of a Fundamental Theory". Oxford University Press, 1996, p. 313.

Davies, Paul. "Why the Hard Problem Will Never Be Solved." *ResearchGate*, https://www.researchgate.net, accessed 29 July 2023.

Dennett, Daniel C. "Consciousness in Human and Robot Minds." *Cognition, Computation, and Consciousness*, edited by Masao Ito,

Yasushi Miyashita, and Edmund T. Rolls, Oxford University Press, 1997, p. 17.
Gabriel, Markus. "Could a Robot Be Conscious? Some Lessons from Philosophy." *Journal of Epistemology, Modern and Contemporary Philosophy*, vol. 5, 2021, pp. 57-59.

Harvey, Inman. "Evolving Robot Consciousness: The Essay Problems and the Rest." *Journal of Cognition and Computing Sciences*, 2002, p. 13.

Harrison, M. "Scientists Say They Are Now Actively Trying to Build Conscious Robots." *Futurism*, https://futurism.com, accessed 23 July 2023.
Kaess, Gerald. "Could Consciousness Emerge from a Machine Language?" *Macalester Journal of Philosophy*, vol. 20, 2012, p. 144-160.

Koch, Christof. *The Quest for Consciousness: A Neurobiological Approach*. Roberts and Company Publishers, 2004.

Konstantin, R. "The Philosophical Issues of the Idea of Conscious Machines." *Journal of Philosophy*, vol. 6, 2017, p. 105.

Locke, John. *An Essay Concerning Human Understanding*. Oxford University Press, 1999, p. 98.

MacDorman, Karl F., and Stephen J. Cowley. "Long-Term Relationship as a Benchmark for Robot Personhood." *Journal of Personhood*, vol. 1, 2006.

Metzinger, Thomas. *Being No One: The Self-Model Theory of Subjectivity*. MIT Press, 2003.

Morisse, Beatrice. "Will Machines Ever Become Conscious?" *FirstRand*, https://www.firstrand.co.za, accessed 22 July 2023.
Noë, Alva. *Action in Perception*. MIT Press, 2004.

Noë, Alva. “Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness”. Hill and Wang, 2009.

Nath, R. "Machine Consciousness, Mind, and Consciousness." *Academia.edu*, https://www.academia.edu, accessed 21 July 2023.

Ng, Gary W., and Li W. Chi. "Strong Artificial Intelligence and Consciousness." *Journal of Artificial Intelligence and Consciousness*, vol. 7, 2020, p. 71.

Ng, Yew-Kwang. "Could Artificial Intelligence Have Consciousness? Some Perspectives from Neurology and Parapsychology." *Journal of AI and Consciousness*, vol. 1, 2021.

Oleksiewicz, I., and M. E. Civelek. "From Artificial Intelligence to Artificial Consciousness: Possible Legal Bases for the Human-Robot

Relationships in the Future." *International Journal of Advanced Research*, vol. 1, 2019, p. 260.

Rosenthal, David. "Concepts and Definitions of Consciousness." *Academic Press*, 2009, p. 157.

Searle, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences*, vol. 3, no. 3, 1980, pp. 417-457.

_____. "Minds, Body, and Programs." *Journal of Philosophy*, vol. 3, no. 3, 1980, p. 417.

Sid Ahmed, K. "John Searle - Chinese Room Argument." *Discourse*, https://discourse.clevious.com, accessed 20 July 2023.

Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G. E. M. Anscombe, Basil Blackwell, 1976, pp. 281-283.