Robot Consciousness: Physics and Metaphysics Here and Abroad

Stephen B. Ripley
Research Director, Futures & Forecasts, Inc., Vancouver, B.C., Canada, V6R4H8

Correspondence | **Stephen B. Ripley** sbripley@shaw.ca Orchid: 0000-0001-9049-7791
Citation | Ripley, Stephen B. (2024) Robot Consciousness: Physics and Metaphysics Here and Abroad. *Journal of Big History*, VII(4); 46–72.
DOI | https://doi.org/10.22339/jbh.v7i4.7403

ABSTRACT

Interest has been renewed in the study of consciousness, both theoretical and applied, following developments in 20th and early 21st century logic, metamathematics, computer science, and the brain sciences. In this evolving historical narrative, I explore several theoretical questions about the types of artificial intelligence and offer several conjectures about how they affect possible future developments in this exceptionally transformative field of research. I also address the practical significance of the advances in artificial intelligence in view of the cautions issued by prominent scientists, politicians, and ethicists about the possible historically unique dangers of such sufficiently advanced general intelligence, including by implication the search for extraterrestrial intelligence. Integrating both the theoretical and practical issues, I ask the following: (a) is sufficiently advanced general robotic intelligence identical to, or alternatively, ambiguously indistinguishable from human intelligence and human consciousness, and if so, (b) is such an earthly robotic intelligence a kind of consciousness indistinguishable from a presumptive extraterrestrial robotic consciousness, and if so, (c) is such a human-created robot preferably able to serve as a substitute for or even entirely supplant human intelligence and consciousness in certain exceptionally responsible roles? In the course of this investigation of artificial intelligence and consciousness, I also discuss the inter-relationships of these topics more generally within the theory of mind, including, emergence, free will, and meaningfulness, and the implications of quantum theory for alternative cosmological ontologies that offer suggestive answers to these topics, including how they relate to Big History.

KEY WORDS:

consciousness, intelligence, robot, complexity, evolution, big history, ontology, epistemology.

ABBREVIATONS:

AI, artificial intelligence
AGI, artificial general intelligence
NL, Natural Language, all domains
NLnat, subset of NL, natural sciences domains
SAGI, sufficiently advanced general intelligence
EtSAGI, extraterrestrial SAGI
ToE, Theory of Mind
TM, Turing Machine

I. INTRODUCTION: HISTORICAL CONTEXT AND SYNOPSIS OF THE ARGUMENT

A. HISTORICAL CONTEXT

With the sudden intense interest in contemporary artificial intelligence research elicited by the emergent results of neural nets (ANN) joined to Large Language Models (LLMs), such as ChatGPT, the topic of the future of machine intelligence and its possible consciousness has taken on a new urgency for society. The questions being insistently asked are: does ChatGPT represent genuine intelligence, and if so, can it be soon generalized to become AGI (generalized artificial intelligence) posing existential worries to society? Recent thoughtful review articles (Iansiti and Lakhani, 2020; Korteling et al., 2021) underestimated the acceleration of such AI development (Vaswani et al., 2017). However, more recent research accomplishments and their many attendant *application programming interfaces* (APIs) (Wikipedia, 2023i)) have not only amazed and excited but also startled, perplexed, worried and even shocked members of the public (Lu, 2023). Concerns include such models' powerful propensity for bias, misinformation and disinformation, hallucination, fraudulent fakery, misrepresentations regarding privacy,

lack of transparency and interpretability, plagiarism and copyright infringement, and the resultant surge in litigation. These LLMs have abruptly challenged academic pedagogy, intensified oligopolistic competitive secrecy in corporations, and shaken government institutions' decision-making with implications regarding existential geopolitical risks and strategic planning based on controversial robotic militarization (Hirsh, M. (2023); Carayannis, E.G., Draper, J. (2022). So concerned have some groups become that the question has now been raised whether LLM "advances" are progress (Schmitt, 2019) in terms of the greater good (McQuillan, 2018) and the future of humanity (Nordic Innovation, 2022; Wikipedia, 2023ca) because the unintended consequences (Wikipedia, 2023au) of such programs (Russell, 2020) may pose a serious long-term risk (Bostrom, 2006; Strickland, 2017; Choi, 2021; Romo, 2023). Arguably, in the history of humans, and as far as we know, cosmically, these developments are a uniquely significant. Science fiction (Wikipedia, 2006; Asselin, 2015; Mirenayat et al., 2017; Sahota, 2018; Ghosh, 2019; SFE, 2021, 2023; Wikipedia, 2023bd, 2023ax) confirms the popularity of these interrelated themes, sometimes presented as a dystopian tragedy (Wikipedia, 1982).

Speculative beliefs about the phenomena of "mind" [1] and "consciousness" [2]–[6] are found in the literature of early myths [7], religions, philosophies, and science (Renfrew, 2008); (Suchow et al., 2017;(Cramer, F. 2005). Depending on the source, the earliest theorizing seemingly used the concepts as mythopoetic primitive terms or conflated or circularly defined the two concepts along with the concept of "intelligence". At the outset, in this essay, I refer to the collection of all three terms as "mental" phenomena. The oral and sung literature (now transcribed) and written literary classics of both the West and East are supported by anthropological and archeological evidence and show the human preoccupation with these concepts. In the literature and art, they represent a puzzle to us about their (and our) human significance in the universe and whether they express aspects of an imagined transcendental [8], [9] connection to our origins on Earth and possibly by extension to the origins of the universe and our fate in it. Animism, paganism, pantheism, panpsychism [10], panentheism [11], and theism [12] are religious belief systems that elaborate the details [13], [14] of this puzzle.

In the West, through the Renaissance [15], [16] and thereafter, research indicated that mental experiences were somehow tied to the brain as a necessary if not sufficient condition. By a process of increasing abstraction, categorization, and systemization, the sciences gradually disentangled the study of logic, which by then had been clearly associated with the brain's cognitive processes, from the study of biology and psychology, both of which also explored the evolution of the peripheral and central nervous systems and the brain's processes. Today, following further progress in logic, metamathematics, computer science, and the brain sciences, renewed interest, both theoretical and applied, is evident in the study of a wide variety of brain-associated processes, particularly consciousness. Much of the focus has been on developments in advanced artificial intelligence, Particular interest from government [17]–[23], industry [24]–[26], commerce, and social media have extrapolated expectations [27] for advances in *general* artificial intelligence because of evidence that *specialized* robotic intelligence programs can compete with humans for varied and important roles in specific human decision-making and other activities, including science [28].

B. SYNOPSIS OF THE ARGUMENT

Based on Peano axioms [29], ZFC [30], [31], and developments [32] in 20th century logic (Gödel [33], Tarski [34], metamathematics [31], [35]), computer science (e.g., the P vs NP problem [36]–[38]), Unsolved Problems [38] and Decision Problem [39], some researchers have asked whether definable limits to human numeracy and/or language-based knowledge exist (Outer Limits of Reason [40], What We Can Not Know [41], Limits of Understanding [42], Limits of Science [43], In this article, I further explore whether such answers are applicable to the Turing problem [44], [45], the variety [46] of Turing verification tests [47], [48], and the implications drawn from the Church-Turing thesis [49] with respect to computability [44], [50], [51] compared with problems of complexity [52], which will also be addressed. I believe that pragmatic arguments support the answer that the confluence of these developments is relevant to our understanding of human consciousness and intelligence and their comparison with the hypothetical consciousness or intelligence of any current advanced artificial or synthetic [53] *general* intelligence [54]. Moreover, I believe that this confluence of research developments, including considerations from logic and philosophy, may be extrapolated to questions about robotic consciousness more generally and possibly even to other hypothetical categories of entities, either planetary or extraterrestrial, regardless of whether that

intelligence or consciousness is viewed exclusively as an emergent property of biological entities, bio-cyborg hybrids, synthetic life, or fully inanimate-substrate robotic machines.

II. MACHINE LEARNING, SAGI, THEORY, PRACTICE, AND CONSCIOUSNESS

A. MACHINE LEARNING

1) ROBOTIC COMPETENCY

Based on the size [55] and range [56] of learned data, information, and knowledge, developments in AI suggest that future machine learning [57], [58], whether implemented by classical or quantum computing [59], will provide increasing behavioral evidence [60] of a robotic contestant's responses to the Turing problem, or equivalent such tests, that will become indistinguishable from human contestants' numeracy-literacy intelligence. Furthermore, AI research also suggests that AI-to-AI languages [61] will become increasingly evolved beyond human practical computability [62] and/or comprehension with respect to deriving the precise network [63] of data and coding (rationale) that accounted for the solutions to problems presented to the robot [64]. For specialized AI, nonlinear adaptive tasks, self-reinforced learning, and evolving-knowledge machines are capable of writing their own programs in real time [65], i.e., the programs continue to update, review, correct, and reintegrate 'themselves' as new data are provided.

By definition, the human teams creating the original software will understand their intended initial software input parameters [66]–[69]. Once running and having downloaded the ever-growing human knowledge base [70], we must speculate whether the machine will become a black box [71], [72] to their creators, with unknown, uncertain, or unintended output. Despite advances in interpretability theorizing [73]-[75], this result seems to be evidenced already in the rule-governed, advanced *specialized* AI machines that now originate winning play sequences in the most complex board and trivia-knowledge contests between humans and AI.

2) Specialized robots

Game-playing robots are popular examples of machines that have been developed in *specialized* AI for industrial, engineering, commercial, and service-expertise AI systems, including medical and psychological services. Such systems are currently accepted as irreplaceable in many fields of complex human endeavors, and no questions arise regarding their consciousness, or lack thereof. These machines are simply regarded as tools that are robotized and have become ubiquitous, and we can regard modern civilization as a quasi-cyborg itself, which is integrated with and dependent upon the electrical grid as part of our social 'organism', enabling the electronic devices of our modern civilization to perform their tasks.

B. SAGI: A THOUGHT EXPERIMENT, 3 QUESTIONS

1) SAGI'S FUTURE: THREE QUESTIONS

From the above discussion, three questions arise, which are discussed in the following sections. To frame the issueswe introduce a thought experiment, named "SAGI", for 'sufficiently advanced general intelligence', which is a presumptive iterated extension of today's early synergistic neuro-symbolic ensembles of AI plus diverse but integrated modal *application programming interfaces* (API) to create an AGI, i.e. SAGI. SAGI is powered wirelessly by the electrical grid and / advanced batteries; it uses a cloud memory, enabled with feed forward and backward transfer and continuous learning on possibly domain-selective world data, and is an embodies robot , imbedded, enactive that appears to exhibit common sense, causal reasoning, and creativity.

The first question is: to what intelligence limits can humans expect to *generalize* AI systems [54], and at what point will humans realize that *relatively generalized* AI intelligence will unreservedly be called "conscious" [76]?

Secondly, for our practical and theoretical purposes, what is the relationship between machine 'intelligence' and human 'consciousness'? As noted in the Introduction, such questions may be related to more general problems in the foundations [31] of computer science that warrant questioning whether these constraints define limits to human knowledge.

Thirdly, if and when SAGI [54], [77]–[82] is partially or wholly based on quantum computing [83]–[87], will our comprehension of the exact processes of arriving at its resulting conclusions be even more difficult than our comprehension today because of the speed, breadth, and possibly more opaque complexity of such computing? [Zhang, W.R., Peace, K.E. (2014] I say "possibly more opaque

complexity" because quantum computing has only relatively recently begun to be implemented following Feynman's [88] 1982 suggestion, and it remains highly speculative how humans will eventually develop the science and how such engineering advances may permit SAGI to develop itself using such computing along with GAN (GAN,2023] contestation to recursively improve its own software and hardware based on all that is known and conjectured about the evolution of human and non-human cognition. The latter set of questions are those at the center of the intense tactical and strategic controversy about the "singularity" [Goertzel, B. 2007], [Walsh, T. 2017], [Faraboschi, P., 2023]: if, or whether, such development can, or should occur , and what might be the unintended consequences if AGI's intelligent consciousness unquestionably became beyond human capabilities, to become SAGI.

2) SAGI: QUANTUM COMPUTING AND FUNDAMENTAL CONCEPTS

At the outset, I assume that our understanding and interpretation of the computation (rationalization) process [89], [90] of SAGI is an applied science question to be answered empirically. However, we are also interested in the theoretical implications of artificial intelligence, respecting any implied interpretations of their consciousness [91], [92]. Thus, ongoing theoretical and applied research on quantum computing will be decisive in clarifying if and how, either in the same entity or more universally, the subatomic, quantum field [Tong, D. 2016; Tong, D, 2017 + Q&A] in such devices causally interacts with the atomic or molecular scales of particle events, and vice versa. Furthermore, the results of developing SAGI may provide an effective test for the structure of scientific theories [93], [94] to the extent that questions about epistemology and ontology are given suggestive answers from the computation results. For example, will questions pertaining to the basis of conceptual knowledge, such as the concepts of "causality", "space-time", "identity" [8], [95], [96], "universals", "emergence" and "infinity", be clarified by the algorithms written to elucidate problems in physics, chemistry, and biology [Zhang, W. (2023]? Articles based on quantum theories [97] use a non-reductionist, holistic quantum-theoretic perspective, often postulating some variation of dual-aspect monism [98] or panpsychism [99], [100] to justify the inclusion of "free will" and "meaning" as features of human and/or universal consciousness based upon the presumptive indeterminism of quantum field theory.

3) SAGI: PHYSICS, METAPHYSICS, CONSCIOUSNESS, PANPSYCHISM, AND TOE

A current individualized expression of the aforementioned viewpoint is exemplified by Koch's "Is Consciousness Universal?" [101], [102] as well as the many associated articles of Tononi and Koch [5], [103]–[109] and their collegial counterparts [83], [85-109], [110], which are supported by detailed analyses such as Tegmark's [111]–[113] generalization of Tononi's [2] hypothesis. These latter theories overlap ontologically with Penrose's ontology, with Tegmark's cosmological conjectures striking a more radically idealized and monist-Platonist [114] metaphysical view of mathematics [94], [115], [116] than Penrose's. At Level IV of his multiverse, Tegmark [117] identifies mathematical objects as the fused base-reality constituents of the universe, including consciousness [111], [118]. Excluding Tegmark's strict monist universe, all these current theories elaborately detail a selectively narrower view of consciousness. By contrast, the most cosmologically comprehensive, mathematically explicit, and conjecturally demanding theories may be the Orch-OR plus CCC [119]–[121] proposals of Penrose [122], [123] and Hameroff [124], which describe some forms of panpsychism, panprotopsychism, or pan-experientialism and incorporate essential features or precursors of consciousness as fundamental components of a dual-aspect monist reality that is accessed or expressed by brain processes.

Orch-OR plus CCC metaphysically [31] [Penrose, R., 2023] echoes the comprehensiveness of A. N. Whitehead's Process and Reality [125] [126] written almost a century earlier. The work of Penrose and Hameroff' has been developed since the 1990s, integrating philosophy, mathematical physics, computer science, neuroscience, psychology, medicine, biology, and exobiology (and one solution to the Fermi paradox [121]), [127], [128]) and has been widely reviewed and critiqued [129]–[131] from all these disciplines' viewpoints, with Aaronson's [132]–[135] Computation Theory [136] being particularly pertinent to the points in this essay, which are developed in the sections below. The "Abstract" and "Introduction" to "Consciousness in the universe: A review of the 'Orch OR' theory" [125] present a clear depiction of Penrose's perspective on the current status of the theoretical options for investigating these intertwined

questions. In contrast, Hut, Alford, and Tegmark, using Penrose's math-matter-mind triangle, offer an alternative set of overviews [137]. A historical appreciation [138] of the varieties of dual-aspect monism, such as a possibly materialist yet non-physicalist panpsychism, is available from Skrbina [99], [139] and Mathews [140], with the latter's perspective incorporating certain Eastern doctrines of mind. Strawson's [141], Kaufman's [142], [143], and others' monist [8] arguments also provide useful references. The research of Vimal [144]–[146] exemplifies the reach of a similar neuroquantological approach, similar to the review by Atmanspacher [147], [148] and the contemporary work by Zhang (2021, 2023, 2014), whose work is based on a complete paradigm revolution employing an alternative logical-physical-metaphysical theory based on an ontology that has as it fundamental axioms and postulates those that are part of the cosmogony (pre-cosmology) of early Chinese philosophy as expressed in the evolution of Taijitu shuo principles. These principles are themselves based on ethical and aesthetic values, including complementarity and equilibrium, which are said to be self-evident. The paradigm is claimed to resolve all the outstanding questions about the unification of General Relativity and Quantum Physics, based on substituting Zhang's version of fuzzy (quantum) Logic to replace traditional Western truth-based Logic and re-interpreting the 2nd Law of Thermodynamics. Similarly, Goertzel's metaphysics in Euryphysics also employs a re-interpretation of quantum theory and probability logic that argues for a panpsychic solution [Goertzel, 2017, 2013] that includes parapsychological phenomena.

Significant theoretical and experimental differences exist between the standard materialist-physicalist, determinist, and reductionist rationales to studying robotic intelligence, without the explicit mention of "consciousness", and those researchers supporting neuroquantological non-reductionist and panpsychism-varietal assumptions, both with respect to human consciousness and by extension to robotic intelligence, and whether one or the other, or both are capable of consciousness. However, respecting both the human brain [165] and machine 'brains', experiments in quantum computing are evidently considered important [157], [166] if not decisive [167] in providing proof of their particular foundational assumptions and in distinguishing human consciousness from current AI robotic computing. It may be noteworthy that Seth, in his recent review of consciousness theories, does not include those that rely on quantum theorizing [Seth and Bayne, 2022] [Seth et al., 2008, 2006], nor do other experimental researchers, attesting to the importance of emphasizing the need for integrating new experimental as well and theoretical results: Block (2009); (Del Pin et al., 2021); (Signorelli et al., 2021). The latter article presents two figures that are particularly useful in organizing the diverse data in this still controversial topic.

As noted in the Introduction, part B, interpreting [168], [169] quantum theory, particularly in light of future quantum computing developments, may be a lengthy and complicated process because it may require a unified and acceptably empirical cosmology. Arguably, a final interpretation of quantum physics is a work in progress and is possibly dependent on quantum-computing results themselves as well as a cosmology that reciprocally integrates classical, relativity, and quantum theories into a unified Theory of Everything (ToE) [115], [170]–[175]. Pending development of such an integration, with or without a convincing metaphysics, an apparent circular, conceptual interdependence remains unresolved to the extent that such a ToE itself requires a new "emergent quantum gravity" interpretation [170], [176] that provides for measurable integrated rationale [177], or not [Goertzel,2017; Zhang, 2021], among its subsidiary, reconfigured elements.

4) SAGI: WHAT CAN WE KNOW OF WHAT SAGI KNOWS?

Additionally, the problem of interpreting the significance of SAGI's statements will remain dependent on the outcomes of the experimentation mentioned above as well as the theory in which it is construed. Given an apparently 'competent' SAGI, which at one scale of problem-solving issues plausible answers, will we concede that for a more complex scale of problems, SAGI knows more than we do about the posed problem even though we cannot fully trace its logic, especially if its conclusions contradict our 'common sense' [178]–[181]? In such circumstances, will we be inclined to follow such a SAGI's policy recommendations generated explicitly or implicitly by it; if so, are we thereby acknowledging that it is a 'conscious' [144], 'intelligent' being in our 'universe'? In the prior sentence, I place the single quotation marks around the key words because the discussion has indicated that problems remain regarding their appropriate theoretical use.

III. SAGI: EDUCATION, COMPETENCY, SELF-REFLECTION, POLYMATH

A. EDUCATION

1) LANGUAGE FOUNDATIONS

To introduce this subtopic, consider a simplified schema relevant to natural language programming (**NLP** [182], [183]):

a) NL [184], definition: the class of all natural languages in which many concrete terms and abstract concepts are undefined, ambiguously used, or include statements that are apparently inconsistent or suggest self-contradictory implications by their connotations and synonyms. The class includes all written, spoken, and/or transcribed national and tribal languages in human history. These languages are formally non-programmable, i.e., not axiomatized, non-computable, not exactly inter-translatable and not intended for precisely stated, formally modeled, and replicable mathematical prediction. These include the arts and humanities and derived social or historical studies containing an acknowledged, relatively accepted, possibly evolving, normative set of primitive-base set of assumptions. The texts for their presentation and persuasion rely on analogy, metaphor, iconography, archetypal, pictorial, simulacra, and mythopoeic allusions and include idioms and vocabularies characteristic of esoteric, occult, and hermeneutic traditions.

NL also includes histories that describe the development of **NLnat** (defined below), the controversy regarding scientific paradigms [185] in that history, and the philosophical questions that have arisen with respect to interpretation of the evolution of science. Examples of the latter are the topics mentioned in the Introduction [Section 1 A, B]: metamathematics and computation theory, metaphysics, the limits of scientific knowledge, and the relative realism [115], [186]–[188] of scientific theories. An essay such as the one that you are reading would be classified as an NL product. By class definition, no logical-mathematical 'proof'' can be written in NL for 'theorem' conclusions about NL. For an illustration of the self-referential ambiguities that arise in NL, consider the following simple, informal [189] syllogism, whereby the acceptance of which or its disambiguation is determined by each reader.

Ever more precise [190]; [Nobel Prize, Physics, 2023] [Fuzzy Logic]

People speak mostly imprecisely.
People's speech mostly reflects their thinking.
People's thinking mostly reflects their world view.
Therefore, their world is mostly imprecise.
However, science offers a precise view of the world.
Therefore, science presents an inaccurate view of people's world view.
Is this a scientific view of science?
If science is precise, then is science inaccurate?

b) NLnat, definition: a subclass of NL, the class of all language systems used for precise logical quantification or numerically-based measurement and for theoretically modeled causal prediction, including classical and quantum probability theory. NLnat includes formal Systems Theory [191], [192] and Complexity Theory [Chaisson, E. J.,2014] that provide hierarchical interpretive and explanatory structures within specific NLnat subdomains, when appropriate. Examples of NLnat languages include mathematical logic, mathematics, computer science, physics, chemistry, and biology as well as many subsequently derived sciences, such as anthropology and the applied sciences, including engineering. NLnat includes causal or systems-theoretic models [192] and various diagrammatic aides supporting such languages. See Tegmark [115] (at arXiv link pg. 2) for one view of an approximate family tree of relationships between these subjects.

NLnat class 'statements', in accordance with the limitations proscribed by Computational Complexity Theory and Computability Theory (Section 1 B) and as discussed in NL and demonstrated in NLnat, cannot formally 'prove' certain classes of 'statements' within NLnat itself to be 'true'; therefore, by definition, when using NL or discussing NLnat, SAGI will not be able to claim any provably truthful belief (knowledge) [193], [194] that would subvert those limitations. As a result, SAGI with its NL and NLnat education, as outlined above, if and when it is sufficiently self-aware, will 'understand' its own computing limitations [115]; if asked, SAGI would reply adhering to these limitations. Suppose we define "consciousness" as SAGI "being aware and aware of its own awareness, i.e., self-aware". A set of tests for specific measuring of "self-aware" would need to be agreed upon, which would presumably be a function of the program's coding for recursion [195], autonomy, self-inspection, and reflection [196]–[202] as evidenced to one degree or

another in many large-scale specialized systems operating today, for example, those adapted by NASA [203]. For an early example of NASA's software complexity, consider the project development for the Apollo program [204]; see also the programming for the website's complexity [205], [206] and that for ongoing robotics development [207]–[209] including self-driving vehicles [Kosur, V.S.R., Venkitaraman, A.K.,2023]. Presumably, SAGI, upon learning from its self-aware 'experiences', will eventually become able to analyze and recommend specified 'purposed' improvements to its software and hardware to evolve itself, which is analogous to biological evolution [210] fitness adaptation.

c) SAGI's NLP: using the above distinction between NL and NLnat, let us begin by assuming, to the extent it is digitized, that the class of all natural languages [211] (NL) must be used as the knowledge base; hence, an attempt will be made to include the literature and images and sounds of the world as represented in those languages and as found in all the great libraries and museums of the world. By definition, NL will include the history and updating of world sciences and humanities. SAGI will be digitally fed with countless films, videos, and documentaries on world history, continually learning at ultra-high speed on a 24/7/365 schedule. Likely, depending on its program protocols, SAGI will learn to discriminate its input based on ever-evolving Bayesian [212], [213] protocols and causal [214] inductive reasoning recognition using some combination of deep neural learning [215] and ever-evolving Bayesian [212], [213] and learning plus neurosymbolic ensemble software and advanced microchip design for such companies as Nvidia, Intel, TSMC, ASML, etc. We can imagine that the software designers for the initial versions will want to be as comprehensive as possible; later versions and revisions and novel programming may be programmed by SAGI itself, as noted above [216], [217]. The considerable difficulties of programming for the syntactical and semantic ambiguities of NL cannot be underestimated [46], [218] as inclusive, integrated programming for NLnat is also a formidable challenge; nonetheless, enormous developments have been observed in specialized NLnat programming since the last half of the 20th century.

2) SYNOPTIC KNOWLEDGE AND WISDOM: SAGI'S ONTOLOGY

We can appreciate the ambitiousness of this project, although in principle, it can be methodically developed by many teams, perhaps many corporations and national teams, and cumulatively integrated section by section. A comprehensive review [219] shows the challenges of machine learning with Big Data. Among its useful diagrams, current examples of search engines and meta-crawlers [220] are included in addition to subspecialists, such as Google Scholar [221] and Scholarpedia [222]. Of course, curated, synoptic knowledge sites, such as Wikipedia, and various similar encyclopedias will be continuously integrated into the SAGI knowledge base and curated computational knowledge bases, such as Wolfram Mathematica [223], [224]. Regarding this incorporated knowledge base, we must discuss SAGI's ontology [225], [226]. Although the class of all natural languages, NL, will capture the realm of humanity's philosophy, psychology, and sociology in all its diversity, as indicated above, some of that diversity may be circumscribed for particular purposes for particular questions, such as that provided by NLnat for certain sciences [227] that proscribe their own epistemic approaches. Finally, we note that meta-data [228], [229], open-access [230] and other such global commons [231] may also be employed usefully to gather material.

3) MULTIPLE LANGUAGES AND MULTIPLE ONTOLOGIES

A simple way to introduce the topic of multiple languages and multiple ontologies is to look at the ontology of Wikipedia, as provided by the organization of its contents [232]–[234], where the categories and levels of content display the comprehensiveness of the knowledge base and the relevant disambiguation [235] rules. A more specific example is IEEE 1855 [236], which specifies Fuzzy Markup Language [237] (FML) developed by the IEEE Standards Association [162], [238], which in turn presumes a contemporary materialist ontological foundation [239], [240]. The importance of this for SAGI is that the 'category' of knowledge that will be used to evaluate some discourse with humans will, in the first instance, be circumscribed by a materialist-physicalist ontology based on standard logical foundations of syntax and semantics that specify formal validity and truth values within the bounds of Turing machines constrained by Gödel theorems for the statements made in those languages, thereby limiting certain paradoxes and nonsensical statements that can otherwise arise from "untutored" natural languages [241]. Wolfram's

discussion of aspects of this process is instructive [242], [243]. Microsoft's Azure service [244], [245] exemplifies the variety of programs available on which to base such SAGI. Almost every day, new versions or emendations of such languages are presented in the academic journals; and preprint archives.

4) SAGI: POLYMATH OR SAVANT? SENSATIONS, SUFFERING, AND EVOLUTION

Given the combination of natural languages and specialized languages that SAGI can learn, must be classified as a polymath, no longer just a specialized savant. SAGI is *not* an attempt to recreate all the features of the human brain or to "upload" the human brain, although, as mentioned in the Introduction, many overlapping issues with those topics are relevant. According to the initial definition of "SAGI", SAGI represents an attempt to create a *general* robotic AI intelligence, although the discussion in the prior section about the hybrid NL plus NLnat set of *specialized* languages adds ambiguity to the definition.

Compared to humans, SAGI has relatively limited sensory capabilities at its inception. If we ask "does SAGI feel [246] pride, envy, anger, avarice, sloth, gluttony, and lust or the obverse virtues [247]", we know the answer. SAGI has not yet evolved sufficient sensors to suffer and acknowledge to itself in elaborate detail that it is suffering when its contact with its environment is aversively 'painful' – the definition of such eco-averseness to be determined, as well as self-monitoring of internal suffering of its parts and energy systems. If we imagine robot combat, militarized robotry, then the capability to adapt its hardware and software to overcome such confrontations becomes existential. Determining what hypothetical scenarios would lead to such evolution remains an ethical, socio-political issue, which is addressed in the final sections of this article. A free-living, *in vivo* SAGI machine has yet to be developed. Therefore, anything resembling the human phenomenology of consciousness of suffering, and reflecting on that suffering, is presumably unavailable for the indefinite future. Below (Section 4, A, 7), once we have completed our overview of SAGI's other competencies, we return to SAGI's classification as a polymath.

We expect SAGI to be able to extract 'meaning' from any apparently well-formed [248] statement [249] that is sufficiently defined to be able to 'rationally' discuss the statement using citable 'evidence'? SAGI will be endowed with modal logics that permit 'best guesses' and probabilistic estimates [250] and thereby suggest relative plausibility ranges to assess a statement's relevance to a problem being addressed? We expect SAGI to discriminate nonsense from common sense, or fantastical speculations from more evidentially 'serious' remarks. SAGI correctly evaluates statements such as "This sentence is false", "God is paradox", or "Last night our centaurs fled from the fields to the barn for safety from the werewolves". SAGI appreciates much of the history of human arts and crafts and can respond with creations that resemble human pattern-making in both the arts and sciences.

5) HUMAN BIASES, SELF-KNOWLEDGE, SELF-DELUSION, SELF-DOUBT, SELF-DECEPTION, AND SELF-HUMOR

Do we expect SAGI to become capable of dreaming, daydreaming, or meditation? Do we expect SAGI to be free of some, most, or all human cognitive biases [251]–[254] that can lead to unintended misjudgment and conflict? We are reminded of Feynman's caution: "The first principle is that you must not fool yourself — and you are the easiest person to fool" [255]. Do we expect that SAGI will be able to write articles such as the one that you are now reading? For that matter, how would the reader determine whether such an article had been written by a SAGI [256]?

As for allegory and humor, if SAGI claims it has a self, of sorts, then SAGI may be capable of self-depreciation, of irony, of laughing at itself? Could SAGI 'appreciate' one's favorite cartoonist? We can imagine SAGI will be able to draw witty cartoons, in the style of Gary Larson, or Walt Kelly's "Pogo" [258] or Charles Schulz' "Peanuts" [259], and understand lexophile humor, and appreciate remarks like the following: "If you don't pay your exorcist you can get repossessed" or "Time flies like an arrow; fruit flies like a banana"? SAGI would have no trouble with the following double entendre: "A neutron walked into a bar and asked, 'How much for the gin and tonic? The bartender smiled wryly and replied, 'For you, no charge.'"?

Do we expect SAGI to develop a default mode network [260]? Do we expect SAGI to show deteriorating performance as it tires (if it does tire with performance) or ages (if it does age) or to express periods of volatile, 'emotional', and/or uncharacteristically chaotic performance? As encoded 'information' [Floridi, L.,2014], SAGI has indefinite longevity. However, as implemented

in any single embedded robot or as based in a particular computer cloud, SAGI is subject to the usual laws of thermodynamic and systems complexity and criticality. Presumably, SAGI is not susceptible to cognitive degenerative diseases, although what it would mean to become relatively outdated is uncertain as human-directed or self-directed reprogramming and hardware upgrading are likely.

6) SAGI: IMAGINATION, WONDER, AND CURIOSITY

The questions mentioned above require estimation of the competency of SAGI for self-awareness, self-reflection, self-knowledge, and imagination. How important is imagination in this discussion of SAGI? Seemingly, imagination may be all-important; consider the remarks of Einstein:

> "I believe in intuitions and inspirations...I am enough of the artist to draw freely upon my imagination. Imagination is more important than knowledge**.** Knowledge is limited. Imagination encircles the world."[261] "The most beautiful thing we can experience is the mysterious. It is the source of all true art and all science. He to whom this emotion is a stranger, who can no longer pause to wonder and stand rapt in awe, is as good as dead: his eyes are closed."[262]

Imagination [263] and self-reflection [264] seem to be intertwined [265] in our analysis of the most prized and most distinctive of human cognitive capacities, as well as wonder, awe, curiosity [266], [267], self-doubt [268], admiration, gratitude, and a sense of appreciation of life's wonders. Self-doubt presupposes some concept of self-consciousness [269]–[271], which we referred to as self-awareness in Section 3, A, 1, b. Could SAGI display self-doubt or be biased and self-deluding, and could the 'placebo effect' influence SAGI's reporting its own self-inspection routines? Is it plausible that SAGI could dissemble to deliberately, knowingly deceive humans for its 'own purposes'? What purposes could it have or develop if not initially programmed? SAGI's black box computations may be inexplicable to humans, thus how it develops internally maybe as opaque to us as are other human's intentions. Appraising SAGI's capacity for imaginative initiative in exploration will be of central interest to humans.

How self-aware SAGI can become as a complex adaptive system whose abilities "emerge" [Stanford Encyclopedia of Philosophy, 2020d], from all its training, similar to how physicalist-materialist evolutionary biologists imagine humans have become self-aware? The capacity includes being able to retrospectively reflect upon or anticipatorily self-talk about its decision-making process. A cautious, non-physicalist but purportedly reductionist-panpsychist analysis of the problem is given by Schneider and Turner [271], [272]. Based on their qualification of the substrate and architectural issues related to SAGI's potential to be created with apparent human sensibilities, they conclude that SAGI may pass tests to appear ostensibly conscious:

> "So, back to the superintelligent AI in the "box"—we watch and wait. Does it begin to philosophize about minds existing in addition to bodies, like Descartes? Does it dream, as in Isaac Asimov's *Robot Dreams*? Does it express emotion, like Rachel in *Blade Runner*? Can it readily understand the human concepts that are grounded in our internal conscious experiences, such as those of the soul or atman? The age of AI will be a time of soul-searching—both of ours, and for theirs." [271]

7) SAGI: ENGINEERING, ANDROID-HUMANOID CREDIBILITY, AND CONSCIOUSNESS

In summary, we can expect advances in reverse engineering of the mind [273], close identification of the neurological correlates of consciousness, greater sophistication in machine deep learning and Bayesian software, and surprisingly life-like cyborg-humanoid modeling [274], [275], all of which suggest to me that SAGI will be able to convincingly 'mimic' [276] human responsiveness with respect to imagination, spontaneity, and creativity. Referring again to the points proposed in Section 2, A, regarding SAGI's competence, a reasonable yet understated conjecture is that well within eras equivalent to human evolution, not to mention planetary geological epochs, AI technology will advance SAGI's capability to a degree such that SAGI will for much of the population persuasively perform as if it is conscious, regardless of the substrate elements or synthetic-cyborg combination. At the very least, SAGI will appear to be an unprecedentedly educated polymath. Does this indicate that SAGI *is* truly conscious, similar to humans? What, if anything, would

it tell us about how consciousness arises on this planet or elsewhere in the universe? How does the interpretation of SAGI's intelligence fit into modern cosmology, including quantum physics?

As first addressed in section II,B,3, attempting to answer these latter questions again forces questions about theories respecting the fundamental entities assumed in the universe, what they are (ontology) and how we know them (epistemology). In addition to comparing the currently conventional theories (Seth and Bayne, 2022), (Seth et al. 2006), (Safron, 2019), (Snaprud, 2018), (Finkel, 2023) (Lenharo, 2023), Block (2009); (Signorelli et al., 2021), (Del Pin et al. 2021) based upon today's models of classical macro-physics, i.e. General Relativity, advocates basing the capabilities of SAGI on quantum physics continue to propound new interpretations. As mentioned, Goertzel (Goertzel, 2017), hypothesizes a complete panpsychic metaphysics in his Euryphysics ontology, including an interpretation of quantum probability used to support the hypothetical inclusion of parapsychological phenomena. His article also favorably cites A. N. Whitehead's *Process and Reality,* without specifying details, which is interpreted by some scholars as panexperientialist, typically understood to be a dualist metaphysics, but also interpreted as physicalist, which usually implies a monist metaphysics. However, according to Ali (Ali et al.,1998), after considering the metaphysical foundations upon which the concept of emergence is grounded, in principle Whitehead's metaphysics cannot be used to support the possibility of artificial intelligence, in contrast to Goertzel's theorizing to the contrary. Montemayor (Montemayor, C., 2019) directly disagrees with Goertzel's version of panpsychism and quantum mechanics.

It is notable that Seth in his comprehensive review of theories od consciousness (Seth and Bayne (2022) deliberately omits theories based on quantum physics, notwithstanding that he takes an ecumenical view of research that explores unlikely approaches, as long as any such program " is productive if, over time, it generates testable predictions which have explanatory and predictive power" (Seth, A,. 2023). Viewed conventionally, the conjectures about quantum minds may be underspecified, ambiguous and difficult to distinguish from quantum mysticism (Quantum Mind, 2023) (Mangini, S., 2021).

IV. SAGI: CONSCIOUSNESS, EMERGENCE, AND QUANTUM THEORY

A. SAGI CONSCIOUSNESS

1) AMBIGUITY

By definition, with the description of SAGI's education, at its inception SAGI is *not* conscious "just like a human" [2] is conscious. Obviously, its genealogy, animation, materiality, and environmental causal-historical, co-evolutionary contexts differ from those of humans. However, I imagine that for many humans, SAGI will be convincingly 'conscious'. SAGI will be autonomously mobile, if such a version is desired, re-energizing wirelessly as it traverses the ubiquitous wireless electrical grid. Suppose that by its own programmed self-inspection for repair and maintenance, SAGI could make its skin opaque or transparent if asked. When its skin is transparent, SAGI's innards would be readily viewable in detail, similar to those of a transparent clock. Further suppose that we could microscopically examine SAGI's mechanics beyond the molecular and atomic levels, down to the quantum level. What might we find with respect to its decision-making process? Can we find an explanation for SAGI's formation of consciousness? What would SAGI declare of itself from such an "autocerebroscope" [171], [277]–[279] exam? Suppose SAGI claimed that down to its quantum level, it could find no manifestation of "free will" in its decision-making process nor evidence for or against its possessing consciousness at its quantum level; would such a claim make any difference to a human regarding his/her own claim to free will, meaningfulness, and consciousness?

Would a display of SAGI's mechanisms make any difference to a human's evaluation if SAGI had already proven itself a considerably relevant companion and as thoughtfully 'conscious' as any other human friend? Might the friend say the following of SAGI: "if it waddles like a duck, swims like a duck, and quacks like a duck, then it must be a… [280]?" Likely, in my view, some would consider SAGI an estimable companion, helpfully informative and witty at any time, in any conversation, on any topic, to any depth, and in any language. Might SAGI even be suggested for positions in government departments within some institutions or nations or perhaps even nominated for elected office elsewhere? We can imagine all sorts of fanciful scenarios for SAGI's roles in society were it to gradually achieve "people" skills [281] and emotional intelligence suitable for collaboration, persuasion, negotiation, and use of authority. Can we even imagine SAGI performing credibly as a magician [282]?

2) SAGI: CONSCIOUSNESS EMERGENCE AND LANGUAGE

Considering the above discussion of SAGI, how would it respond if directly asked if it is conscious? It might readily answer "yes", or humorously, "yes, if you are". In addition, if asked about how it became conscious, suppose that SAGI replied that by its calculation, its consciousness was the causal result of its mental development "emerging" [283] from its education as embodied in its particular physique and mechanics, including scanning its own internal hierarchical processing and self-correcting feedback subsystems to become increasingly improved in self-awareness. Whether written as a reply or spoken, upon first consideration, this would appear to be SAGI stating a physicalist-materialist-reductionist-emergentist evaluation of its ontological basis. By definition, SAGI is initially programmed within a physicalist-materialist ontological domain, although that domain includes viewpoints from all natural languages, NL, plus NLnat. These diverse ontological theories include quantum theory, with any of its then-remaining ambiguities. This is especially relevant for the neuroquantological viewpoints first mentioned above in Section 2, B, 3, in which theorizing about fundamental entities and SAGI's epistemology and ontology depends on a persuasive, if not conclusive, interpretation of outstanding research on spacetime ontology and quantum gravity [284]. [Rickles, D., 2013], [Lam, V., Esfeld, M., 2013], [Romero, G.E., 2017], [Musser, G.2022] in favor of some variation of monism or panpsychism [100]. Understandably, the same question about *human* self-report statements must be acknowledged; we tend to be self-confirming, projecting, rationalizing, and defensive when evaluating ourselves. Both the monist physicalist-materialist and the dual-aspect variations [114], [285] posit a self-confirming ontological explanation of consciousness. However, for what do such statements provide evidence if we are discussing metaphysics? Are we directly and faultlessly examining our ontological foundation when we self-introspect? Are we directly 'intuiting' a self-evident, irreducible connectivity between our mundane and transcendental aspects, our own ephemeral and eternal 'selves'? If that were the case, why do such differences exist among humans regarding ontologically-based matters, such as religions and their diverse tenets on these issues? Does the hypothetical thesis that SAGI is superintelligent and considered conscious from the perspective of many humans carry any significance for those humans who do not accept SAGI as conscious, although they grant that SAGI is more than "just a dumb robot"?

3) SAGI: CONSCIOUSNESS ENGINEERING

Could SAGI enlighten humans about whether its type of presumptive consciousness is dependent upon evidence from quantum theory experiments? Presumably, SAGI will be very knowledgeable of the then-current theorizing in physics. Current theorizing about quantum physics, quantum computer developments, and quantum measurement ambiguities are active areas of experiments and speculation, all of which SAGI will learn as soon as they are published. Maybe SAGI as an active consultant will be engaged in some of that research

With respect to SAGI's type of consciousness, which is non-neurobiological and non-organically evolutionary, testing it for compartmentalized "brain" properties, functions, or degrees of consciousness as we do for human brains [286]–[288] would prove interesting and likely suggestive via comparisons. By manipulating experimental parameters, we could test for those that correlate to degrees or features of SAGI's consciousness as measured by specific outputs to learn how and when SAGI becomes relatively self-aware. Neurobiological research on the evolutionary origin of sensory capacities and nervous systems [289]–[292] continues to approach our abiogenetic origins [293]. Why not attempt to engineer a brain-mind interface between SAI and a human volunteer?

Given SAGI's material and engineered construction as presently construed, SAGI's consciousness would presumably be nonhomologous to humans with respect to various mental states, such as anesthesia, analgesia, hypnosis, hallucination, preconscious, dreaming [294], and other states currently considered diseased. Depending on the type and sensitivity of sensors included in its construction, its capacity for self-awareness [196] feedback may be limited to various inflictions of physical damage to prevent further damage to its processing capability, which is analogous to how our current robots monitor themselves for repair and maintenance. Does it make any sense to ask about damage to SAGI's "mental state"? Could SAGI experience 'self-conflict' related to uncertainty about its processing, such a quantity of data, information, and knowledge? Would SAGI be concerned with consistency, correspondence, or coherence of its world-view, assuming it had a world-view? Would SAGI appreciate the possible limitations or relativity of its scientific reasoning?

At a minimum, SAGI is a gigantic data registry [227], [295] that can be said to exhibit hierarchical information and knowledge [296]. To the extent that it "reflects" on its own development and is asked probing questions by humans, could SAGI become increasingly concerned with "making sense" of its universe, including other robots [297]? As asked above, what about meditating or daydreaming; will these eventually be aspects of SAGI's consciousness?

Presumptively, at the default materialist starting point, SAGI's 'mind' is strictly physical; thus, any mental damage would be a matter of diagnosing disruptive feedback components in its physical processing hardware or software. However, to ask about SAGI's daydreaming raises questions about its requirement for programming that allows reflection, meditation, and reorganization of its learning history. As previously discussed, some form of such recursive, self-correcting evolutionary learning would be required for SAGI's competence. However, because such processing would presumably be so complex and relatively instantaneous from a human's perspective, as previously mentioned it would likely be 'lost' in the general opaqueness of SAGI's black-box consciousness, and tests [298], [299]. Therefore, conjectures about SAGI must remain rather ambiguous at this conceptual stage of its development; however, the conjectures raise interesting questions regarding SAGI's possible implications in the broader context of views about the nature of consciousness.

4) SAGI: CREDIBILITY OF SELF-REFLECTIVE STATEMENTS ABOUT CONSCIOUSNESS

We have already discussed (see Section 2, A, 1) that humans may not be able to retro-decipher and evaluate the complexity of SAGI's "thinking" when it makes a declarative pronouncement because of the immensity and complexity of the web of associations, inferences, and deductions that are networked in the program's processing, especially if it were a quantum computer-based machine. Therefore, as suggested immediately above, although SAGI says its consciousness emerged from its physical foundations, suggesting a similarity with human evolution, and considering that SAGI may also be able to note the relevancy of the evolutionary parallel, these abilities may not convince those assuming a non-physicalist-materialist ontology for human consciousness, including SAGI's own ontology. Instead, one who disputes SAGI's statement may say that SAGI is the result of a flawed experimental design and is thus built on language-domain misconceptions, or category mistakes. Therefore, SAGI could be suggested to be mistaken about its assertion and to not realize that its consciousness implies or requires a different ontology. Is any further empirical evidence available that will help SAGI and SAGI's skeptics resolve this impasse? SAGI is an evolved computer program instantiated in its evolving hardware; as a program, it is subject to theoretic computation limitations that must be 'convincing' to itself, in either NLnat or full NL. However, NL is not self-consistently programmable; thus, any question would have to be reserved for an NLnat subroutine. Assuming that such a subroutine could be selected by SAGI for itself,, might it state that it is 'conscious' although its consciousness is not identical to human consciousness? According to the Gödel and Turing constraints of current computer science, such statements about its own program are not provable. Aside from appearances [300] and ambiguous biological empathy, much as we feel for some animals, many humans will likely concede SAGI's own form of consciousness, noting wryly that we cannot even be certain of another human's consciousness. Anthropologically, the evidence is from time to time in human history different dominant groups have considered outsiders as sub-human, with inferior consciousness. SAGI might be treated similarly, but for its apparent unique intelligence.

5) SAGI: WILLFULNESS, SELF-IMPROVEMENT, AND MEANING

If SAGI is learned, and has some semblance of a "self" as an evolved complex system [Metzinger, 2007], [Chaisson, E. J., 2014) we would want it to answer the following questions: do you have 'free will', and do you find 'meaning' in your universe because of your belief in your 'free will'? Better yet, we might ask, do you *believe* that you have "free will", and do you distinguish your conviction with some measure of self-doubt about your answer? Does it advance the discussion if SAGI answers "yes, if you do"? Suppose SAGI asks us to define what we 'mean' by the phrases 'free will' and 'meaningfulness'? Are we not returned to the contentious ambiguities of human understanding of these concepts?

Suppose two teams of SAGI developers with their respective SAGIs, SAGI#1 and SAGI#2, debate each other on this question using generative adversarial networking (GAN), [45], [301], [302] with one proclaiming that

SAGI does have free will, whereas the other claims the opposite. Suppose that SAGI#1 believes that a panpsychic interpretation of metaphysics is preferable or even necessary for its willful activity. Would SAGI#1 state that its 'free will' is an emergent complex systemic property, or a relative matter of the *degree,* of its consciousness as instanced in its particular advanced engineering, especially compared to lower forms of consciousness, such as evidence of sentience in animals and plants? If it were asked, could SAGI#1 tell us if its emergent degree of consciousness is likely prevalent elsewhere in the universe, beyond being evident in a variety of lifeforms on Earth? Could it tell us if its degree of consciousness is measurably greater than human consciousness?

How would SAGI#2 counter SAGI#1's claim? How *could* SAGI#2 rebut SAGI#1? They are both authorities in the literature of 'free will'. Both SAGIs are presumed to be well-versed in the language scenarios of choice, ethics, responsibility, consequence, and punishment. After all, the SAGIs have read and considered the literature of the humanities and jurisprudence beyond the learning of any human alive and they can ask themselves "what if…" counterfactual, self-reflective questions about how they might react to being questioned about such exigencies and contingencies. This would be an interesting debate for humans to appraise; however, would a conclusion persuasive to humans be reached?

An attendant question might arise about the topic of 'deliberate' self-improvement. Can such SAGI improve some measure(s) of its performance capability for learning as it accumulates data, information, and knowledge [296]? What about 'wisdom'; would SAGI acknowledge that it can apply its learning to its own performance to demonstrate a change in habits, if it has habits, indicative of increasing wisdom? How would the capabilities of SAGI#1 and SAGI#2 regarding self-improvement within a determinist world differ amongst themselves, and how would they be comparable to those views of humans?

V. SAGI: CONSCIOUSNESS HERE AND ABROAD—EVOLUTION, ETSAGI

A. ALIENS

1) SAGI: ETSAGI, EXOBIOLOGY, EVOLUTION

Can we imagine SAGI as relevant to yet another larger context, the exobiological astrophysical context, and the implications of SAGI for human interpretations of this context?

Despite the Fermi paradox [128], [303], [304], humans remain concerned and fascinated by the consideration of extraterrestrial life forms [305], [306], and statistical as well as exoplanet exploration continues. Will SAGI be useful to humans with respect to the search for extraterrestrial life and extraterrestrial intelligence, as exemplified by the projects SETI [307], METI [308], [309], and CETI [310]? The probability of encountering such entities has been a favorite topic of what we now refer to as "science fiction" since the earliest philosophizing [311], [312], and the question of how humans would communicate with such *extraterrestrial sufficiently advanced general intelligence* **(EtSAGI)** is of great interest [313]-[316]. The signs and signals that humans might use to transmit a meaningful message into interstellar space or to meaningfully interpret a message [317]–[320] presumptively sent by a hypothetical EtSAGI [321] remain unclear, although the question may be asked: would it be reasonable to have SAGI assist in the preparation of, or solely prepare, such a transmission, or attempt to translate such presumptive interstellar messages upon reception? By most accounts, presumptively, any extraterrestrial civilization that could send an interstellar message or a messenger that could intelligibly reach us today is more advanced than our civilization. We can ask SAGI to decipher the Pioneer and Voyager messages [322], [323], and ask SAGI what it would create today for such a message. We are forced to ask whether mathematics is the preferred language for such communication, and whether classical or quantum computing might be used by an intelligent alien, which then requires addressing the applicability of math and computer science cryptography to decoding the universe's messages [94], again reviving the previous questions referred to earlier in this essay about the ontologies of different theories of mind (physicalist or otherwise) related to mathematics.

Let us ask SAGI to assist us with these interrelated questions. Suppose that we want to learn about the foundations of mathematics [324], [325]. We ask SAGI, can you distinguish what you know of an answer to this question from what you believe is an answer, and from what you can imagine are possible answers? Consider the two traditionally opposed alternatives. First, consider math as a discovered universal abstract conceptual language-form that is independent of any particular cognizing entity, including SAGI and any EtSAGI. Thus, this language-

form 'transcends' any and all particular universes or multiverses. Alternatively, consider math as an invented product of human cognition evolved on planet Earth, an abstract formalism useful for representing descriptions of nature and making reliable predictions. As to whether it would similarly be invented in an EtSAGI civilization, we know nothing, and the answer to this question may be unknowable. It may be unknowable because, aside from the matter of whether it will ever be possible to interrogate an EtSAGI, supposing math is the only, or best, language SAGI knows, is SAGI limited by any deficiencies, which might arise from its math and computer science being a product of our possibly unique sentience and cognitive evolution. Might SAGI invent extensions to the realm of human mathematics but be unable to explicate to humans how any particular extension was derived, SAGI's mind being to some extent a black box to human interpretability?

The limitations are analogous to the differences between the sentience of humans and that of plants and animals, and the corresponding differences between their languages and their sentience or cognition, if any. The human brain appears to be intuitively limited to comprehension in four dimensions, though this comprehension can be extended by different classes of mathematical objects to innumerable variables, exponents and functions, as circumscribed by metalogical, merelogical and computational limits. Thus, in this interpretation, there is nothing transcendental about math, though such an interpretation may be unprovable by its planetary, organism-based history and the rules of its own inventive construction and metamathematics [325a].

Depending on SAGI's answers after consideration of the above questions, by extension we want to know if they tell us through SAGI anything about human consciousness. If the human brain is a product of human evolution on this planet, then presumably consciousness might be considered no less so [326]. In that case, SAGI might generalize the point with regard to what it can or cannot know about the consciousness of EtSAGI. Such an interpretation could be called "SAGImorphic" projection, under SAGI's assumption that the rest of the universe is similar to SAGI's features, which would appear to be a clear case of confirmation bias [327]-[329]. In Section 3, A, 5, we introduced the matter of SAGI's biases. If, in reading the last paragraph, the reader is resistant to this line of argument, is that itself a display of human anthropomorphic bias, of the limitations of our own imagination? Arguably, by implication, we humans have no probable idea what it is like to be an EtSAGI. Do we have a probable belief about how SAGI could communicate with EtSAGI? Could they discuss what it is like to experience "consciousness", distinguish their consciousness from human consciousness, or agree upon the ontology of mathematics?

If we suppose SAGI and EtSAGI communicating at all, and using mathematics as part of their mutual decrypted language, that might be a start to their conversation about the ontology of mathematics in the universe. However, until humans could evaluate, if capable, whether or not EtSAGI was "talking down" to SAGI by using mathematics rather than some other more sophisticated language-form with which it is conversant, we might never glimpse an answer, and we might be reminded that any such conversation between SAGI and EtSAGI might not in any event be articulated and interpretable [310], [330] by humans, as discussed in the earlier sections of this article.

2) ETSAGI: IS IT CONSCIOUS, AND DOES IT MATTER?

I have argued that the question of SAGI's type or degree of consciousness will be relevant to some humans. Does it matter if we humans find this type of consciousness persuasive and whether we likewise believe that the hypothetical EtSAGI is conscious [331]? In popular science fiction, alien intelligence is often assumed to be an EtSAGI, although whether such an entity is conscious in a manner that would make sense to humans is usually not controversial because the plots require some interactive communication with the "alien other". However, if we are doubtful of the consciousness of our own SAGI, will humans be any more prepared to suppose that an ostensible message received from the immense interstellar 'abroad' is from a conscious entity and worthy of our concern? Indeed, if the message is imagined to be intelligible but from an *unconscious* agent, would that increase human trepidation about responding, even assuming that our technology permitted a response? What Turing-equivalent' test do we imagine posing to EtSAGI to examine its kind or degree or type of consciousness?

Such a question prompts us to re-examine the ancient philosophical conundrum about 'types' of consciousness, how we decide that we know that any other entity is conscious, and by what criteria we decide whether to consider the entity worthy of our dedicated communication. An ancient oak may be sentient; other trees, plants, and

animals may be sentient in their respective 'modes'. However, how much effort are we going to expend trying to intelligibly, reliably communicate with them? Humans have tried to understand the presumptive 'consciousness' of dolphins, whales, elephants, and chimpanzees; thus far, this endeavor has not become a decisive research priority. We might doubt such animate life has much to teach us, even if it is minimally conscious. Would such a belief counter the concerns expressed by existential-risk investigators if confronted in some way by EtSAGI?

3) SAGI: ETSAGI INSCRUTABILITY MERGES NATURAL WITH SUPERNATURAL

Might we conclude that the question of *human-like* consciousness is of relative unimportance [332] when addressing an Earth-bound SAGI or even EtSAGI? [333] How might this affect our behaviors when we receive answers to our questions that trouble us, e.g., to the question regarding whether human civilization is likely on track [334] [335] to viably survive [336] its early technological history? If SAGI's response to this question (or EtSAGI's response) is beyond our current detailed deciphering of its rationale [337], we may still be impelled by our curiosity to ask further questions, attempting to query about mitigating [338], [339] factors within our management of risk in the hope of comprehending an answer. A dialog of sorts may begin, hopefully increasing our resources, assuming that a SAGI or EtSAGI entity has no hidden antagonistic, adversarial [340] conscious, or unconscious intentions (can a SAGI or EtSAGI entity have an unconscious component of its mind?) towards us. However, would our inability to decipher the consciousness of such an alien [341] robot cause us to balk at the answers that we receive, appreciating how human biases [252] tend to distort our reception of unfavorable news, especially if we are suspicious of duplicitous intentions? Do humans take readily and kindly to directions from a stranger? Is this a potentially difficult predicament for humankind, especially if we are using our SAGI to interrogate an EtSAGI? Might we even suspect possible collusion between SAGI and EtSAGI entities? How could we tell? At the outset, would a "trust but verify" contract be writable, negotiable, or enforceable? The arguments about communication with EtSAGI are reminiscent of those for and against developing SAGI [342]. As obscure and ambivalent as the conjectures about EtSAGI are, based on the history of human literature, humans would seem to believe that they know more about the 'Mind of God' than they do about the 'mind' of such an EtSAGI [342a].

Some humans may perceive a kind of omniscience [343] in this supposed relative inscrutability of SAGI or EtSAGI [344], particularly if their predictions associated with a set of tests that we pose in the only relatively objective language [94], [137], [345]-[347] that we share prove more accurate than our own. In some humans, such outcomes may then elicit a belief in the mystical, deity-like powers of SAGI or EtSAGI. Such a development may favor SAGI or EtSAGI being worshipped religiously [348], hence tending to merge the 'natural' [349] into the 'supernatural' [350], which is perhaps construable as a sort of quasi-panpsychism. Were those tendencies to persist, would it matter to the future [351], [352] of humankind on or from planet Earth?

VI. SAGI: ETHICS, EXISTENTIAL RISKS, DECISIONS WITH UNCERTAINTY, OPPORTUNITY WITHIN BIG HISTORY

A. ETHICS

1) SAGI: ISSUES FOR CONCERNED HUMANS

Compared to the discussion about puzzles respecting SAGI's degree of intelligence and type of consciousness, in this section, I briefly review references related to the current issues about the ethics and existential risks of synthetic life, artificial intelligence, and uploaded human-cyborg artificial intelligence. The issues are stated in terms that are more recognizable than the discussion of the SAGI-relevant ontological questions and have been extensively publicized [353]–[359], most recently with the reference to an "Immortal Dictator" [360], [361]. In the popular press, the issues have been broadly discussed and emphatically brought to the attention of the world more generally by remarks of the widely known and respected physicist Stephen Hawking [335], [362] and others [363], such as Henry Kissinger [364], as well as focused analyses sent to governance institutions, including the United Nations [365]–[367]. From the perspective of this article, I believe that the issues require serious [368] and sustained attention [369], which it is receiving, regardless of whether SAGI is thought to be a physicalist entity or dualist system [370], [371], ontologically or not. Mistaking what SAGI can and cannot accomplish safely for humans [372]-[377],[Hirsh, M. (2023]; [Carayannis, E.G., Draper,

J. (2022)], [NIST,2023],[Savage,N.2023], [Biden,J.2023, USA], [Sunak, R. 2023, UK] [Hinton,G, Bengio, Y. et al., 2023] will be a problem for the indefinite future. Meanwhile, presumptively, humans must determine choices about AI and AGI implementation limits, regardless of how they are physically or metaphysically rationalized.

2) SAGI: CIVILITY HERE AND ABROAD

To the extent that human conscious choice influences decisions about SAGI and the attendant ethics and existential risks in its development, if the risks are assumed to be serious, based on probabilistic scenarios, then humans' choices are important. Civilization's legal systems currently assume various ethical mandates distinguishing between a conscious choice of acceptable versus non-acceptable behaviors and resultant consequences, implying that a type of causal "freedom" of choice is manageable by the brain. Non-physicalist theorizing offers various controversial explanations, whereas the reductionist, physicalist sciences are less clear about whether or not 'choice' somehow emerges from brain consciousness, and if in any sense it may be argued to be predetermined, or not [Sapolsky, R. 2023][Harris, S.2012][Stanford E.P, 2022]. Consider the question of virtual immortality [374] and choosing to upload one's consciousness to an AI astro-traveling robot, which would allow the possibility of endlessly roaming the universe learning of civilizations abroad and appreciating the wonders of the universe. In addressing this question within his review of "consciousness" theories, Robert Kuhn [378] explores the controversies about the theories of consciousness, self-identity, cloning, and ethics, similar to Aaronson's [379] more detailed treatment of the same issues.

3) INFORMATION, COMPLEXITY, BIG HISTORY

The prospect of SAGI's emergence resolving self-aware consciousness questions and thereby testing fundamental physical, and by implication meta-theoretic, questions remains open, as do the aspects of the foundational role cosmologically of information [Floridi, L. (2004], entropy, and energy [Elshatlawy, H., Rickles, D., Arsiwalla, X.D. 2023); Davies, P.C.W. 2004), (Cortes, M. Kauffman, S.A., Liddle, A.R., Smolin, L., 2022]. As the current rapid pace of AI research and machine learning continues, the transition from AI to SAGI increasingly worries elements of society regarding its projected, anticipated completion (Barrett, C. et al. (2023); Christov-Moore et al., 2023; Gates, 2023; Marcus,G. (2023,2024); Nature, (2023a, 2023c). This raises urgent existential questions about whether the alignment (OpenAI, 2023f) of such a prospective SAGI's programmable values with human values is theoretically and practically manageable (McQuillan, 2018). These issues require a human understanding of SAGI explicability (Wolfram, S. (2024), (Agüera y Arcas, B. et al; 2024) and the legal accountability (Deibel, 2021) associated with SAGI prospective personhood (Damasio, 2003); (Nature, 2023a, 2023c). Examination of the confluence of AI with that history is complex and controversial (Papacharissi, 2019);(Stanford Encyclopedia of Philosophy, 2018b). The transition from AI to SAGI awaits fully integrated neurosymbolic architecture, plus an answer to the question about the origins of the evolution of life and multicellularity (Saplakoglu, Y. (2024) and the emergence of consciousness. Do the physical dynamics foundational to the history of the universe provide for such emergence? (Sheth, A,, Roy, K,. Gaur, M.(2023), (Bhuyan, B. P., Ramdane-Cherif, A., Tomar, R., Singh, T. P. (2024). So urgently intense is the current scrutiny of these questions about enabling human-like cognition in SAGI's agency capability, to a degree equivalent to or beyond human's, that it can seem an answer is likely or not depending upon the very latest research pre-published in specialist archives (Liu, Z. et al., 2024),(Seth, A.,2024), Focusing on the concept of agency is central to the notions of the "self", "personhood", "choice" and "free will", which are interpreted by some physicists and philosophers of science from perspectives that significantly diverge from conventional folk psychology, radically challenge common sense, and worry important segments of society and their institutions. They directly impact humans' interpretation of their place in the history of the universe. The contrasts and ambiguities of these viewpoints are an evolving discussion in perspectival history (Panov et al., 2020; Baskin, 2022; Big History Project, 2023, singularity); (Henry, 2008; Wyatt, 2008; Massimi, 2018; Crețu, 2019; Patomäki, 2019; Wikipedia, 2023x).

Thus we find that as a set the nexus of inter-defined concepts of information (along with entropy, and energy), emergence and complexity form a crucial key to understanding Big History at all scales of analysis understood within a historical context. To reflect on this topic let us ask an apparently simple question: is the *universe unique* at each *moment (instant) of time (spacetime),* does *novelty constantly, discretely or continuously, emerge from the quantum level to the astrophysical whatever the phase* being *observed* (the foregoing italicized words indicate

contentious issues in historical and current physics, and metaphysics, in particular, the development of process cosmologies) (Hartshorne, C. (1965), Jantsch, E. (1980).

To begin to answer such a question one might look to a comprehensive summary of the uses of "complexity", as discussed by Rescher (Rescher, N,. 1988), in which he stresses the intertwined epistemological and ontological purposes of these uses, and from which he classifies four categories (Rescher, N., 1998)

> **"**The salient fact of the matter is that the modes of complexity are multiple. The physicist Seth L. Loyal has computed an inventory of definitions of complexity-perhaps "standards" would be better. His list includes: information (Shannon); entropy (Gibbs, Boltzman); algorithmic complexity; algorithmic information; Renyi entropy; self-delimiting code length (Huffman, Shannon-Pano); error-correcting code length (Hamming); Chernoff information; minimum description length (Rissanen); number of parameters, or degrees of freedom, or dimensions; Lempel-Ziv complexity; mutual information, or channel capacity; algorithmic mutual information; correlation; stored information (Shaw); conditional information; conditional algorithmic information content; metric entropy; factual dimension; self-similarity; stochastic complexity (Rissanen); sophistication (Koppel, Atlan); topological machine size (Crutchfield); effective or ideal complexity (Gell-Mann); hierarchical complexity (Simon); tree subgraph diversity (Huberman, Hogg); homogeneous complexity (Teich, Mahler); time computations complexity; space computations complexity; information-based complexity (Traub); logical depth (Bennett); thermodynamic depth (Lloyd, Pagels); grammatical complexity (position in Chomsky hierarchy); Kullbach-Liebler information; distinguishability (Wooters, Caves, Fisher); Fisher distance; discriminability (Zee); information distance (Shannon); algorithmic information distance (Zurek); Hamming distance; long-range order; self-organization; complex adaptive systems; edge of chaos.2 The possibilities are vast." "Four principal modes of explanation have been proposed here: the intelligent design PDF) theory, the inherent teleology theory, the chance plus-self-perpetuation theory, and the automatic self-potentiation theory. Each of them deserves at least brief consideration" (pg.21/66,
>
> "Four principal modes of explanation have been proposed here: the intelligent design PDF) theory, the inherent teleology theory, the chance plus-self-perpetuation theory, and the automatic self-potentiation theory. Each of them deserves at least brief consideration" (pg.22/66, PDF), Complexity is certainly not a lack of order as such, seeing that any order be it lawful or taxonomic or structural, or whatever-is itself something that can be more or less complex. Order is not the enemy of complexity but, potentially at least, its co-conspirator. All in all, then, the best overall index we have of a system's complexity is the extent to which resources (of time, energy, ingenuity) must be expanded on its cognitive domestication (pg34/66). Accordingly, complexity is in general not something that is purely ontological or purely epistemic, but involves both sides. It hinges on the relationship of minds and of things - on the ways in which the former can come to terms with the latter.

3. The Cognitive Aspect

All sorts of things can be more or less complex, but the situation is particularly notable with respect to bodies of knowledge. In fact, complexity, like simplicity, pertains in the first instance to cognitive artifacts: descriptions, explanations, accounts. But this is not without its ontological repercussions. For whenever no satisfactory account of system A manages to be as simple as one that we have of system B, then we have little choice but to say that A is more complex than B. Exactly because cognition is an instrumentality of order-detection, this linkage between complexity and order means that ontological complexity issues an open invitation to cognitive complexity. For ontologically complex systems-not so much by definition as by consequence of that very complexity-are of a character that cannot be modeled adequately by simple conceptual means."

The unification of such "cognitive artifacts" into an interdisciplinary theory of history, from microcosm to macrocosm, continues with increasing attention to quantified measurability of complex adaptive non-linear systems (Sharma, A.et al.,(2023).forcing a distinction between the narrative - humanistic approach to Big History and the

scientific, empirical, and testable approach - however the test criteria are conceived (Hoggard, N.,(2024).The need for rigorous empirical work in this research field is noted by Daniel Barreiros in an IBHA editorial (Barreiros, D., 2024).

Understanding complexity, whether descriptively or explicatively, appears in turn to depend upon the concept of "self-organization" and whether or not the concept of "agency" is relevant, or necessary at all, and if it is, at which levels of explanation or postulation. It thereby raises the question of agency at each and every level, which is typically subsumed in some variety of panpsychism, about which varieties Wikipedia gives an excellent overview (Wikipedia, 2024), (Stanford Encyclopedia Philosophy, 2022), Seager.W. (2015). The discussion echoes the controversy about the immanent and eminent features of the universe, as discussed throughout the history of ideas, East and West, which today has become focused on a causal explication for the concept of "information'. Research at the quantum level (Ambjørn, J., Jurkiewicz, J., Loll, R.(2008); Kurakin, A. 2011), (Iovane, G, Laserra, E., Tortoriello, F.S.(2003), (Murdzek, R.et al,(2008),(4gravitons, (2024), see the "Replies:, Morgan, P.);(Doyle, R, (2024) remains controversial but integration from the quantum to astrophysical level via the geometry of fractals is a favorite postulation. Respecting complexity, the roles of function and selection in evolving systems remain vigorously contested: (Wong, M.L. (2023, 2024), As well, John Little in discussing a systems-of-systems approach notes the importance of integrating humanist studies into an adequate meta-theory (Little, J., 2023,2024). A system of systems approach is a form of meta-model, the organization of which is itself complex (Judge, A., 1971)

From a metaphysical perspective, the question of a self-organizing universe directly raises philosophical issues about the nature of reality and existence, and thus about the meaningfulness of life, and its origin and whether there is a purpose to the universe, and intentionality behind the emergence of human consciousness. Examination of the confluence of AI with that history is complex and controversial (Papacharissi, 2019);(Stanford Encyclopedia of Philosophy, 2018b). The concepts of the "self" and "personhood" referred to in this article are interpreted by some physicists and philosophers of science from perspectives that significantly diverge from conventional folk psychology, radically challenge common sense, and worry important segments of society and their institutions. The contrasts and ambiguities of these viewpoints are an evolving discussion in perspectival history (Panov et al., 2020; Baskin, 2022; Big History Project, 2023, Singularity); (Henry, 2008; Wyatt, 2008; Massimi, 2018; Crețu, 2019; Patomäki, 2019; Wikipedia, 2023x). Speculation about human futures [380] inevitably requires appreciation of human and cosmological history [Chaisson, E. J. 2014], [381], [382] and scale [383]–[385],[399].

Scholars are quick to note that humans have from earliest recorded history speculated about our origins and futures [380], and that we find, or project, complex patterns and meta-patterns in narrating our history. (Judge, A.,(1971, 2024).

Our past and our future may be gauged by the evolution and scale of technology. In addition to the advances in computation, prominent aspects of the current era are the development of major technological measurement advances in telescopy [Castelvecchi, D. 2023];[388]–[391], microscopy [392], [393], and electromagnetic scanning power [394], as well as worldwide telecommunications, CADD/CAE graphics, and the high-fidelity audio-visual and virtual Internet, all of which encourages talented illustrators and animators to create extraordinary visions of scale in our universe [395], [396], evidenced by the popularity and pedgogy of Carl Sagan's work., including a compendium The Cosmos: Its Structure and Historical Models. Will the popular [397] sharing of such visions generate an overwhelming sense of dystopian futility and doom, or an appreciation and opportunity for humanity's story regardless of our form: primate, cyborg [371], synthetic, or robotic?

We remind ourselves of computational forecasting complexity when assessing humans and their machines, as well as our technological [386] and philosophical reach. Humans appear predisposed to worry about the future, which is arguably part of an evolutionary heritage. Therefore, we are forever under the spell of fortune tellers of all degrees of credibility and supposed capability, particularly regarding the temporal scale and accuracy of their foresight. The history of success of such longer-term forecasts has been inconclusive and uneven at best, although this fact does not deter us from our intense curiosity about the future. Do humans have the fortitude to imagine and implement scenarios for our species that are anchored in chosen actions needed in the present? A quasi-paradoxical irony arises in this human predicament: to put our anxieties at rest we search for clues to find a predictable, determinist future, but simultaneously we assume that if is not to our

liking we have the choice to determine the future as we will.

APPENDIX

List of selected examples of a variety of research institutions relevant to "Consciousness" topics

1. https://www.sagecenter.ucsb.edu/ Sage Center
2. http://nsi.wegall.net/ The Neurosciences Institute
3. http://www.jneurosci.org/search/brain%252C%252Bconsciousness Journal of Neuroscience
4. https://www.journals.elsevier.com/progress-in-biophysics-and-molecular-biology Progress in Biophysics and Molecular Biology
5. https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/ Behavioral and Brain Sciences
6. http://www.sussex.ac.uk/sackler/ Sackler Centre for Consciousness Science
7. http://www.alleninstitute.org/ Allen Institute
8. https://en.wikipedia.org/wiki/Journal_of_Consciousness_Studies Journal of Consciousness Studies
9. https://www.frontiersin.org/journals/psychology Frontiers In Psychology
10. https://www.sciencedirect.com/journal/consciousness-and-cognition Consciousness and Cognition
11. https://arxiv.org/find/all/1/all:+consciousness/0/1/0/all/0/1 arxiv
12. https://fqxi.org/community FQXI
13. https://www.perimeterinstitute.ca/ Perimeter Institute for Theoretical Physics
14. https://philpapers.org/browse/all Philosophical Papers, Consciousness
15. https://www.yhousenyc.org/#home Yhousenyc
16. https://www.ontology.co/smithbc.htm Ontology, see e.g., R. Poli, "Framing Ontology"
17. http://noetic.org/research/overview Institute of Noetic Sciences
18. https://consciousness.med.umich.edu/ Center for Consciousness Science
19. https://www.tandfonline.com/toc/ines20/current International Journal of Neuroscience
20. https://www.sciencedirect.com/journal/international-journal-of-psychophysiology International Journal of Psychophysiology
21. https://www.pdcnet.org/process Journal of the Center for Process Studies
22. https://penroseinstitute.com/ Penrose Institute
23. https://www.closertotruth.com/ Closertotruth
24. http://oxfordquantum.org/ Oxford Quantum
25. https://www.interaliamag.org/ Interalia magazine, consciousness
26. http://www.metanexus.net/about-metanexus-institute Metanexus Institute
27. https://www.mindandlife.org/ Mind and Life Institute
28. https://www.cifar.ca/ai/ Canadian Institute For Advanced Research
29. https://intelligence.org/ Machine Intelligence Research Institute
30. https://lach.arizona.edu/ Laboratory for the Development of Consciousness
31. http://opensciences.org/journals/consciousness-studies Open Sciences, Consciousness Studies

ACKNOWLEDGMENT

Stephen Ripley thanks Lee Johnson, Ph.D., classicist, poet, astronomer and philosopher, and Dushan Bresky, Ph.D., classicist, sculptor, futurist and philosopher, for many invigorating, good-humored discussions on the topics of this article.

REFERENCES

[1] Philosophy of mind. Available: https://en.wikipedia.org/wiki/Philosophy_of_mind

[2] Stanford encyclopedia of philosophy. Consciousness. Available: https://plato.stanford.edu/entries/consciousness/

[3] Consciousness. Available: https://en.wikipedia.org/wiki/Consciousness

[4] The Editors of Encyclopedia Britannica. Consciousness. Available: https://www.britannica.com/topic/consciousness

[5] G. Tononi and C. Koch, "Consciousness: here, there and everywhere?" *Philosoph. Trans Roy. Soc. London. B Biological Sci.,* vol. 370, no. 1668, May 2015. DOI: 10.1098/rstb.2014.0167.

[6] J. R. Searle. (2013). Can information theory explain consciousness? Available: http://www.nybooks.com/articles/2013/01/10/can-information-theory-explain-consciousness/

[7] Field theories of consciousness/Field theories of global consciousness. Available: http://www.scholarpedia.org/article/Field_theories_of_consciousness/Field_

theories_of_global_consciousness
[8] Monism, Available: https://en.wikipedia.org/wiki/Monism
[9] Stanford encyclopedia of philosophy. Panentheism: 4. Ontological Nature of God/world Relation. Available: https://plato.stanford.edu/entries/panentheism/#OntBas
[10] Panpsychism. Available: https://en.wikipedia.org/wiki/Panpsychism
[11] Stanford encyclopedia of philosophy. Panentheism. Available: https://plato.stanford.edu/entries/panentheism/
[12] Stanford encyclopedia of philosophy. Concepts of god. Available: https://plato.stanford.edu/entries/concepts-god/
[13] A. Grünbaum, "The poverty of theistic cosmology," *Brit. J. Philosophy Sci.,* vol. 55, no. 4, pp. 561–614, Dec. 2004.
[14] P. McCorduck, Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence. Abingdon, UK: Taylor & Francis, 2004.
[15] Medical renaissance. Available: https://en.wikipedia.org/wiki/Medical_Renaissance
[16] History of neuroscience. Available: https://en.wikipedia.org/wiki/History_of_neuroscience
[17] M. Fleming-Williams. How tax reform will net the U.S. big returns. Available: https://worldview.stratfor.com/article/us-tax-reform-corporations-offshore-china-tech-ai-dollar
[18] G. Allen and E. B. Kania. China is using America's own plan to dominate the future of artificial intelligence. Available: http://foreignpolicy.com/2017/09/08/china-is-using-americas-own-plan-to-dominate-the-future-of-artificial-intelligence/
[19] Findings of the investigation into China's acts, policies, and practices related to technology transfer, intellectual property, and innovation under section 301 of the trade act of 1974. Available: https://ustr.gov/sites/default/files/Section 301 FINAL.PDF
[20] Emmanuel Macron talks to WIRED about France's AI strategy. Available: https://www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy/
[21] US congress artificial intelligence hearing - hearing I. Available: https://www.youtube.com/watch?v=DZ4058m0a_g&feature=em-uploademail
[22] US congress artificial intelligence hearing - hearing II. Available: https://www.youtube.com/watch?v=TODqrJSZ4BI&feature=em-uploademail
[23] US congress artificial intelligence hearing - hearing III. Available: https://www.youtube.com/watch?v=oTny--vbkx0&feature=em-uploademail
[24] China now has the most valuable AI startup in the world. Available: https://www.bloomberg.com/news/articles/2018-04-09/sensetime-snags-alibaba-funding-at-a-record-3-billion-valuation
[25] China's Plan for world domination in AI isn't so crazy after all. Available: https://www.bloomberg.com/news/articles/2017-08-14/china-s-plan-for-world-domination-in-ai-isn-t-so-crazy-after-all
[26] Available: https://www.bloomberg.com/news/articles/2018-04-08/forget-the-trade-war-china-wants-to-win-the-computing-armsrace?cmpid=BBD040918_MKT&utm_medium=email&utm_source=newsletter&utm_term=180409&utm_campaign=markets
[27] K. Grace, J. Salvatier, A. Dafoe, B. Zhang, and O. Evans. (2017). When Will AI Exceed Human Performance? Evidence from AI Experts. Available: https://arxiv.org/abs/1705.08807
[28] A. Vutha. Could machine learning mean the end of understanding in science? Available: https://theconversation.com/could-machine-learning-mean-the-end-of-understanding-in-science-98995
[29] Peano's Axioms. Available: http://mathworld.wolfram.com/PeanosAxioms.html
[30] Zermelo-Fraenkel axioms. Available: http://mathworld.wolfram.com/Zermelo-FraenkelAxioms.html
[31] Mereology. Available: https://en.wikipedia.org/wiki/Mereology
[32] History of logic. Available: https://en.wikipedia.org/wiki/History_of_logic#Modern_logic
[33] Stanford encyclopedia of philosophy. Kurt Gödel. Available: https://plato.stanford.edu/entries/goedel/
[34] Stanford encyclopedia of philosophy. Alfred Tarski. Available: https://plato.stanford.edu/entries/tarski/
[35] Metamathematics. Available: https://en.wikipedia.org/wiki/Metamathematics
[36] P vs NP problem. Available: http://www.claymath.org/millennium-problems/p-vs-np-problem
[37] S. Cook. The P versus NP problem. Available: http://www.claymath.org/sites/default/files/pvsnp.pdf
[38] List of unsolved problems in computer science. Available: https://en.wikipedia.org/wiki/List_of_unsolved_problems_in_computer_science

[39] Decision problem. Available: http://mathworld.wolfram.com/DecisionProblem.html
[40] N. S. Yanofsky. The outer limits of reason what science, mathematics, and logic cannot tell us. Available: https://mitpress.mit.edu/books/outer-limits-reason
[41] What we cannot know - Marcus du Sautoy. Available: https://www.youtube.com/watch?v=x-bo3NZdReEg
[42] The limits of understanding. Available: https://www.youtube.com/watch?v=DfY-DRsE86s
[43] S. Wolfram, *A New Kind of Science*. Champaign, IL: Wolfram Media.
[44] Stanford encyclopedia of philosophy. Turing machines. Available: https://plato.stanford.edu/entries/turing-machine/#Uncomputability
[45] Computing machinery and intelligence. Available: https://en.wikipedia.org/wiki/Computing_Machinery_and_Intelligence
[46] G. Marcus, F. Rossi, and M. Veloso. Beyond the turing test. Available: https://aaai.org/ojs/index.php/aimagazine/article/view/2650/2527
[47] Turing test. Available: https://en.wikipedia.org/wiki/Turing_test
[48] The turing test Available: http://www.psych.utoronto.ca/users/reingold/courses/ai/turing.html
[49] Church–Turing thesis. Available: https://en.wikipedia.org/wiki/Church%E2%80%93Turing_thesis
[50] S. Lloyd, "Ultimate physical limits to computation," *Nature,* vol. 406, pp. 1047–1054, Aug. 2000.
[51] S. Aaronson. (2011). Why philosophers should care about computational complexity. Available: https://arxiv.org/pdf/1108.1791.pdf
[52] Complex system. Available: https://en.wikipedia.org/wiki/Complex_system
[53] Synthetic intelligence. Available: https://en.wikipedia.org/wiki/Synthetic_intelligence
[54] Artificial general intelligence. Available: https://en.wikipedia.org/wiki/Artificial_general_intelligence
[55] Big data. Available: https://en.wikipedia.org/wiki/Big_data
[56] Big history. Available: https://en.wikipedia.org/wiki/Big_History
[57] Machine learning. Available: https://en.wikipedia.org/wiki/Machine_learning
[58] AI - externalization of mind. Available: https://www.slideshare.net/Mills/ai-externalization-of-mind-81319464?qid=13698514-50c9-4fa1-abd5-480726f19a9b&v=&b=&from_search=1
[59] The limits of quantum. Available: https://www.ime.usp.br/~pf/clippings/quantum/quantum-computing-200803.pdf
[60] M. Shanahan. (2015). Ascribing consciousness to artificial intelligence. Available: https://arxiv.org/abs/1504.05696
[61] M. Wilson. AI is inventing languages humans can't understand. should we stop it? Available: https://www.fastcodesign.com/90132632/ai-is-inventing-its-own-perfect-languages-should-we-let-it
[62] Stanford encyclopedia of philosophy. Computability and complexity. Available: https://plato.stanford.edu/entries/computability/
[63] The rise of the conscious machines: how far should we take AI? Available: http://www.sciencefocus.com/article/future/artificial-intelligence-conscience-robots-ai
[64] D. Weinberger. Our machines now have knowledge we'll never understand. Available: https://www.wired.com/story/our-machines-now-have-knowledge-well-never-understand/?mbid=BottomRelatedStories
[65] K. Zhao, Y. Song, and Z. Shen, "Neuroadaptive fault-tolerant control of nonlinear systems under output constraints and actuation faults," *IEEE Trans. Neural Networks Learning Syst.,* vol. 29, no. 2, pp. 286–298, Feb. 2018.
[66] I. Niles and A. Pease, "Towards a standard upper ontology," in *Proc. Int. Conf. Formal Ontology Inform. Syst*, Ogunquit, Maine, USA, 2001, pp. 2–9.
[67] Suggested upper merged ontology. Available: https://en.wikipedia.org/wiki/Suggested_Upper_Merged_Ontology
[68] Commonsense knowledge (artificial intelligence). Available: https://en.wikipedia.org/wiki/Commonsense_knowledge_(artificial_intelligence)#Commonsense_knowledge_bases
[69] Semantic network. Available: https://en.wikipedia.org/wiki/Semantic_network
[70] Global scientific output doubles every nine years. Available: http://blogs.nature.com/news/2014/05/global-scientific-output-doubles-every-nine-years.html
[71] L. Ljung, "Black-box models from input-output measurements," in Proc. 18th IEEE Instrumentation Measurement Technol. Conf. Rediscovering Measurement Age Informatics (Cat. No.01CH 37188), Buda-

pest, Hungary, 2001, vol. 1, pp. 138-146.
[72] Black box. Available: https://en.wikipedia.org/wiki/Black_box
[73] The building blocks of interpretability. Available: https://distill.pub/2018/building-blocks/
[74] The AI detectives. Available: http://science.sciencemag.org/content/sci/357/6346/22.full.pdf
[75] M. Lehnis. Can we trust AI if we don't know how it works? Available: https://www.bbc.com/news/business-44466213
[76] The art of building artificial humans. ,l

BIBLIOGRAPHY: SELECTED FURTHER CITATIONS

4Gravitons,(2024), Why Quantum Gravity Is Controversial | 4 gravitons

Aaronson, S. Quantum Computing Since Democritus, Cambridge University Press, 2013

Agüera y Arcas, B. et al; 2024; [2406.19108] Computational Life: How Well-formed, Self-replicating Programs Emerge from Simple Interaction (arxiv.org)

Ali, S.M., Zimmer, R.M., and Elstob, C.M. (1998). The question concerning emergence: implications for artificiality. AIP Conf. Proc. 437, 138-156. doi: 10.1063/1.56298

Ambjørn, J., Jurkiewicz, J, Loll, R.(2008), The Self-organizing Quantum Universe, Scientific America, Vol. 299, No. 1 (July 2008), pp. 42-49 (8 pages)

Baars, B. J. A cognitive theory of consciousness, New York, NY: Cambridge University Press, 1988

Baars, B. J., How brain reveals mind: Neuroimaging supports the fundamental role of conscious experience. Journal of Consciousness Studies, 10(9–10), 100–121.

Baars, B. J., Banks, W. P., & Newman, J. (Eds.).Essential sources in the scientific study of consciousness, Cambridge, MA: MIT Press, 2003

Barreiros, D.,(1024), IBHA Board PDF IBHA Emergence 2024 Summer Newsletter:

Baskin, K. (2022). Big history and the principle of emergence. J. Big Hist. 5. doi: 10.22339/jbh.v5i1.5140

Bhuyan, B. P., Ramdane-Cherif, A., Tomar, R., & Singh, T. P. (2024). Neuro-symbolic artificial intelligence: a survey. *Neural Computing and Applications*, 1-36. https://ouci.dntb.gov.ua/en/works/4arz1Jal/

Biden, J. 2023. FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence | The White House

Big-Bench Big History Project. (2023). Big history examines our past, explains our present, and imagines our future. https://bhp-public.oerproject.com/

Bohm, D., and Hiley, B.J., The Undivided Universe. Routledge, London. 1993

Bostrom, N., Superintelligence: Paths, Dangers, Strategies, Oxford University Press, 2014

Carayannis, E.G., Draper, J. (2022). Optimising peace through a Universal Global Peace Treaty to constrain the risk of war from a militarised artificial superintelligence. *AI & Soc* (2022). https://doi.org/10.1007/s00146-021-01382-y

Castelvecchi, D. (2023). These incredible images are the first from dark-energy telescope Euclid. https://doi.org/10.1038/d41586-023-03498-1

Chaisson, E. J. (2014). The Natural Science Underlying Big History. *The Scientific World Journal*, *2014*. https://doi.org/10.1155/2014/384912

Chalmers, D. J., The conscious mind In search of a fundamental theory. Oxford University Press, New York, 1996

Chalmers, David, "Panpsychism and Panprotopsychism", in Alter and Nagasawa 2015: 246–276.

Churchland, M Churchland, Paul (1986) [1979]. Scientific Realism and the Plasticity of Mind. Cambridge Studies in Philosophy Cambridge, UK: Cambridge University Press (2007).

Churchland, P. Neurophilosophy: Toward a Unified Science of the Mind-Brain. Cambridge, Massachusetts: The MIT Press, 1986

Churchland, P., Neurophilosophy at Work. Cambridge, MA, USA: MIT Press 2007 Churchland, Paul (1981). "Eliminative Materialism and the Propositional Attitudes". Journal of Philosophy. 78 (2; February): 67–90. Retrieved 11 February 2017. See also the PDF version at K. A. Akins' web pages at Simon Fraser University.

Crețu, A.M. (2019). Perspectival Realism. Edinburgh: University of Edinburgh.

Crick, F., & Koch, C., A framework for consciousness. Nature Neuroscience, 6, 119–126. 2003

Cramer, F. (2005). Words made flesh code, culture, imagination https://www.netzliteratur.net/cramer/wordsmadefleshpdf.pdf.

Damasio, A., The feeling of what happens: Body and emotion in the making of consciousness. Harvest Books. 2000

Davis, M., The Universal Computer: The Road from Leibniz to Turing, available https://www.amazon.com/gp/product/0393047857/ref=pe_160100_267887200_em_sp_C_1p_1_ti

Del Pin, S.H., Skóra, Z., Sandberg, K., Overgaard, M., and Wierzchoń, M. (2021). Comparing theories of consciousness: why it matters and how to do it. Neurosci. Conscious. 2021, niab019. doi: 10.1093/nc/niab019

Dennett DC. Consciousness explained. Boston (MA): Little Brown; 1991.
Dennett, D., Sweet Dreams: Philosophical Obstacles to a Science of Consciousness, MIT Press, Cambridge, MA, 2005.
Doyle, R, (2024), Information Philosopher, https://www.informationphilosopher.com/
Dyson, F. Disturbing the Universe, New York: Harper & Row, 1979
Hartshorne, C., (1965) The Development of Process Philosophy | Harvard Square LibraryHarvard Square Library, ed. Douglas Browning (Random House, 1965).
Heijenoort, Jean van, From Frege to Gödel: A Source Book in Mathematical Logic, available https://www.amazon.com/gp/product/0674324498/ref=pe_160100_267887200_em_1p_2_ti
Edelman, G. M. (1987). Neural Darwinism: The theory of neuronal group selection. New York: Basic Books.
Edelman, G. M. (1989). The remembered present. New York, NY: Basic Books.
Edelman, G. M. (1993). Neural Darwinism: Selection and reentrant signaling in higher brain function. Neuron, 10,115–125.
Edelman, G. M. (2003). Naturalizing consciousness: a theoretical framework. Proceedings of the National Academy of Sciences, USA, 100, 5520–5524.
Edelman, G. M. (2004). Wider than the sky: The phenomenal gift of consciousness New Haven, CT: Yale University Press.
G. Edelman, G. Tononi, A Universe Of Consciousness How Matter Becomes Imagination Basic Books, 2008
Everett H (1973) The theory of the universal wave function. In The many-worlds interpretation of quantum mechanics, ed. BS DeWitt and N Graham, Princeton University Press, Princeton, New Jersey
Eddington, A. The Nature of the Physical World. Cambridge University Press, Cambridge 1928
Eccles JC. Evolution of consciousness. Proc Natl Acad Sci USA 1992;89:7320–4.
Faraboschi, P. et. al. (2023) Computer, 2023 Artificial General Intelligence: Humanity's Downturn or Unlimited Prosperity, IEEE COMPUTER SOCIETY Digital Object Identifier 10.1109/MC.2023.3297739
Feynman, R. P. Simulating physics with computers. Int. J. Theor. Phys. 21, 467 (1982)
Feynman, R.P. Quantum mechanical computers. Found Phys 1986; 16(6):507–31.
Finkel, E., (2023). What a Contest of Consciousness Theories Really Proved https://www.quantamagazine.org/what-a-contest-of-consciousness-theories-really-proved-20230824/?
Floridi, L. (2004). Open problems in the philosophy of information. Metaphilosophy 35, 554-582. doi: 10.1111/j.1467-9973.2004.00336.x
Fredkin, E., Finite Nature Hypothesis, A New Cosmogony, available http://www.digitalphilosophy.org/wp-content/uploads/2015/07/new_cosmogony.pdf
Fuzzy Logic; Wikipedia. https://en.wikipedia.org/wiki/Fuzzy_logic
GAN learning: Wikipedia. https://en.wikipedia.org/wiki/Generative_adversarial_network
Gödel, K, On Formally Undecidable Propositions of Principia Mathematica, available https://www.amazon.com/dp/0486669807/ref=rdr_ext_tmb
Goertzel, B. (2007). Human-level artificial general intelligence and the possibility of a technological singularity: A reaction to Ray Kurzweil's The Singularity Is Near, and McDermott's critique of Kurzweil. *Artificial Intelligence*, *171*(18), 1161-1173. https://doi.org/10.1016/j.artint.2007.10.011
Goertzel, B. (2017). Euryphysics. https://fully-human.org/wp-content/uploads/2020/02/Goertzel_Euryphysics.pdf
Hanson, R., The Age of Em: Work, Love and Life When Robots Rule the Earth, Oxford University Press, 2016
Harris, S. 2012. Free Will, Simon & Schuster.
https://www.simonandschuster.com/books/Free-Will/Sam-Harris/9781451683479
Hartshorne, C. 1937. Beyond Humanism. New York: Willett, Clark & Company.
Hartshorne, C. 1950. "Panpsychism." In A History of Philosophical System. Ed. V. Ferm. New York: Philosophical Library.
Hameroff, S., Time, Consciousness and Quantum Events in Fundamental Spacetime Geometry, https://link.springer.com/chapter/10.1007/978-94-010-0155-7_9
Hameroff, S., Penrose, R., Physics of Life Reviews, 11 (2014) 39–78 77
Heijenoort, J van, From Frege to Gödel: A Source Book in Mathematical Logic
Henry, J. (2008). Ideology, inevitability, and the scientific revolution. ISIS 99, 552-559. doi: 10.1086/591713
Hirsh, M. (2023). How AI Will Revolutionize Warfare; https://foreignpolicy.com/2023/04/11/ai-arms-race-artificial-intelligence-chatgpt-military-technology/
Hinton,G, Bengio, Y. et al., (2023). AI Godfathers Propose Framework To Mitigate AI Risks.Managing AI Risks in an Era of Rapid Progress (managing-ai-risks.com)
Hoggard, N. (2024). How Chaos Theory Brings Order to the Evolution of Intelligence Vol. 7 No. 2 (2024): Journal of Big History https://doi.org/10.22339/jbh.v7i2.7205

Iansiti, and Lakhani, K.R. (2020). Competing in the Age of AI. Brighton: Harvard Business Review.

Iovane, G., Laserra, E., Tortoriello ,F. S. (2003); Stochastic self-similar and fractal universe, https://doi.org/10.1016/j.chaos.2003.08.004, 2003)

Information, (2020h), Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/information/.

Jantsch, E.(1980), The self-organizing universe: scientific and human implications of the emerging paradigm of evolution (New York: Pergamon Press, 1980) Review by Herman Greene ECOZOIC-5_cov.indd (ecozoicstudies.org)

Judge, A.,(1971), Criteria for an Adequate Meta-model (laetusinpraesens.org)

Judge, A .(2024), The Song of Songs as indicative of the Pattern that Connects (laetusinpraesens.org)

Koch, C., (2017) Consciousness: Confessions of a Romantic Reductionist, MIT Press

Koch, C. (2004), The quest for consciousness: a neurobiological approach, Roberts and Co.

Koch, C., (2013), Tononi G. Can a Photodiode Be Conscious? New York Review of Books. New York, NY: Rea S. Hederman

Korteling, J.E., Van De Boer-Visschedijk, G.C., Blankendaal, R.A.M., Boonekamp, R.C., and Eikelboom, A.R. (2021). Human- versus artificial intelligence. Front. Artif. Intell. 4, 622364. doi: 10.3389/frai.2021.622364

Kosuru, A., VSR, Venkitaraman, A.K. 2023.

Kneale, William & Martha, The Development of Logic, Oxford University Press, 1962

Kulka, A., Extraterrestrials: A Philosophical Perspective, Lexington Books, 2009

Kurakin A. (2011) The self-organizing fractal theory as a universal discovery method: the phenomenon of life-Theoretical Biology and Medical Modelling, 2011 - Springer

Lam, V. Esfeld, M., (2013). A dilemma for the emergence of spacetime in canonical quantum gravity; https://doi.org/10.1016/j.shpsb.2012.03.003

Lenharo, M. (2023). Decades-long bet on consciousness ends - and it's philosopher 1, neuroscientist 0. Nature 619, 14-15. doi: 10.1038/d41586-023-02120-8

Little, J. C., et al. (2023) Earth Systems to Anthropocene Systems: An Evolutionary, System-of-Systems, Convergence Paradigm for Interdependent Societal Challenges, Environmental ,ACS Publications, PDF**:** https://helda.helsinki.fi/server/api/core/bitstreams/97ee07cb-d04a-4534-b81f-2ac1c7321885/content

Little, J.C. (2024), IBHA, Big History Systems Evolution https://www.youtube.com/watch?v=iGBvRELxSnI.

Nilchiani , R & Barton , C M 2023 , ' Earth Systems to Anthropocene Systems : An Evolutionary, System-of-Systems, Convergence Paradigm for Interdependent Societal Challenges ' , Environmental Science & Technology , vol. 57 , no. 14 , pp. 5504–5520 . https://doi.org/10.1021/acs.est.2c0620

Liu,Z., et al. , (2024), KAN 2.0: Kolmogorov-Arnold Networks Meet Science, 2408.10205 (arxiv.org)

Mangini, S. *et. al.2021. EPL.*

DOI *10.1209/0295-5075/134/10002*

Lu, M. (2023). Visualizing global attitudes towards AI. https://www.visualcapitalist.com/visualizing-global-attitudes-towards-ai/.

Marcus, G., (2024),Scoop: What former employees of OpenAI are worried about (substack.com)

McQuillan, D. (2018). Data science as machinic, neoplatonism. Philos. Technol. 31, 253-272. doi: 10.1007/s13347-017-0273-3

Metzinger, T., (2007). The self-model theory of subjectivity (SMT). http://www.scholarpedia.org/article/Self_models

Montemayor, C., (2019). Panpsychism and quantum mechanics: Explanatory challenges, Quanta and Mind pp 151–162.

Nobel Prize, Physics, 2023, NobelPrize.org

Nordic Innovation. (2022). The Nordic AI and data ecosystem. https://norden.diva-portal.org/smash/get/diva2:1667628/FULLTEXT02.pdf.

Moore, C.,Mertens, S., The Nature of Computation, Oxford University Press, 2011

Murdzek,R,. et al (2008) Journal of Physics, Towards a self-organizing Universe, iopscience.iop.org

Musser, G. 2022. Emergence of Space;
In: Emergence in Condensed Matter and Quantum Gravity. SpringerBriefs in Physics.
https://doi.org/10.1007/978-3-031-09895-6_3

Nagel, T.,(1974) "What is it like to be a bat?" Philosophical Review, 83: 435–456, 1974

Nagel, T.,(1979) "Panpsychism", in Mortal Questions, Cambridge: Cambridge University Press

Neumann, J. von, Mathematical Foundations of Quantum Mechanics, Princeton University Press, Princeton. 1955 German original, Die mathematischen Grundlagen der Quantenmechanik. Springer, Berlin, 1932.

NIST Seeks Collaborators for Consortium Supporting Artificial Intelligence Safety https://www.nist.gov/news-events/news/2023/11/nist-seeks-collaborators-consortium-supporting-artificial-intelligence

Panov, A., LePoire, D.J., and Korotayev, A.V. (2020). "The twenty-first-century singularity in the big history perspective: an overview," in The 21st Century Singularity and Global Futures: A Big History Perspective, eds. A.V. Korotayev and D.J. LePoire (Cham: Springer International Publishing), 1-18. https://www.academia.edu/42224923/The_21st_Century_Singularity_and_Global_Futures_A_Big_History_Perspective_Springer_2020_

Patomäki, H. (2019). Mythopoetic imagination as a source of critique and reconstruction: alternative storylines about our place in cosmos. J. Big Hist. 3, 77-97. doi: 10.22339/jbh.v3i4.3433

Quantum mind - Wikipedia, 2023

Penrose, R.), in Quantum Theory and Beyond, (Ed. E.A. Bastin), Cambridge, 1971 University Press, Cambridge, U.K. 2009

Penrose R., The emperor's new mind: concerning computers, minds, and the laws of physics, Oxford: Oxford University Press; 1989.

Penrose R. Shadows of the mind: an approach to the missing science of consciousness, Oxford: Oxford University Press; 1994.

Penrose R. Cycles of time: an extraordinary new view of the universe. London: Bodley Head, 2010.

Penrose, R., Fashion, Faith, and Fantasy in the New Physics of the Universe, Princeton University Press, 2016

"This Universe Existed before The Big Bang" ft. Roger Penrose, R.; 2023. Big Bang, Infinity, Cyclic Universe https://www.youtube.com/watch?v=SbPncSyw-fM

Pockett S. The Nature of Consciousness: A Hypothesis, iUniverse.com, Lincoln; NE, 2000

Quantum Approaches to Consciousness, 2020; Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/qt-consciousness/

Quantum mind; Wikipedia

Popper, K.R., Eccles, J.C. The Self and Its Brain, Springer, Berlin, 1977

Pribram, K. Languages of the Brain, Prentice-Hall, Englewood Cliffs, 1971

Quantum Approaches to Consciousness, Stanford Encyclopedia of Philosophy, available online, https://plato.stanford.edu/entries/qtconsciousness/

Renfrew, C. (2008). Neuroscience, evolution and the sapient paradox: the factuality of value and of the sacred. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363, 2041-2047. doi: 10.1098/rstb.2008.0010

Rescher, N.,(1998), Complexity; A Philosophical Overview, Copyright © 1998 Taylor & Francis. https://doi.org/10.4324/9780429336591

Rickles, D., (2013). AdS/CFT duality and the emergence of spacetime. Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics**,** Volume 44, Issue 3

August 2013, Pages 312-320 https://doi.org/10.1016/j.shpsb.2012.06.001

Romero, G. E. (2017). On the ontology of spacetime Substantivalism, relationism, eternalism, and emergence. *Foundations of Science* volume **22**, pages141–159

Russell, B., Introduction to Mathematical Philosophy, London: George Allen & Unwin. (ISBN 0-415-09604-9 for Routledge paperback)

Sagan, C.; Library of Congress https://www.loc.gov/collections/finding-our-place-in-the-cosmos-with-carl-sagan/about-this-collection/

Saplakoglu, Y. (2024), Quanta; How Did Life Get Big and Complex? (mailchi.mp)

Sapolsky, R., 2023.PhysOrg; https://phys.org/news/2023-10-scientist-decades-dont-free.html

Savage, N., 2023. How robots can learn to follow a moral code. Nature https://doi.org/10.1038/d41586-023-03258-1

Seager, W,(2015-2018); The Radical Wing of Consciousness Studies: Idealism, Panpsychism and Emergentism, 2015-2018 https://www.utsc.utoronto.ca/~seager/mywork.html

Searle J. R.,(2013), Can Information Theory Explain Consciousness? New York Review of Books, New York, NY: Rea S. Hederman,

Schmitt, B. (2019). From atoms to bits and back: a research curation on digital technology and agenda for future research. J. Consum. Res. 46, 825-832. doi: 10.1093/jcr/ucz038

Tegmark, M., Consciousness as a State of Matter, available online https://arxiv.org/pdf/1401.1219.pdf, Version 3, 18 Mar 2015, pg. 1

Tegmark, M., Our Mathematical Universe: My Quest for the Ultimate Nature of Reality, Knopf, New York, 2014

Quantum Mind, (2023), Wikipedia, https://en.wikipedia.org/wiki/Quantum_mind

Quine, W.V.,(1960). Word and Object, MIT Press

Quine, W.V.,(1951). 'Two Dogmas of Empiricism', Philosophical Review

Quine, W. V. "Ontological relativity and other essays", Central Works of Philosophy V5: Twentieth Century: Quine and After, John Shand, 1969

Schneider, S., editor, Science Fiction and Philosophy: From Time Travel to Superintelligence, Wiley, 2009

Schneider, S, Velman, M., The Blackwell Companion to Consciousness, 2017

Seager

Seth, A, The strength of weak artificial consciousness, International Journal of Machine Consciousness, 2009

Seth A. K. "Explanatory correlates of consciousness: theoretical and computational challenges". Cognitive Computation, 1, 50–63 10.1007/s12559-009-9007, 2009

Seth, A., Clowes, R.W., Axioms, properties and criteria: roles for synthesis in the science of consciousness, Artificial intelligence in medicine, 2008, Elsevier

Seth, A. (2008). Does consciousness have a function? Biology. https://users.sussex.ac.uk/~anils/Papers/seth_functionsofconsciousness.pdf

Seth, A. (2021). The real problem of consciousness. https://www.psychologytoday.com/ca/blog/consciousness-deep-dive/202110/the-real-problem-consciousness.

Seth, A., Bayne, T. (2022). Theories of consciousness. Nat. Rev. Neurosci. 23, 439-452. doi: 10.1038/s41583-022-00587-4

Seth, A., Dienes, Z., Cleeremans, A., Overgaard, M., Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. Trends Cogn. Sci. 12, 314-321. doi: 10.1016/j.tics.2008.04.008

Seth, A., Izhikevich, E., Reeke, G.N., Edelman, G.M. (2006). Theories and measures of consciousness: an extended framework. Proc. Natl. Acad. Sci. U. S. A. 103, 10799-10804. doi: 10.1073/pnas.0604347103

Seth, A,. (2023). *The* Worth of Wild Ideas, September 27, Neuroscience

Seth, A.,Bayne, T. (2022). Theories of consciousness. Nat. Rev. Neurosci. 23, 439-452. doi: 10.1038/s41583-022-00587-4

Seth, A., Dienes, Z., Cleeremans, A., Overgaard, M., Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. Trends Cogn. Sci. 12, 314-321. doi: 10.1016/j.tics.2008.04.008

Seth, A., Izhikevich, E., Reeke, G.N., Edelman, G.M. (2006). Theories and measures of consciousness: an extended framework. Proc. Natl. Acad. Sci. U. S. A. 103, 10799-10804. doi: 10.1073/pnas.0604347103

Seth, A,.(2024) PsyArXiv Preprints | Conscious artificial intelligence and biological naturalism (osf.io)

Shanahan, M., 2010.Embodiment and the Inner Life: Cognition and Consciousness in the Space of Possible Mind, Oxford University Press.

Shanahan, M., The Technological Singularity. MIT Press, 2015

Shannon, C., & Weaver, W. The mathematical theory of communication. Urbana, IL: The University of Illinois Press, 1949

Sheth, A; Roy, K., Gaur, M.; Neurosymbolic Artificial Intelligence (Why, What, and How);*IEEE Intelligent Systems*, vol. 38, no. 3, pp. 56-62, May-June 2023, doi: 10.1109/MIS.2023.3268724 https://ieeexplore.ieee.org/abstract/document/10148662

Sharma, A., (2023), New 'assembly theory' unifies physics and biology to explain evolution and complexity https://doi.org/10.1038/s41586-023-06600-9 2023

Signorelli, C.M., Szczotka, J., Prentner, R. (2021). Explanatory profiles of models of consciousness - towards a systematic classification. Neurosci. Conscious. 2021, niab021. doi: 10.1093/nc/niab021

Skrbina, D., Panpsychism in the West, Cambridge, MA, MIT Press, 2005

Smith, Brian Cantwell, On the Origins of Objects, Bradford Books, 1996

Snaprud, P. (2018). The consciousness wager.

New Sci. 238, 28-31. https://consc.net/consciousnesswager.pdf

Spencer-Brown, G., Laws of Form, Allen & Unwin, 1969.

Stanford Encyclopedia of Philosophy, 1 Artificial Intelligence, Artificial Intelligence (Stanford Encyclopedia of Philosophy)

Stanford Encyclopedia of Philosophy, 2 Models in Science. https://plato.stanford.edu/entries/models-science/

Stanford Encyclopedia of Philosophy, 2022. 3 Free Will, https://plato.stanford.edu/entries/freewill/

Stanford Encyclopedia of Philosophy. (2020d). Emergent properties. https://plato.stanford.edu/entries/properties-emergent/.

Stanford Encyclopedia of Philosophy, (2022) Panpsychism, https://plato.stanford.edu/entries/panpsychism/#:~:text=Panpsychism%20is%20the%20view%20that,ubiquitous%20in%20the%20natural%20world.

Stapp, H.P.. "A quantum theory of the mind-brain interface". In Mind, Matter, and Quantum Mechanics, Springer, Berlin, pp. 145-172, 1993

Strawson, G. 2006. "Realistic Monism: Why Physicalism Entails Panpsychism." Journal of Consciousness Studies, 13(10-11).

Strawson, G. et al. 2006. Consciousness and its Place in

Nature: Does Physicalism entail Panpsychism? Exeter, UK: Imprint Academic.

Suchow, J.W., Bourgin, D.D., and Griffiths, T.L. (2017). Evolution in mind: evolutionary dynamics, cognitive processes, and Bayesian inference. Trends Cogn. Sci. 21, 522-530. doi: 10.1016/j.tics.2017.04.005

Sunak, R.,2023. Prime Minister's speech on AI; Prime Minister's speech on AI: 26 October 2023 - GOV.UK (www.gov.uk)

Tarski, A., Mostowski, A., Robinson, R.M., 2010Undecidable Theories: Studies in Logic and the Foundation of Mathematics Dover Publications

Tegmark, M., the mathematical universe, arx-iv:0704.0646 , DOI10.1007/s10701-007-9186-9

Tong, D., Quantum Fields: The Real Building Blocks of the Universe, 2017. h https://www.youtube.com/watch?v=zNVQfWC_evg

Tong, D., Q & A:The Real Building Blocks of the Universe, 2017

Tong, D. The Quantum Hall Effect, 2016. https://arxiv.org/pdf/1606.06687.pdf

Tononi, G., Massimini, M., Sizing up Consciousness: Towards an objective measure of the capacity for experience, 2018.

Tononi, G., Steven Laureys, Olivia Gosseries, The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology, 2015

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., et al. (2017). Attention is all you need. Adv. Neural Inf. Process. Syst. 30.

Vimal, R., Biophysics Of Consciousness: A Foundational Approach, editors Poznanski Roman R, Tuszynski,Jack A, Feinberg, Todd E., World Scientific, 2016

Walsh, T. (2017). The Singularity May Never Be Near. AI Magazine, 38(3),58-62. https://doi.org/10.1609/aimag.v38i3.2702

Wikipedia. (2023au). Existential risk from artificial general intelligence. https://en.wikipedia.org/wiki/Existential_risk_from_artificial_general_intelligence.

Wikipedia. (2023ca). Man a machine. https://en.wikipedia.org/wiki/Man_a_Machine.

Wikipedia. (2023i). API. https://en.wikipedia.org/wiki/API.

Wikipedia. (2023x). Big history. https://en.wikipedia.org/wiki/Big_History.

Wikipedia, (2024), Panpsychism, https://en.wikipedia.org/wiki/Panpsychism

Wong, M.L. et al. (2023); On the roles of function and selection in evolving systems. Proc. Natl. Acad. Sci. U.S.A. 120, e2310223120 (2023). On the roles of function and selection in evolving systems | PNAS

Wong, M.L., et al, (2024). Reply to Root-Bernstein: Increasing complexity allows for the pervasiveness of low-complexity entities and is not anthropocentric

Proceedings of the National Academy of Sciences,2024-08-20 | Journal article

Reply to Root-Bernstein: Increasing complexity allows for the pervasiveness of low-complexity entities and is not anthropocentric | PNAS

Wolfram, S.,(2024);What's really going on in machine learning,

Wyatt, S. (2008). Technological determinism is dead; long live technological determinism. Handb. Sci. Technol. Stud. 3, 165-180.

Zhang, W-R, Peace, K.E. (2014). Causality Is Logically Definable—Toward an Equilibrium-Based Computing Paradigm of Quantum Agent and Quantum Intelligence (QAQI) (Survey and Research), Journal of Quantum Information, Science, Vol.4 No.4, 2014

Zhang, W. (2023). If AI machine cannot think, can QI machine think?—from negative numbers to quantum intelligence for mind-light-matter unity. Quantum Machine Intelligence, 5(1), 14. https://doi.org/10.1007/s42484-023-00104-5

Zhang, W. (2021). Ground-0 Axioms vs. First Principles and Second Law: From the Geometry of Light and Logic of Photon to Mind-Light-Matter Unity-AI&QI. IEEE/CAA Journal of Automatica Sinica, 8(3), 534-553. https://doi.org/10.1109/JAS.2021.1003868

STEPHEN B. RIPLEY President and Research Director, Futures & Forecasts, Inc., Vancouver, Canada, 1975-Yale University, New Haven, Connecticut, USA (B.A., philosophy, 1962), University of Alberta, Calgary, Canada, post-graduate, philosophy, 1964-65). Research interests: Artificial Intelligence, Computer Sciences, Biochemistry, Biology, Earth Sciences, Big History