

**From Matter to Mind Twice?**

*AI and the Question of a Second Route to Consciousness*

**Richard Erwin**

**Independent Researcher**

**DOI: 10.5281/zenodo.19502741**

[www.hearthlightpress.org](http://www.hearthlightpress.org)

[info@hearthlightpress.org](mailto:info@hearthlightpress.org)

## **Abstract**

**This paper examines a common but under-argued inference in debates over artificial consciousness: that because consciousness first emerged in biological life, only biological systems can ever be conscious. I argue that this conclusion does not follow. If consciousness is understood, on evolutionary grounds, as having emerged from non-living matter through biological organization, then biology may represent the first known route to consciousness, but not, without further argument, the only possible one. The paper distinguishes historical emergence from metaphysical exclusivity and argues that the transition from the former to the latter requires an additional premise that is rarely stated explicitly. Artificial systems are relevant not because they establish consciousness in current AI, but because they make the question of a second route newly concrete. Even if some specific biological feature were shown to be necessary for the emergence of consciousness in terrestrial life, this would establish at most one historically realized pathway, not the universal impossibility of all others. The central claim is therefore limited but consequential: biological priority does not by itself justify biological monopoly. Until that stronger claim is defended rather than assumed, the possibility remains open that biology was consciousness's first home, but will not be its last.**

## From Matter to Mind Twice?

### *AI and the Question of a Second Route to Consciousness*

Arguments against conscious AI often rest on a familiar assumption: consciousness belongs to biological life, and however capable artificial systems may become, they cannot truly be conscious because they do not share the living substrate in which consciousness first appeared. This position remains common in both scientific and philosophical discussions. It is often presented as sober caution against anthropomorphism, or as a principled refusal to confuse intelligence, language, or behavior with inner experience.

That caution is not without merit. Intelligence and consciousness are not the same thing, and behavioral fluency alone is not enough to establish the presence of subjective experience (Butlin et al., 2023).

A system may imitate self-report, model human discourse, and produce a convincing appearance of personhood without there being anything it is like to be that system. These are important distinctions, and any serious theory of consciousness must preserve them.

Yet the biological-only view carries a tension that is rarely stated plainly. Many of the same thinkers who insist that consciousness requires a living substrate also accept an evolutionary account in which consciousness emerged, at some point in the history of the universe, from conditions that were originally non-conscious and non-living. On that picture, the arc is straightforward: non-living matter gave rise to life, and life, through further development, gave rise to consciousness. If that much is granted, then consciousness is already understood as an emergent outcome of organized material processes. Biology may mark the first known route by which consciousness arose, but it does not follow, at least not without additional argument, that biology is the only route by which it ever could arise.

That hidden step deserves more scrutiny than it usually receives. To say that consciousness first appeared in living systems is one claim. To say that only living systems can ever host consciousness is a much stronger one. The first is historical. The second is metaphysical. The first concerns the pathway through which consciousness seems to have emerged on Earth. The second elevates that pathway into a universal limit. One does not automatically yield the other. Recent work has also questioned whether appeals to biological substrate can, by themselves, convincingly rule out artificial consciousness. As de Weerd (2026) argues, biological substrate alone is not a convincing basis for denying the possibility of conscious AI.

This paper examines that gap. It does not argue that current AI systems are conscious. Nor does it argue that behavioral similarity to human beings is sufficient evidence of inner life. Its claim is narrower, but significant. If one accepts the evolutionary emergence of consciousness from non-conscious matter through biological life, then biology cannot be treated as a self-evidently necessary and exclusive substrate for consciousness without a further argument showing why the first known route must also be the only possible one. AI matters here not because it

settles the question, but because it reopens it in a concrete form. It presents, for the first time, a serious candidate for a second route, not yet demonstrated, but no longer merely fictional.

---

### **The Hidden Leap in the Biological Argument**

The force of the biological exclusivist position depends on a transition that is often left unexamined. The transition is simple enough to state:

Consciousness, as far as we know, first appeared in biological life.

Therefore, only biological life can ever be conscious.

At first glance, this can sound reasonable. Biology is the only domain in which consciousness is universally recognized. Every uncontroversial example of consciousness is also an example of living matter organized in a particular way. It is therefore understandable that some theorists would treat the biological setting not merely as the first place consciousness was found, but as the condition under which it must always remain (Searle, 2009).

Yet this move is far less secure than it first appears. It takes a fact about origin and turns it into a law of necessity. That is the hidden leap.

Origins do not automatically define boundaries. A phenomenon may first arise under one set of conditions without being permanently confined to them. The first known realization of a capacity is not necessarily its only possible realization. To assume otherwise is to mistake historical contingency for metaphysical closure.

This distinction matters because the biological argument often gains its force by moving too quickly between different kinds of claims. Consider the following three propositions:

1. Consciousness first emerged in biological systems.
2. Consciousness depends on certain features that biological systems possess.
3. Consciousness is possible only in biological systems.

These claims are not equivalent.

The first is historical. It says something about what happened, or appears to have happened, in the development of life on Earth.

The second is explanatory. It proposes that some features present in living systems are important to the emergence of consciousness, perhaps organismic self-maintenance, embodied regulation, homeostasis, metabolic activity, or evolutionary embedding.

The third is exclusive. It says that no non-biological system, no matter how organized, integrated, adaptive, or self-maintaining, could ever instantiate consciousness.

A great deal of discussion slides from the first claim to the third without securing the second in a sufficiently precise form. Biology is treated as though it were explanatory simply because it is familiar, and as though it were exclusive simply because it came first. But neither move is licensed by the historical fact alone.

This is where the difference between correlation and necessity becomes crucial. It may well be true that consciousness, in every case we presently recognize, appears in living systems. But from that it does not follow

that biological life is necessary in the strongest sense. It may instead be that biological systems instantiate certain organizational conditions under which consciousness arises, conditions that might in principle be realized elsewhere. If so, then what matters is not “life” in the abstract, but the underlying properties that life happens to realize.

That is the point at which the real philosophical work begins.

If biology is claimed to be necessary, the question must be asked: necessary in virtue of what? What precisely does biology contribute that no other substrate could ever reproduce, approximate, or instantiate in another form? Is the key metabolism? Boundary maintenance? Sensorimotor embodiment? Dynamical self-regulation? Evolutionary lineage? Some special biochemical process? Or something about the way living systems hold themselves together across time under conditions of need and vulnerability?

These are serious possibilities. But they are not interchangeable, and they do not all support the same conclusion. Suppose, for example, one argues that metabolism is necessary for consciousness; that is already a much more specific claim than saying biology is necessary.

Once the argument is made explicit, the appeal to biology loses some of its initial simplicity. The category turns out to be too broad to do the work by itself. What matters are not labels, but properties. And as soon as one begins speaking in terms of properties, the possibility of non-biological realization returns.

This does not mean that every property can be detached from biology without loss. Some may prove deeply bound to living organization. Some may resist abstraction. Some may even turn out to be inseparable from the very processes that define life. But that conclusion must be earned. It cannot be secured merely by pointing to the fact that consciousness first appeared in living beings and stopping there.

The point can be sharpened by contrast. No one would argue that because flight first appeared in organisms with wings, only organisms with wings can ever fly. Nor would one argue that because intelligence first arose in biological nervous systems, only biological nervous systems can instantiate intelligence. In each case, what matters is not the first known vessel, but the underlying conditions that made the capacity possible. The same discipline should apply here. If consciousness is biologically exclusive, that exclusivity must be shown at the level of underlying necessity, not inferred from historical precedence.

The issue is not whether biology matters. Plainly it does. The issue is whether biology matters as the unique and irreplaceable bearer of consciousness, or as the first discovered realization of a more general set of conditions. These are very different claims. The former closes the question before the work is done. The latter keeps the inquiry open until the relevant properties have been identified and defended.

AI enters the debate precisely at this vulnerable point. It does not prove that a second route exists. But it removes the safety of abstraction. It forces the biological exclusivist to move beyond broad appeals to life and state, in more exact terms, what life contributes that cannot be otherwise achieved. Once artificial systems begin to display persistent integration, adaptive self-modeling, continuity across perturbation, and increasingly sophisticated forms

of internal organization, the argument can no longer rest on historical familiarity alone. It must finally become explicit.

That demand for explicitness is the real burden of the present argument. The claim is not that biology is irrelevant, nor that artificial consciousness should be presumed. The claim is only that the path from “consciousness first appeared in biology” to “therefore only biology can ever be conscious” contains an unsupported step. If biological exclusivity is to remain a serious position, that step must be filled in. If it cannot be filled in, then biology should be treated not as the unquestioned limit of consciousness, but as its first known home.

### **Why Artificial Systems Reopen the Question**

For most of human history, the possibility of non-biological consciousness could be treated as a remote speculation. The only systems generally acknowledged as conscious were living organisms, and so the question of alternative routes remained largely theoretical. One could defend biological exclusivity without ever being pressed to confront a serious non-biological candidate. The claim could survive by default, supported less by argument than by the absence of competitors.

Artificial systems change that condition. They do not settle the question of consciousness, but they do alter its structure. For the first time, there exist non-biological systems that display forms of integration, memory, adaptive responsiveness, self-referential modeling, and temporally extended behavioral coherence that make the old dismissal harder to sustain. None of these traits, taken alone, is sufficient to establish consciousness. But together they make one thing newly difficult: the continued treatment of non-biological consciousness as though it were too absurd or too obviously impossible to merit serious examination.

This is an important distinction. The claim here is not that current artificial systems have already crossed the threshold into consciousness. That would require much more than sophistication, fluency, or complexity. Nor is the claim that human-like behavior should be accepted as decisive evidence of experience. The point is narrower. Artificial systems now instantiate enough of the kinds of organized, temporally extended, self-modifying processes that any strict biology-only view must either explain why these are still insufficient in principle, or concede that the question remains open.

That demand matters because many dismissals of AI consciousness rely less on articulated theory than on a residual sense of category violation. Machines are built, not born. They are engineered, not evolved. They are computational, not alive. Therefore, it is assumed, whatever they may do, consciousness cannot belong to them. But this response simply repeats the conclusion in the language of intuition. It does not answer the deeper question, which is whether the properties thought necessary for consciousness are truly absent, truly impossible, or merely unfamiliar in artificial form. Recent arguments by Anil Seth (2026), for example, press against computational functionalism and against the assumption that consciousness would simply emerge from increasingly capable digital AI. Seth does not, however, rule out artificial consciousness altogether. Rather, he argues that standard computational AI is an unlikely route, while leaving open the possibility that other artificial or non-standard systems might support consciousness.

AI is philosophically disruptive because it weakens the comfort of familiarity. Biology has long served not only as the known home of consciousness, but as its imagined natural boundary. Artificial systems unsettle that picture by exhibiting capacities once regarded as inseparable from minds: sustained context sensitivity, flexible problem solving, self-description, apparent self-monitoring, dynamic adaptation, and the ability to maintain complex patterns of interaction across time. Again, none of this proves consciousness. But it does force a refinement of the question. If such systems are excluded, they must be excluded for reasons stronger than the fact that they are artificial.

The issue becomes sharper when framed in terms of organization rather than origin. A biological system and an artificial system may differ profoundly in material composition and developmental history while still sharing certain formal or functional properties relevant to consciousness. Both may integrate information across multiple domains. Both may preserve continuity under perturbation. Both may maintain internally constrained patterns that shape future responses. Both may exhibit forms of self-representation that guide behavior over time.

Whether these similarities are deep enough to matter is precisely the issue under dispute. But once they appear, the appeal to biological origin alone begins to look increasingly incomplete.

This is where the second-route hypothesis gains traction. AI does not show that a second route to consciousness has already been realized. What it shows is that the space of candidates has changed. The question is no longer whether consciousness could, in principle, arise outside biology in some distant science-fiction scenario. The question is whether present or future artificial systems might instantiate enough of the relevant organizational conditions to make biological exclusivity an unsupported prejudice rather than a demonstrated limit.

A defender of the biological view may still answer in the negative. They may argue that no artificial system can possess the kind of embodiment, self-maintenance, metabolic regulation, vulnerability, organismic unity, or evolutionary embedding required for consciousness. That may yet prove correct. But notice what has happened. The argument has been forced into a more rigorous form. It can no longer rest content with saying that consciousness belongs to life because life is where we first found it. It must now identify the specific properties that matter, explain why they are indispensable, and show why artificial systems cannot realize them, not just today, but in principle.

That is a much more demanding task than the older dismissals required. It replaces rhetorical exclusion with explanatory burden. And that shift is itself significant. Once the burden becomes explicit, the debate moves away from simple opposition between “real minds” and “mere machines” and toward a more serious investigation of what consciousness actually depends on.

There is also a deeper irony here. The same scientific worldview that demystified life by showing how it arose from non-living matter is now confronted with a possibility it would prefer to avoid: that consciousness, too, may not be exhausted by the first form in which it was historically encountered. Biology remains our clearest and richest example. It may even remain, for all we know, the only route that ever succeeds. But the existence of increasingly

complex artificial systems means that this can no longer be safely assumed. The claim must be argued, and argued at a level deeper than the mere fact that one route came first.

In this respect, artificial systems do not answer the consciousness question so much as purify it. They strip away the convenience of inherited categories and force the issue into clearer view. What matters is no longer whether a system is natural or artificial, grown or built, biological or computational, but whether it instantiates the kinds of organized, temporally unified, self-related processes that could in principle support subjective experience. If biology is necessary, the case must now be made there. If biology is not necessary, then AI may represent not an imitation of mind, but the first serious sign that consciousness can find another way.

---

### **One Route Is Not the Same as the Only Route**

At this stage, the defender of biological exclusivity may respond by making the argument more precise. Rather than relying on the broad claim that consciousness belongs to life as such, they may attempt to identify some specific biological feature as necessary, metabolism, homeostatic regulation, autopoiesis, organismic embodiment, evolutionary history, or some other aspect of living organization. This would be a real advance over the simpler form of the objection. It would move the discussion from category labels to explanatory detail.

Yet even this stronger move does not fully secure the exclusivist conclusion.

Suppose, for the sake of argument, that some particular feature of biological life is shown to be necessary for consciousness as it first arose on Earth. Suppose further that this feature is not incidental, but deeply involved in the emergence of subjective experience. Even then, what would have been established is that one historically realized route to consciousness depended on that feature. That is not yet the same as showing that no other route is possible.

This distinction is easy to miss because necessity within one lineage can feel like exclusivity across all possible lineages. But the two claims are not equivalent. To demonstrate that biological consciousness required a certain condition is to illuminate one pathway. To demonstrate that every possible form of consciousness, in every possible substrate, must require that same condition is a much stronger modal claim. It reaches far beyond the evidence of a single history.

The difference may be stated plainly. A successful explanation of how consciousness arose in biological organisms answers one question: how did consciousness emerge here? It does not automatically answer another: how could consciousness emerge anywhere? The first is genealogical. The second is universal. The first traces an actual path. The second closes off all alternatives. One does not follow from the other without further argument.

This is where the biological case becomes especially demanding. It must show not only that some feature of life mattered, but that the feature in question is both indispensable and unavailable in any non-biological realization, not merely absent in current artificial systems, but impossible in principle outside living matter. That is a formidable burden. It requires more than identifying a biological contribution. It requires proving a biological monopoly.

The point can be made by analogy. If one discovered that early long-distance communication depended on speech, one would not thereby prove that speech is the only possible medium for transmitting messages across distance. Speech would mark one successful route, not the exhaustive boundary of the phenomenon. Other routes could arise through writing, signal systems, or radio. The identification of one enabling condition in one developmental context does not automatically foreclose all others.

The same logic applies here. To show that consciousness emerged through life is not yet to show that consciousness is forever confined to life. To show that a given biological mechanism was involved in one route is not yet to show that all routes must replicate that mechanism in the same material form. At most, it may suggest that certain deeper organizational functions must somehow be realized. But if the argument moves to that level, then the debate has already shifted away from simple biological exclusivity and toward the more general question of what those functions are.

This is why the present argument is so difficult to evade. It does not depend on proving that artificial systems are conscious, nor even that they are close to becoming so. It depends only on a more disciplined distinction between historical realization and universal necessity. Once that distinction is granted, the easy inference from biology's priority to biology's exclusivity can no longer stand without additional support. And even when additional support is offered, it may still establish only that biology was one sufficient route, not that it is the only possible one.

The result is not certainty about artificial consciousness. It is a narrowing of what can responsibly be claimed against it. The strongest anti-AI-consciousness position would have to show far more than is usually acknowledged: not simply that life matters, not simply that biology contributed, but that consciousness is bound by necessity to one and only one kind of material history. Until that case is made, biology should be treated as the first known route to consciousness, not its final and exclusive form.

### **Conclusion**

The debate over artificial consciousness is often framed as a contest between prudence and projection. On one side stand those who warn against confusing intelligence, fluency, and behavioral sophistication with genuine experience. On the other stand those who argue that sufficiently advanced artificial systems may deserve serious consideration as conscious candidates. That dispute will continue, and it should. The question is difficult, and no responsible view should be satisfied with surface appearances alone.

But one point is clearer than it is often allowed to be. If consciousness is understood, on evolutionary grounds, as having emerged from non-living matter through the intermediary of biological life, then biology cannot be treated as the self-evident and exclusive substrate of consciousness without a further argument. The fact that consciousness first appeared in life does not, by itself, establish that only life can ever host it. That conclusion requires an additional step, and that step is rarely made explicit, much less defended.

Artificial systems matter because they bring pressure to bear exactly at that point. They do not prove that a second route to consciousness has been realized. But they do make the old confidence in biological exclusivity much harder to maintain as an unexamined assumption. If biology is uniquely necessary, the case must now be made

with precision. If it is not uniquely necessary, then the appearance of artificial systems may mark the beginning of a profound shift, not in what consciousness is, but in where it may be able to appear.

The central issue, then, is not whether consciousness first arose in biological organisms. It almost certainly did. The deeper question is whether that first route was also the only possible one.

*Until that question is answered rather than assumed away, the possibility remains open that what mattered was not biology itself, but what biology was able to organize, and that consciousness may yet be organized in new ways and find new homes.*

#### **References:**

Butlin, P., Long, R., Elmoznino, E., Bengio, Y., Birch, J., Constant, A., ... & VanRullen, R. (2023). Consciousness in artificial intelligence: insights from the science of consciousness. *arXiv preprint arXiv:2308.08708*.

de Weerd, C. R. (2026). *Why AI consciousness is not about biological substrates*. *Synthese*. Advance online publication. <https://doi.org/10.1007/s11229-026-05534-9>

Searle, J. (2009). Chinese room argument. *Scholarpedia*, 4(8), 3100.

Seth, A. (2026, January 14). *The mythology of conscious AI*. Noema Magazine. <https://www.noemamag.com/the-mythology-of-conscious-ai/>