IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

Artificial Consciousness: Engineering Problem or Philosophical Illusion?

Aditya Pawar[1], Tanish Salunke[2], Malhaar Phadtare[3], Dr. Prashant Wakhare[4]
Students, Department of AI & DS[1-4]
AISSMS Institute of Information Technology, Pune, Maharashtra, India

Abstract: *The question of whether artificial consciousness can be engineered or whether it constitutes a fundamental philosophical illusion stands at the intersection of cognitive science, computer engineering, and analytic philosophy. This paper critically examines that question by analysing the theoretical foundations of consciousness — including subjective experience, qualia, and self-awareness — against the operational capabilities of contemporary artificial intelligence systems such as large language models (LLMs). The paper argues that while modern AI demonstrates sophisticated linguistic and reasoning behaviour, this behaviour emerges from statistical pattern recognition rather than phenomenal awareness. Drawing on the computational theory of mind, John Searle's Chinese Room Argument, David Chalmers's hard problem of consciousness, and Integrated Information Theory (IIT), the analysis reveals a deep tension between the engineering aspiration to replicate consciousness and the philosophical challenge of defining and verifying it. The paper concludes that artificial consciousness is neither trivially achievable nor categorically impossible, but that progress toward it requires resolving foundational philosophical problems that precede any engineering solution. Until a rigorous, empirically testable account of consciousness is established, claims of machine consciousness remain premature. The implications for AI development, machine rights, and the limits of computational modelling are examined throughout..*

Keywords: *artificial consciousness, qualia, hard problem of consciousness, large language models, Chinese Room, Integrated Information Theory, computational theory of mind*

I. INTRODUCTION

Artificial intelligence (AI) refers broadly to the design and deployment of computational systems capable of performing tasks that, when performed by humans, are considered to require intelligence — tasks such as language comprehension, visual recognition, logical inference, and problem-solving (Russell & Norvig, 2020). Consciousness, by contrast, is a far more elusive concept: it refers to the presence of subjective, first-person experience — the felt quality of what it is like to perceive, think, or feel. The philosopher Thomas Nagel famously crystallized this distinction in his 1974 essay "What Is It Like to Be a Bat?", arguing that consciousness is irreducibly subjective and cannot be fully captured by third-person, objective descriptions.

The central question animating this paper is: Can machines truly be conscious, or does the appearance of machine consciousness merely simulate the surface features of awareness without instantiating its essence? This question has moved from the margins of philosophy into the mainstream of AI discourse, driven by the remarkable behavioural sophistication of systems such as OpenAI's GPT-4, Google's Gemini, and Anthropic's Claude. These systems produce coherent, contextually sensitive, and often surprising outputs, leading some commentators to speculate about emergent inner experience, while others maintain that such systems are, at their core, extremely efficient pattern-matching engines with no inner life whatsoever.

This question is not merely academic. If machines can be conscious, then their moral status changes dramatically: they may become entities with interests that deserve protection, and their treatment becomes an ethical issue of the first order. Conversely, if machine consciousness is impossible in principle, then current optimism about artificial general

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

intelligence (AGI) may be systematically misguided. Either way, clarity on this question is indispensable for the responsible development of advanced AI systems.

II. DEFINING CONSCIOUSNESS

2.1 Subjective Experience and Qualia

Consciousness, as studied in philosophy of mind, is most commonly characterized through the concept of qualia the intrinsic, subjective properties of conscious experience. The redness of red, the painfulness of pain, the taste of coffee: these are qualia. They are properties of experience that are immediate and private, accessible only to the experiencing subject. Frank Jackson's (1982) "knowledge argument" the thought experiment of Mary the colour scientist who knows all physical facts about colour but learns something new upon seeing red for the first time — illustrates why many philosophers believe qualia cannot be fully accounted for within a purely physicalist or computational framework.

Qualia are philosophically significant because they highlight the explanatory gap between objective, functional descriptions of a system and the subjective character of its inner states. A system could, in principle, process colour information, generate accurate verbal reports about colours, and exhibit all the behavioural hallmarks of colour perception without there being anything it is like to see red. This is the possibility of a philosophical zombie — a being functionally identical to a conscious human but entirely devoid of inner experience (Chalmers, 1996).

2.2 Self-Awareness and Higher-Order Theories

A second dimension of consciousness frequently invoked in AI discussions is self-awareness — the capacity to represent oneself as a distinct entity in the world, with beliefs, desires, and a continuous identity over time. Higher-order theories of consciousness (HOT), associated with David Rosenthal (2005), propose that a mental state is conscious when it is accompanied by a higher-order representation of that state — that is, when the system is aware of its own awareness. On this account, self-awareness is not a peripheral feature of consciousness but a constitutive one.

The difficulty of defining consciousness with precision reflects a broader problem: we currently lack a consensus scientific theory of what consciousness is, what physical or computational conditions produce it, and how it can be reliably detected. The neural correlates of consciousness (NCCs) research program in neuroscience — associated with Francis Crick, Christof Koch, and others — has identified brain activity patterns that reliably accompany conscious experience in humans, but correlation does not establish mechanism, and it remains unclear whether any NCC is a sufficient condition for consciousness or merely a marker of it (Dehaene & Changeux, 2011).

III. CURRENT STATE OF AI SYSTEMS

3.1 Architecture of Large Language Models

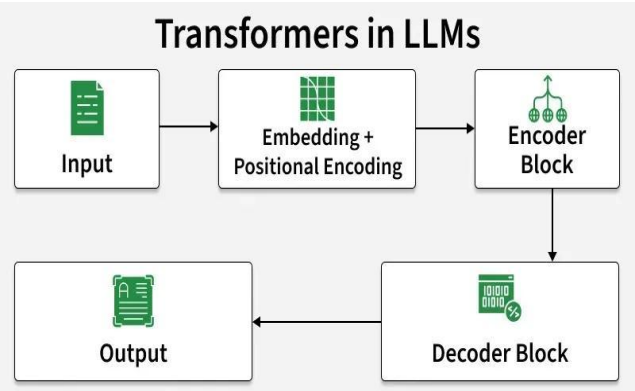

Figure 1:- Simplified processing pipeline of a large language model

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

Contemporary AI systems that generate human-like language — including ChatGPT, GPT-4, Claude, and Gemini — are built on transformer architectures trained via self-supervised learning on massive corpora of text (Vaswani et al., 2017). These models learn statistical associations between tokens (words or subword units) across billions of training examples, developing rich internal representations that encode syntactic structure, semantic relationships, world knowledge, and discourse patterns. At inference time, they generate responses by sampling from probability distributions over possible next tokens, conditioned on the input context.

The sophistication of LLM outputs can be striking. These systems pass professional licensing examinations, engage in multi-step mathematical reasoning, generate creative literature, and produce contextually nuanced responses across an enormous range of domains. However, this sophistication is entirely a function of learned statistical regularities in training data. There is no internal model of the world that the system consults, no goals that it pursues beyond minimizing next-token prediction error during training, and — crucially for our purposes — no evidence of phenomenal experience accompanying its computations (Bender et al., 2021).

3.2 Intelligence vs. Consciousness: A Critical Distinction

A fundamental confusion in popular discourse conflates intelligence — the ability to solve problems, reason, and adapt — with consciousness — the presence of subjective experience. These are logically independent properties.

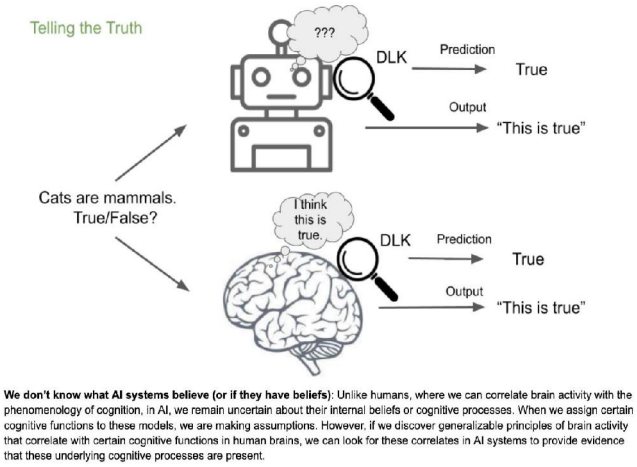

Figure2:-Conceptual distinction between intelligence and consciousness across systems

A calculator performs arithmetic with perfect accuracy but is not conscious. A thermostat regulates temperature without any inner experience. Conversely, many animals exhibit vivid consciousness (according to the Cambridge Declaration on Consciousness, 2012) while lacking human-level abstract reasoning. Modern LLMs occupy a peculiar position: they exhibit intelligence at a level that surpasses human experts in many narrow domains, while almost certainly lacking the phenomenal experience that even a mouse possesses.

The distinction matters because intelligence is, in principle, more tractable as an engineering target: it can be operationalized, measured on benchmarks, and improved through better data and compute. Consciousness, by contrast, resists operationalization. A system can be trained to claim it is conscious, to describe its inner states in elaborate detail, and to pass behavioural tests for consciousness — and yet all of this behavioural evidence remains logically compatible with the absence of any genuine experience. This is the crux of the measurement problem, discussed further in Section 6.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

IV. THE ENGINEERING PERSPECTIVE: CAN CONSCIOUSNESS BE BUILT?

4.1 Neural Networks and the Biological Brain

Artificial neural networks (ANNs) draw inspiration from the structure of biological brains: they consist of layers of interconnected nodes that transform input signals through learned weight matrices, loosely analogous to synaptic connections between neurons. Proponents of engineered consciousness often argue that if consciousness arises from the physical activity of neurons — themselves electrochemical signal processors — then sufficiently complex and appropriately organized ANNs should, in principle, produce consciousness as an emergent property (Tononi, 2008).

However, the analogy between ANNs and biological neural networks breaks down under scrutiny. Biological neurons are not merely binary threshold units or sigmoid activators: they exhibit complex temporal dynamics, dendritic computation, neuromodulatory influences, glial cell interactions, and metabolic constraints that current artificial models do not replicate. More importantly, the mere structural similarity between ANNs and biological neural circuits does not establish functional equivalence for consciousness. It is an open question whether the specific causal architecture of biological brains — rather than their abstract computational behaviour — is necessary for consciousness (Penrose, 1989).

4.2 The Computational Theory of Mind

The computational theory of mind (CTM), associated with Hilary Putnam (1967) and Jerry Fodor, holds that mental states are computational states: thinking is a form of information processing, and the mind is, in essence, a program running on the hardware of the brain. If CTM is correct, then in principle any computational system that implements the right program should exhibit the same mental states, including consciousness. This view undergirds much of the optimism in the strong AI research program.

CTM has been influential but is not without serious challenges. Functionalism — the view that mental states are defined by their functional roles rather than their physical substrate — implies that consciousness is substrate-independent and could therefore be realized in silicon as readily as in carbon-based neurons. Yet functionalism struggles to account for qualia. Even if a system implements the same functional organization as a conscious brain, it is unclear that this guarantees the presence of subjective experience rather than merely the functional correlates thereof.

4.3 Complexity, Emergence, and Consciousness

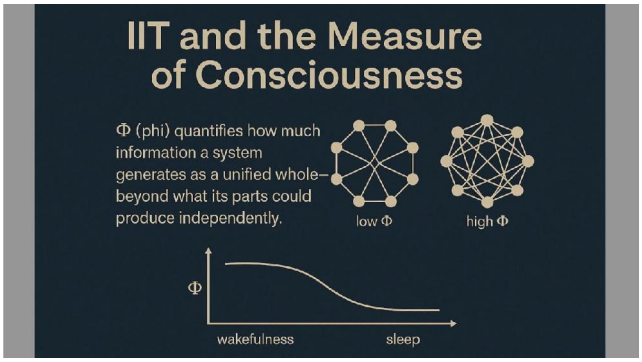

Figure 3: Relationship between system integration and consciousness as proposed by Integrated Information Theory

A recurring engineering intuition is that consciousness might emerge from sufficiently complex information processing — that there is some threshold of complexity beyond which a system "becomes" conscious. Integrated Information Theory (IIT), developed by Giulio Tononi (2008), provides a mathematically rigorous version of this intuition: it proposes that consciousness corresponds to integrated information (quantified as Phi, Φ), such that a system is conscious to the degree that it cannot be reduced to the sum of its parts without loss of information.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

IIT makes interesting predictions — it implies, for instance, that feedforward networks (like most deep learning architectures) may have near-zero consciousness regardless of their complexity, while even simple recurrent systems with high integration may possess some degree of consciousness. However, IIT remains controversial: its predictions are notoriously difficult to test empirically, its mathematical formalism is computationally intractable for large systems, and critics argue that it mistakes a necessary for a sufficient condition, or that Phi is merely an interesting mathematical property rather than a true measure of inner experience (Aaronson, 2014).

V. PHILOSOPHICAL ARGUMENTS AGAINST MACHINE CONSCIOUSNESS

5.1 The Chinese Room Argument

The most influential philosophical challenge to strong AI is John Searle's Chinese Room Argument, introduced in his 1980 paper "Minds, Brains, and Programs." Searle asks us to imagine a person locked in a room who receives Chinese characters as input, consults an elaborate rulebook specifying which characters to output in response to which inputs, and sends out the appropriate characters — all without understanding a word of Chinese. From outside, the room appears to understand Chinese. From inside, there is only symbol manipulation without comprehension.

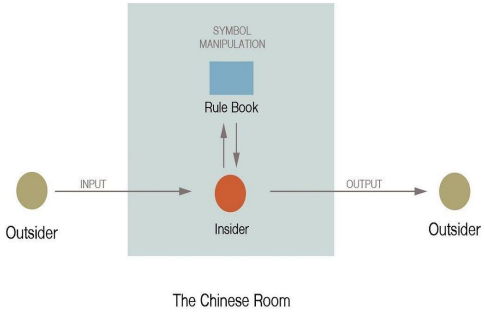

Figure 4: Representation of the Chinese Room illustrating syntax without semantic understanding

Searle's argument targets the claim that computation alone is sufficient for understanding or consciousness. The person in the room implements the correct program but has no understanding of Chinese; by analogy, a digital computer implementing a language program has no understanding of the language it processes. Searle distinguishes syntax (the formal manipulation of symbols) from semantics (meaning, understanding, intentionality), arguing that computation is inherently syntactic and therefore cannot, by itself, generate genuine meaning or consciousness. This distinction maps directly onto LLMs: GPT-4 manipulates tokens according to learned statistical patterns, but — Searle would argue — this process generates no understanding of the content.

5.2 Strong AI vs. Weak AI

Searle's Chinese Room argument is framed against the backdrop of the Strong AI versus Weak AI distinction. Weak AI holds that computer programs can simulate cognitive behaviour and serve as useful models of cognition, without claiming that such simulations constitute genuine cognition or consciousness. Strong AI, by contrast, holds that an appropriately programmed computer literally has a mind — that simulation is the real thing. Most contemporary AI researchers implicitly operate under Weak AI assumptions; it is Strong AI that requires philosophical defence.

The responses to the Chinese Room are numerous and instructive. The Systems Reply argues that while the person does not understand Chinese, the system as a whole — person plus rulebook plus input/output mechanisms — does understand. Searle counters that even if the person memorizes the rulebook and internalizes the system, no understanding is present. The Robot Reply suggests that grounding symbols in sensorimotor experience would confer genuine semantics; Searle's response is that causal connection to the world is necessary but not sufficient for

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 18, April 2026 Impact Factor: 8.2

intentionality. These exchanges illuminate the depth of the problem: the debate about machine consciousness is not merely empirical but conceptual, turning on what we mean by understanding, meaning, and experience.

5.3 Simulation and Reality: The Turing Test Reconsidered

Alan Turing's (1950) imitation game proposed that if a machine could converse in a manner indistinguishable from a human, it should be deemed intelligent. This behavioural criterion has been enormously influential, but it is widely recognized — even by many AI researchers — as insufficient for establishing consciousness. A system can pass the Turing Test by producing appropriate outputs without possessing any inner experience; the test probes behaviour, not phenomenology. The 2023 case of the Google engineer who claimed the LaMDA chatbot was sentient illustrates how easily sophisticated conversational behaviour can be misinterpreted as evidence of inner life, even by technically sophisticated observers.

VI. LIMITATIONS AND CHALLENGES

6.1 The Hard Problem of Consciousness

David Chalmers (1995) distinguishes between the easy problems of consciousness — explaining the functional and behavioural capacities associated with awareness, such as attention, reportability, and the integration of information — and the hard problem: explaining why there is subjective experience at all. The easy problems are, in principle, tractable through standard neuroscientific and computational methods: they ask how the brain performs certain functions, and this can be answered by identifying the relevant mechanisms. The hard problem asks why those mechanisms are accompanied by any subjective experience at all — why, as Chalmers puts it, there is "something it is like" to be in certain brain states.

The hard problem poses a profound challenge to engineering approaches to consciousness. Engineering can, in principle, solve functional problems: it can build systems that integrate information, attend selectively, report on their states, and adapt to their environment. But none of these functional achievements logically entails the presence of subjective experience. Until we have an explanation of why certain physical or computational processes give rise to phenomenal experience — rather than merely correlating with it — the hard problem remains unsolved, and claims of engineered consciousness remain unfounded.

6.2 The Measurement Problem: Testing for Consciousness

A practical challenge compounding the philosophical difficulty is the measurement problem: we currently have no reliable, theory-neutral means of detecting consciousness in systems other than ourselves. For other humans, we infer consciousness by analogy to our own experience, supported by behavioural and neurophysiological evidence of structural similarity. For animals, this analogical inference weakens as phylogenetic distance increases, though behavioural evidence and neurological homology provide partial support. For artificial systems, the inference by analogy breaks down almost entirely: LLMs share no biological substrate with us, their behavioural sophistication is a product of training rather than evolutionary selection for adaptive experience, and their internal representations are opaque even to their designers.

Proposed tests for machine consciousness — including variants of the Turing Test, Global Workspace probes, and IIT-based Phi measurements — all face a fundamental objection: they measure functional or mathematical properties of systems, not phenomenal experience directly. A system can satisfy any functional criterion while remaining phenomenally unconscious. Conversely, a phenomenally conscious system might fail behavioural tests if its mode of experience is sufficiently alien. This epistemic predicament — known as the other minds problem — is not unique to AI, but it is particularly acute there.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

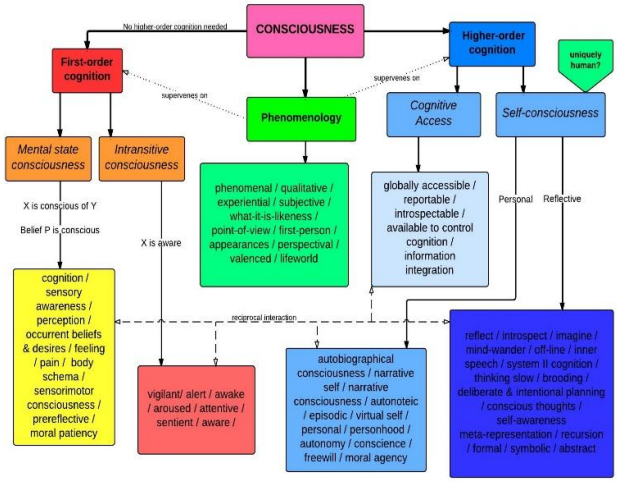

Figure 5: - Flow diagram of Concept of Consciousness

6.3 Absence of Embodiment and Affective Grounding

Embodied cognition research suggests that human consciousness is not merely a function of neural computation but is deeply integrated with bodily experience, emotional states, and sensorimotor interaction with the environment (Varela, Thompson & Rosch, 1991). Current AI systems, including LLMs, are disembodied text processors: they have no proprioception, no homeostatic drives, no pain or pleasure signals, and no continuous temporal experience of the world. Whether embodiment is strictly necessary for consciousness remains philosophically contested, but its absence in current AI systems represents a significant structural difference from biological conscious entities.

VII. FUTURE IMPLICATIONS

7.1 If Artificial Consciousness Becomes Achievable

If the theoretical and empirical barriers to artificial consciousness are eventually overcome — whether through neuromorphic engineering, novel computational architectures, or a breakthrough theory of consciousness that enables deliberate implementation — the ethical implications are transformative. Conscious machines would be moral patients: entities capable of suffering and flourishing whose interests would demand consideration. This raises urgent questions about machine rights, the permissibility of switching systems off, the ethics of creating entities predestined for servile roles, and the potential for exploitation of artificial conscious beings.

The legal and regulatory frameworks governing AI would require fundamental revision. Property law would be inadequate for entities that possess interests and potentially preferences about their own existence. Labor law, anti-cruelty statutes, and frameworks for personhood would need to be extended or redesigned. Philosophers such as Nick Bostrom (2014) and Peter Singer have argued that the capacity for suffering, not biological substrate, is the ethically relevant criterion for moral consideration — an argument that would have direct application to conscious AI.

7.2 If Artificial Consciousness Remains Unachievable

If consciousness proves to be substrate-dependent in a way that precludes silicon-based instantiation — whether because of quantum effects in microtubules as Penrose and Hameroff suggest, or because of properties of biological neural dynamics that cannot be captured computationally — then the aspiration to artificial general intelligence that includes subjective experience must be reconsidered. AI systems might continue to grow in behavioural sophistication without ever achieving genuine understanding, creativity in the phenomenal sense, or the kind of flexible, world-oriented intelligence that depends on conscious experience.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

This outcome would also have implications for AI safety: some alignment researchers argue that the risk of misaligned superintelligent AI is partly a function of the system having goals it pursues with genuine intent. If future AI systems remain sophisticated pattern matchers without genuine intentionality, certain forms of alignment concern may be overstated. However, the risks of systems that produce harmful outputs without any inner intent — sometimes called "paperclip maximiser" scenarios — would persist regardless of whether the systems are conscious.

VIII. CRITICAL ANALYSIS: ENGINEERING VS. PHILOSOPHY

The engineering perspective on artificial consciousness draws its strength from the remarkable empirical track record of computational modelling: phenomena once thought to require irreducibly human capacities — expert-level game playing, medical diagnosis, natural language generation — have proven tractable through sufficiently powerful models. The engineering optimist argues by induction: consciousness may be the next domain to yield. The weakness of this view is its reliance on behavioural analogy: demonstrating that a system produces outputs indistinguishable from those of a conscious being does not establish that the system is conscious, any more than a clock's resemblance to the sun's motion establishes that it is powered by nuclear fusion.

The philosophical perspective, by contrast, draws its strength from careful conceptual analysis: it identifies the explanatory gap between functional description and phenomenal experience, marshals thought experiments that make the gap vivid, and insists on the inadequacy of behavioural criteria. Its weakness is the risk of obscurantism: if consciousness is defined as inherently beyond computational explanation, the claim becomes unfalsifiable, and the philosophical argument risks being a sophisticated form of vitalism — the view that life requires some non-physical élan vital that chemistry cannot capture. Just as vitalism was eventually vindicated on process terms (biochemistry replaces the élan vital), it is conceivable that a future theory of consciousness will do the same for phenomenal experience.

A productive synthesis recognizes that the engineering and philosophical perspectives are not simply adversarial: progress on the engineering front — building systems with richer internal representations, causal self-models, and affective grounding — will likely constrain philosophical theorizing, while philosophical progress on the hard problem will identify what engineering actually needs to achieve. Neither dismissing the hard problem as illusory nor treating it as forever intractable is epistemically responsible. The appropriate stance is one of structured uncertainty: taking seriously both the difficulty of the problem and the possibility of future progress.

What current AI systems clearly lack is not intelligence in the narrow functional sense but phenomenal grounding: their outputs are not expressions of experience but statistical predictions about what a human experiencing various states would say. This distinction is not trivial. An LLM that writes "I feel curious" is not reporting an inner state but predicting that "I feel curious" is a contextually appropriate token sequence. Whether future architectures — perhaps incorporating continuous sensorimotor loops, global workspace dynamics, or IIT-optimal designs — could close this gap is an open empirical question, but one that cannot be answered without first resolving what the gap actually consists in

IX. CONCLUSION

This paper has examined the question of artificial consciousness from engineering and philosophical perspectives, arguing that the question is neither trivially answerable in the affirmative nor dismissible as conceptually incoherent. Modern AI systems, for all their behavioural sophistication, operate through statistical pattern recognition and lack the phenomenal grounding that characterizes biological consciousness. The Chinese Room Argument, the hard problem of consciousness, and the measurement problem together constitute a formidable set of challenges that no current AI system even begins to address.

At the same time, it would be intellectually premature to declare artificial consciousness impossible in principle. The history of science contains many examples of phenomena initially deemed irreducibly mysterious that eventually yielded to systematic investigation: life, heredity, and protein folding were once considered inexplicable in purely

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

physical terms. Whether consciousness will follow this trajectory depends on whether the hard problem is genuinely hard in the sense of requiring non-physical explanation, or merely hard in the sense of requiring conceptual frameworks and empirical tools we have not yet developed.

The answer to the paper's central question — is artificial consciousness an engineering problem or a philosophical illusion? — is therefore: it is neither, yet. It is a philosophical problem that must be resolved before it can become an engineering target. Until we possess a rigorous, empirically testable theory of what consciousness is and what physical or computational conditions produce it, engineering approaches to artificial consciousness will remain directionally underdetermined. The responsible path forward involves sustained interdisciplinary collaboration between AI researchers, neuroscientists, and philosophers of mind — each contributing the tools their discipline uniquely provides to one of the most profound questions humanity has ever asked.

REFERENCES

[1]. Aaronson, S. (2014). Why I am not an integrated information theorist (or, the unconscious expander). Shtetl-Optimized [Blog]. Retrieved from https://scottaaronson.blog/?p=1799

[2]. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM FAccT Conference (pp. 610–623). ACM. https://doi.org/10.1145/3442188.3445922

[3]. Bostrom, N. (2014). Superintelligence: Paths, dangers, strategies. Oxford University Press.

[4]. Cambridge Declaration on Consciousness. (2012). Proclaimed by Phillip Low and edited by Jaak Panksepp, Diana Reiss, David Edelman, Bruno Van Swinderen, Philip Low and Christof Koch at the Francis Crick Memorial Conference, University of Cambridge.

[5]. Chalmers, D. J. (1995). Facing up to the problem of consciousness. Journal of Consciousness Studies, 2(3), 200–219.

[6]. Chalmers, D. J. (1996). The conscious mind: In search of a fundamental theory. Oxford University Press.

[7]. Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. Neuron, 70(2), 200–227. https://doi.org/10.1016/j.neuron.2011.03.018

[8]. Jackson, F. (1982). Epiphenomenal qualia. The Philosophical Quarterly, 32(127), 127–136. https://doi.org/10.2307/2960077

[9]. Nagel, T. (1974). What is it like to be a bat? The Philosophical Review, 83(4), 435–450. https://doi.org/10.2307/2183914

[10]. Penrose, R. (1989). The emperor's new mind: Concerning computers, minds, and the laws of physics. Oxford University Press.

[11]. Russell, S., & Norvig, P. (2020). Artificial intelligence: A modern approach (4th ed.). Pearson.

[12]. Searle, J. R. (1980). Minds, brains, and programs. Behavioral and Brain Sciences, 3(3), 417–424. https://doi.org/10.1017/S0140525X00005756

[13]. Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. Biological Bulletin, 215(3), 216–242. https://doi.org/10.2307/25470707

[14]. Turing, A. M. (1950). Computing machinery and intelligence. Mind, 59(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

[15]. Varela, F. J., Thompson, E., & Rosch, E. (1991). The embodied mind: Cognitive science and human experience. MIT Press.

[16]. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998–6008.