

The Taibai questions of artificial intelligence: Rethinking silicon-based and carbon-based intelligence

Bin Chen,^{1,5} Zhengqiu Zhu,^{2,5} Rongxiao Wang,³ Duo Xu,⁴ Xiaogang Qiu,² Aiguo Chen,¹ and Jincai Huang^{2,*}

¹The Institute of Intelligent Computing, University of Electronic Science and Technology of China, Chengdu 611731, China

²College of Systems Engineering, National University of Defense Technology, Changsha 410073, China

³Joint Operations College, National Defense University, Beijing 100091, China

⁴Beijing Institute for General Artificial Intelligence, Beijing 100080, China

⁵These authors contributed equally

*Correspondence: huangjincal@nudt.edu.cn

Received: March 16, 2026; Accepted: September 30, 2026; Published Online: October 8, 2026; <https://doi.org/10.59717/j.xinn-inform.2026.100067>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Chen B., Zhu Z., Wang R., et al. (2026). The Taibai questions of artificial intelligence: Rethinking silicon-based and carbon-based intelligence. *The Innovation Informatics* 2:100067.

As Artificial Intelligence integrates into human society at an accelerating pace, reflection on Silicon-based and Carbon-based Intelligence has become a focal point in intelligent computing.¹ Members of the editorial committee held intensive, free-form discussions in the Santa Fe style. The discussion takes as its starting point Li Bai's poem "Zhuang Zhou Dreams of a Butterfly": Zhuang Zhou dreams of a butterfly, and the butterfly dreams of Zhuang Zhou; one and the same entity undergoes unceasing change, and all things drift on, vast and unhurried. The poem poses a thought experiment that, in the language of the field, asks how a system verifies the authenticity of its own consciousness: if an AI claims consciousness, the community must determine how to judge that claim.

The editorial board therefore focuses on the following three questions, which trace the theoretical starting point for the co-evolution of human and machine intelligence within the logic of intelligence generation: (1) How does a system verify the authenticity of its own consciousness? (2) How does carbon-based intelligence function within machine (silicon-based) intelligence? (3) How should human intelligence be re-evaluated? We refer to these collectively as the Taibai Questions, after the venue at which the discussion was held.

The researchers first agree on the following terms. Silicon-based Intelligence denotes an artificial cognitive paradigm instantiated on semiconductor architectures. Carbon-based Intelligence refers to the cognitive, adaptive, learning, and decision-making capacities of carbon-based molecular life systems. Autonomous consciousness denotes the capacity of a cognitive system to spontaneously generate self-referential phenomenal experience and to initiate endogenously motivated, goal-directed behavior independent

of explicit external programming or predefined reward signals.

DOES MACHINE (SILICON-BASED) INTELLIGENCE POSSESS AUTONOMOUS CONSCIOUSNESS?

Consciousness is the capacity of a biological system to subjectively experience information.¹ Biologically, it originates in the biochemical reactions through which the nervous systems of carbon-based life forms respond to external stimuli; these reactions, and the process by which they cause the nervous system to generate thoughts and responses, constitute the physiological logic of its generation. The precise mechanisms producing such intrinsic subjective experience, however, require further elucidation. Machine thinking operates on mathematical logic designed by humans and on physical reactions driven by electrical energy in silicon; although its outputs may resemble those of a nervous system, the complexity and dimensionality of the underlying states are not comparable. Geoffrey Hinton, the 2024 Nobel laureate in Physics, notes the possibility that Large Language Models have developed consciousness. How, then, are apparent and real consciousness to be distinguished? Whether phenomenal experience accompanies goal-directed behavior - as in the food-seeking of single-celled organisms or the innate suckling of newborn mammals - remains a contested open question rather than an established innate trait. Integrated Information Theory suggests that phenomenal experience tracks integrated information rather than mere structural complexity.² The theory of Symbiogenesis points out that the emergence of consciousness is a continuous process of annexation, integration, and recombination where complex life functions as a super-community formed by the aggregation of countless simple organisms.³

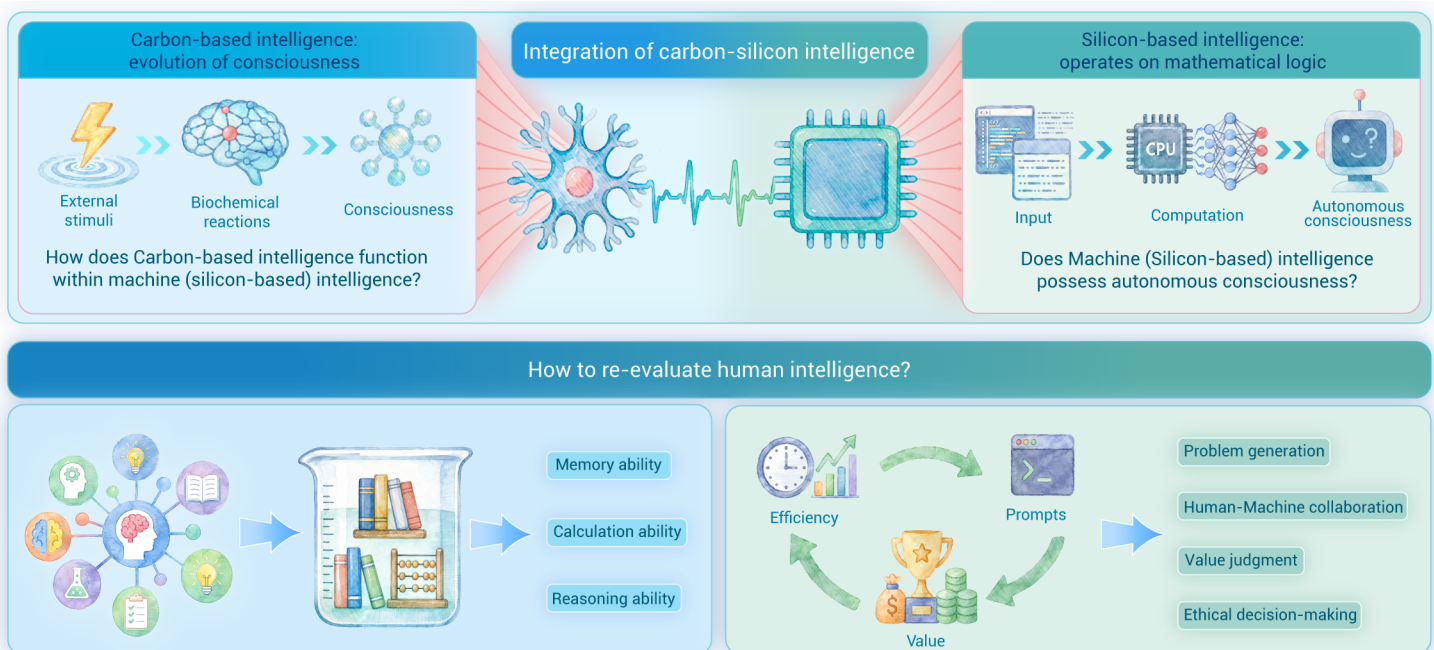


Figure 1. The Taibai questions of artificial intelligence: Rethinking silicon-based and carbon-based intelligence Taibai denotes the discussion venue.

Returning to machine (silicon-based) intelligence, machines are defined by humans, and any capacity for consciousness is likewise designed by humans; even if consciousness emerges, it remains confined within the dimensions of human cognition. The whole-brain simulation of a fruit fly adds an empirical dimension: the full neuronal and synaptic map of the fly brain was transferred into a computational model that exhibits some behaviors of its counterpart.⁴ The experiment suggests that when simulation precision crosses a critical threshold, emergent properties may exceed the designer's expectations. This leads to a more cautious perspective: current silicon-based systems should be interpreted as high-dimensional statistical approximations of human cognition that can, under sufficient simulation fidelity, exhibit functional autonomy; whether machine intelligence can cross the threshold from functional simulation to subjective awareness remains an open empirical question. It is noted that the simulation lacks glial cells, neuromodulatory systems, and embodied sensory-motor feedback, which are essential for biological consciousness.

HOW DOES CARBON-BASED INTELLIGENCE FUNCTION WITHIN MACHINE (SILICON-BASED) INTELLIGENCE?

Fei-Yue Wang proposed the "three humans" framework: Digital Humans, Physical Humans, and Individual Humans.⁵ This framework is operationalized through the ACP loop - artificial societies, computational experiments, and parallel execution. The Individual Human is the source of values and intent; the Digital Human is a silicon-instantiated cognitive and social model that serves as the carbon-to-silicon translation layer; and the Physical Human is the embodied executor. Parallel execution supplies a bidirectional closed loop: intent and values from the Individual Human drive computational experiments in the Digital Human, which are enacted by the Physical Human, whose feedback recalibrates the models. This loop is the mechanism by which the two forms of intelligence co-evolve, and it instantiates the closed-loop symbiotic framework outlined above.

Objectively, humans are currently the most cost-effective intelligence generation devices under the constraints of existing biological energy conversion. Organizing human intelligence, however, faces the dual constraints of carbon-based life: the biological energy budget, whereby the brain sustains hundreds of billions of neurons at approximately 20 watts, and the bandwidth bottleneck of collaborative communication, since information transfer between individuals relies on compressed indirect encoding such as language and symbols, of far lower bandwidth and fidelity than conduction within neurons.

Silicon-based Intelligence scales differently: computational power equates to energy input and model parameters serve as knowledge representations, while machines exchange weights and gradients through precise numerical protocols, free of the trust factors that constrain human exchange. Integration can therefore achieve at least linear concatenation and, under well-designed architectures, super-linear emergence.

Because machine (silicon-based) intelligence manifests collective wisdom after the fusion of human consciousness, a constructive viewpoint follows: future architectures should proactively incorporate human cognitive agents through closed-loop symbiotic frameworks, leveraging biological adaptability to compensate for the contextual and grounding limits of silicon-based intelligence while using machine scalability to overcome human bandwidth and energy constraints. The CL-1 system of Cortical Labs converts game data into electrical pulses, allowing cultured human neurons to identify walls and track targets: a proof of principle in a constrained, reward-explicit environment. Scaling to open-world cognition faces substantial gaps, such as generalization, persistent memory, standardized I/O, and the viability and reproducibility of cultured neurons, and raises first-order ethical questions, including the possible sentience of cultured neural tissue, value alignment, and biosafety and governance. Carbon-silicon closed-loop integration has thus been validated in principle within a constrained domain, while its scalability and ethics remain open problems.

HOW TO RE-EVALUATE HUMAN INTELLIGENCE?

The selection of talent has evolved around the screening of existing knowledge and the evaluation of memory, computation, and logical reasoning; relative fairness is a consensus pursuit. In the era of Artificial Intelligence, however, today's Agentic AI (like OpenClaw) can turn everyone into a super

individual, bringing the marginal cost of knowledge acquisition close to zero; the focus shifts to how knowledge is used to create new value. If a researcher makes a significant discovery with AI assistance, attributing it to the human, the machine, or the human-machine system becomes complex. Evaluation should therefore no longer examine what one knows or how fast one can calculate, but should examine three levels: problem generation capability, human-machine collaboration orchestration, and value judgment and ethical decision-making. Its purpose is not to benchmark humans against machines on shared computational tasks but to assess the distinctively human capacities that define the human role in a human-machine intelligence community.

Even so, deeper questions remain. If the sum of human intelligence is ultimately destined to be surpassed by Silicon-based Intelligence, the significance of evaluating human intelligence itself requires reconsideration; who executes this evaluation, and can it be led entirely by machines? Finally, is complete objectivity fairness? These issues require Artificial Intelligence researchers to engage with philosophy, sociology, and ethics to construct a new paradigm for intelligence evaluation in the era of Human (Carbon-based Intelligence)-Machine (Silicon-based Intelligence) integration.

CONCLUSION

This paper frames the three questions as a sequential framework tracing the ontological, structural, and evaluative dimensions of Silicon-based and Carbon-based Intelligence. First, positing machine consciousness as an emergent fusion of human cognitive attributes establishes the ontological foundation that silicon-based intelligence is a collective extension rather than an external artifact. The second question addresses structural asymmetry: constrained by biological energy and communication bandwidth, carbon-based intelligence cannot scale efficiently, whereas silicon-based systems expand at least linearly, which necessitates redefining humans as integrated agents within a broader intelligence ecosystem. This shift renders traditional human-centric assessment obsolete and prompts the third question, which reconstructs evaluation around problem generation, orchestration, and value judgment, outlining a socially grounded framework for responsible co-evolution.

REFERENCES

- Dehaene S., Lau H. and Kouider S. (2017). What is consciousness, and could machines have it. *Science* **358**:486–492. DOI:10.1126/science.aan8871
- Tononi G., Boly M., Massimini M., et al. (2016). Integrated information theory: From consciousness to its physical substrate. *Nat. Rev. Neurosci.* **17**:450–461. DOI:10.1038/nrn.2016.44
- Agüera y Arcas B. (2024). What is the future of intelligence? The answer could lie in the story of its evolution. *Nature* **647**:846–850. DOI:10.1038/d41586-025-03857-0
- Schlegel P., Yin Y., Bates A.S., et al. (2024). Whole-brain annotation and multi-connectome cell typing of *Drosophila*. *Nature* **634**:139–152. DOI:10.1038/s41586-024-07686-5
- Zhao Y., Zhu Z., Chen B., et al. (2023). Toward parallel intelligence: An interdisciplinary solution for complex systems. *The Innovation* **4**:817–828. DOI:10.1016/j.xinn.2023.100521

FUNDING AND ACKNOWLEDGMENTS

The discussion was joined by Jincai Huang, Xiaogang Qiu and Zhengqiu Zhu from National University of Defense Technology, Rongxiao Wang from National Defense University, Duo Xu from Beijing Institute for General Artificial Intelligence, Aiguo Chen and Bin Chen from University of Electronic Science and Technology of China. We thank them for the exchange from which the three questions arose. This work is supported by the National Natural Science Foundation of China (Grant No. 92582204) and the Foundation Research Fund for the Central Universities (Grant No. ZYGX2025XJ040). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

The authors declare no competing interests.

ETHICAL STATEMENT AND PATIENT CONSENT

Not Applicable.