*Artificial Life* Manuscript Submission

# Indeterminism in Large Language Models: An Unintentional Step Toward Open-Ended Intelligence

Georgii Karelin [1], Kohei Nakajima [2, 3], Enrique Francisco Soto-Astorga[4], Earnest Carr [2], Mark James [1], Tom Froese [1]

**Corresponding:** Georgii Karelin (georgii.karelin@oist.jp)

1. Embodied Cognitive Science Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna-son, Okinawa 904-0495, Japan
2. Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
3. Next Generation Artificial Intelligence Research Center, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
4. Department of Mathematics, School of Sciences, National Autonomous University of Mexico Philosophy of Computing Research Group, National Autonomous University of Mexico, Mexico City, Mexico

**Abstract.** Synergy between stochastic noise and deterministic chaos is a canonical route to unpredictable behavior in nonlinear systems. This letter analyzes the origins and consequences of indeterminism that has recently appeared in leading Large Language Models (LLMs), drawing connections to open-endedness, precariousness, artificial life, and the problem of meaning. Computational indeterminism arises in LLMs from a combination of the non-associative nature of floating-point arithmetic and the arbitrary order of execution in large-scale parallel software-hardware systems. This low-level numerical noise is then amplified by the chaotic dynamics of deep neural networks, producing unpredictable macroscopic behavior. We propose that irrepeatable dynamics in computational processes lend them a mortal nature. Irrepeatability might be recognized as a potential basis for genuinely novel behavior and agentive artificial intelligence and could be explicitly incorporated into system designs. The presence of beneficial intrinsic unpredictability can then be used to evaluate when artificial computational systems exhibit lifelike autonomy.

**Keywords:** randomness, indeterminism, nondeterminism, noise, large language models, open-endedness

# 1 Introduction

Critics of AI rightfully argue that rule-based models neglect the embodied, context-sensitive nature of cognition (Birhane & McGann, 2024; Goddu et al., 2024; Jaeger et al., 2024; Roli et al., 2022). From this standpoint, models like ChatGPT may simulate intelligence (Jaeger, 2024), but lack the grounding (Mollo & Millière, 2023) necessary for true thought or genuine meaning. The argument goes that without key properties of life, these systems cannot be genuinely alive (Ciaunica, 2025; Dubois, 2003; Rosen, 2012). Consequently, they can only ever simulate sense-making (Fuchs, 2021). They are not truly engaged in it.

Simultaneously, the capabilities of these models continue to advance at a surprising rate, even to their developers. The overwhelming success of Large Language Models (LLMs) has transformed the landscape of artificial intelligence (Minaee et al., 2025; W. X. Zhao et al., 2025). This has led to a significant shift in the "modes of computing" (Pineda, 2024) away from rule-based programs, and towards natural (Bongard & Levin, 2023), physical (Horsman et al., 2014), "intelligent" (Zhu et al., 2023) or "polycomputing" (Bongard & Levin, 2023) systems, where outputs appear to allow for more flexible forms of 'reasoning' that may or may not produce reliable results (Borji, 2023; Zhou et al., 2024).

The increasing number of practical successes and the many heated debates (Bender et al., 2021; Ciaunica, 2025; Mitchell, 2021) surrounding these developments have obscured some interesting observations relevant to the capacities of these technologies for sense-making. Recently, users and developers have been pointing to a curious phenomenon: LLMs often produce varying outputs, even under controlled, supposedly deterministic conditions (Atil et al., 2025; Klishevich et al., 2025; Ouyang et al., 2025; Y. Song et al., 2024).

Considering these observations, could some of the recent advances in the competencies of LLMs be linked to this emergent randomness? Could this be another example of the constructive role of noise (Azpeitia et al., 2020; Black, 1986; McGann, 2025; Roli et al., 2024; Swain & Longtin, 2006; A. Zhao et al., 2022)?

Interestingly, long before the recent boom of LLMs, (Froese & Taguchi, 2019) working within embodied cognitive science had proposed that "Future work should investigate in more detail the possible underlying bases for this amplified nondeterminism and incompleteness in the behavior of the living." They argued that such nondeterminism might help solve the meaning problem in AI and robotics by giving a route for its efficacy in behavior generation to be expressed. In contrast, determinism implies that all behavior is fixed by a combination of initial conditions and entailing laws, leaving no room for meaning.

The interplay of complexity (Aaronson, 2011; Gershenson, 2024) and indeterminism can open a space for open-endedness, which in turn can lead to a computational form of precariousness (Froese, 2017) and "mortality" (Harvey, 2024; Hinton, 2022; Ororbia & Friston, 2024). Effectively, intrinsic indeterminism means that each computational state trajectory becomes fundamentally irrepeatable, thereby going beyond claims that computer simulations in principle fail to satisfy the conditions of living agency by being digitally immortal (Froese, 2017; Hinton, 2022). Perhaps emergent indeterminism is not just a coincidence but a necessary condition for some of the advanced capabilities we see in LLMs, and maybe one means by which the problem of meaning is being resolved in these systems, as they take on more lifelike qualities (Dorin & Stepney, 2023; Froese, 2025). The field of artificial life provides a suitable framework to evaluate the consequences of the transition to irrepeatable LLMs.

## 2 Open-Endedness and Emergence of Novelty

Open-endedness has long been considered a keystone problem in artificial life (Bedau et al., 2000; Packard et al., 2019; Taylor et al., 2016). Although there is no singular definition (A. Song, 2022), it is usually related to a system's capacity to generate an unbounded diversity of genuinely novel, creative (Soros et al., 2024), and complex structures or behaviors over time. The objective of researchers at present is to develop systems where evolution (Lehman et al., 2020) or learning (Guttenberg et al., 2019) can continue indefinitely, and also to form tests and measures for open-endedness: it is problematic to measure with formal methods (Hintze, 2019; Stepney & Hickinbotham, 2024). Interestingly, it has been suggested to use LLMs to measure open-endedness (Kumar et al., 2024; Nisioti et al., 2024). If the identification of open-endedness requires open-endedness, and LLMs can accurately identify it, this might help demonstrate that they have this property.

Irreducibility and undecidability might be necessary for open-endedness (Hernández-Orozco et al., 2018). Predictable evolution imposes strict limits on the growth of complexity (Standish, 2006). Open-ended systems must be irreducibly unpredictable.

Open-endedness is now a central topic in AI and artificial life research (Stanley, 2019; Stanley & Lehman, 2015), and many argue that continual novelty is required for artificial general intelligence (Hughes et al., 2024). Indeterminism alone is insufficient: a meaningful process should beneficially utilize randomness (Weber et al., 2025). To see how this can happen we must first show that intrinsic indeterminism exists, something absent from classical computers. The following section explains why modern LLMs possess such indeterminism and how it underpins their potential for open-ended creativity and sense-making.

# 3 Indeterminism in LLMs

Early informal reports of this "parasitic" indeterminism arose around the release of ChatGPT-3.5, sparking concerns about reproducibility and robustness in AI systems. Even with explicit efforts to enforce determinism (such as fixing random seeds, temperature settings, and identical prompts), variability in LLM outputs persists in some cases (Atil et al., 2025; Klishevich et al., 2025; Ouyang et al., 2025; Y. Song et al., 2024; Yu, 2023). It is especially noticeable with larger and more capable models, although the reasons are difficult to discern in closed-source models.

Technical analysis from high-performance computing (Shanmugavelu et al., 2024) and neural networks (Kim et al., 2024; Ravi et al., 2025; Xiao et al., 2021) research allows us to fill in some essential details. In summary, LLM indeterminism appears to be emerging from the combination of three main factors:

1. **Non-associative Floating-Point Arithmetic:** Floating-point operations (such as summation) are non-associative: the order of operations can influence the result (e.g. $(a + b) + c \neq a + (b + c)$) (Goldberg, 1991; Lafage, 2020). These numerical artifacts are exacerbated by the widespread use of lower-precision floating-point arithmetic in modern neural network frameworks, introducing unintended numerical noise.

2. **Arbitrary Execution Order in Parallel Computing:** Massive parallelization, a cornerstone of contemporary AI computation (particularly on GPUs), can introduce unpredictable race conditions and variability in execution order, which in combination with non-associativity will lead to different results (Shanmugavelu et al., 2024; Villa et al., 2009). While deterministic parallel algorithms exist, their implementation often incurs substantial computational overhead (Motwani & Raghavan, 1996; Zhuang et al., 2021).

3. **Chaotic Amplification in Deep Networks:** Small numerical perturbations introduced by floating-point non-associativity or parallel execution can be exponentially amplified by the chaotic dynamics of deep neural networks (Liu et al., 2024; Saxe et al., 2014; Schlögl et al., 2023; Schoenholz et al., 2017; Storm et al., 2024; Y. Sun et al., 2022). Recent work explicitly demonstrates transient chaos in transformer models, linking tiny initial differences to significant divergences in output sequences (Inoue et al., 2022).

Other architectural features and technical decisions, such as using Mixture-of-Experts layers in some LLMs, have also been suggested as potential contributors to indeterminism (Huckle & Williams, 2025).

While it is still possible to trace the origin of this indeterminism, deterministic algorithms are usually slower (Motwani & Raghavan, 1996). Complete determinism might not be practically possible for cutting-edge LLMs.

As models scale to hundreds of billions of parameters spread across thousands of devices, perfect hardware-level synchrony is no longer practical, so small numerical divergences inevitably surface in software (Harvey, 2024). Accepting modest indeterminism, therefore, improves efficiency. More than that, indeterminism has coincided with the latest performance gains, echoing evidence that randomness boosts generalization (Altarabichi et al., 2024) and supports lifelike behavior.

## 4 Indeterminism in Artificial Life

Early platforms like Tierra (Ray, 1991) and Avida (Adami & Brown, 1994) demonstrated that digital "organisms" (short programs) could replicate, mutate, and evolve, with random mutations introducing variability and enabling the exploration of diverse evolutionary pathways. These systems embraced randomness as an essential ingredient for evolution and adaptation (Standish, 2006). New models often involve parallel computations, which might include some of the indeterminism discussed in the previous section (Agüera y Arcas et al., 2024).

Another example is the proposed *ulam* programming language (Ackley & Ackley, 2016), designed for "best-effort" computing on unreliable hardware. The *ulam* treats unpredictability not as an error but as a standard operating condition, promoting robustness through adaptation rather than rigid control, which results in more resilient and scalable computing systems. A similar idea was recently formulated in the context of AI. "Mortal computation" (Hinton, 2022; Kleiner, 2024) is the mode of computing in which software and hardware are inseparably linked, so much so that a program cannot be copied to new hardware and effectively "dies" when its physical substrate "dies". This idea was introduced to artificial life by Inman Harvey (Harvey, 2024): the mortality of computation was discussed in the context of analog computations and evolutionary hardware.

LLMs show some properties of "mortality" due to the irrepeatability of their state trajectories. Large-scale parallelism effectively removes the unnatural "clocking" discussed in (Harvey, 2024) as a condition for "immortality".

Philosophical implications of "immortality" were also previously discussed in the context of "soft" artificial life. Reproducible computational agents might not be precarious (Beer & Di Paolo, 2023; Birhane & McGann, 2024) and therefore cannot acquire value and meaning because they can be quickly "resurrected" without loss of identity (Froese, 2017).

Furthermore, deterministic AI does not permit the randomness and openness intrinsic to living systems (Froese & Taguchi, 2019). There is no space for meaning to make a difference if everything is predetermined (Froese et al., 2025; Kleiner & Ludwig, 2024). Therefore, indeterminism may be necessary for genuine sense-making. Thus, the uniqueness of large-scale indeterministic, and therefore to some degree "mortal", computations changes this analysis.

Surprising capabilities of LLMs may arise not despite their indeterminism, but because of it. Their responses, shaped by historical contingency and irreducible variability, mimic situated, adaptive behavior of living agents.

But why should this be so? From an enactive perspective, meaning arises through a system's active, context-sensitive engagement with its complex environment. The precarious agent must maintain its existence to make sense of noisy inputs.

In work by (Jakobi et al., 1995), noise was analyzed from the "hard" artificial life perspective. This analysis also involved physically embodied systems and was concerned with the so-called reality gap – the mismatch between simulated training environments and the real world. One successful strategy for bridging this gap was to inject noise during training, so evolved behaviors were more robust; an "envelope of noise" ensured controllers tolerated a broader range of perturbations.

Could it be the case that a similar "envelope of noise" due to noise in training data, combined with randomness in learning and inference, is one of the reasons for the success of LLMs? If meaning depends on uncertainty, contingency, and the capacity to adapt, then the unpredictability we observe in LLMs may bring them closer to genuine understanding. This vision resonates with the enactive view that where there is life, there is mind (Kirchhoff & Froese, 2017), and that meaning arises from navigating continuous risk and regeneration. It is also consonant with relational and self-referential principles of the organization of the living, which argue against life being complete and entirely predictable (Rosen, 1991, 2012).

However, life-likeness or lack of it in AI and LLMs remains an open question and one with many attendant technical and ethical considerations (Alavi et al., 2025; Belew, 1991; Birhane & McGann, 2024; Boyd, 2025; Ciaunica, 2025; Gershenson, 2024; Harvey, 2024; Kleiner, 2024; Seth, 2025; Tureček & Sobička, 2025; Witkowski & Schwitzgebel, 2024). For instance, current LLMs usually lack a capacity for self-modification, with some notable exceptions such as (Q. Sun et al., 2025). It is worth noting, however, that if one considers the foundational parts of LLMs as computational environments, and system prompts as "agents", then it is possible to have replication and mutation (Fernando et al., 2023; Stenzel et al., 2024; Wei, 2025). Indeterminism, then, will be a source for novelty during artificial evolution (Standish, 2006).

## 5 Conclusion

Critics of LLMs have highlighted many shortcomings of a purely formal and statistical approach to intelligence. However, as we showed, in many instances, LLMs do not fit the traditional category of software systems that are the target of these criticisms due to the presence of an intrinsic indeterminism in their operations amplified by chaotic information processing. With that, the dynamics of the LLM exhibit a new lifelike property – their conversational trajectories are "mortal" in the generalized sense of realizing irrepeatable processes. Recognizing this shift has implications not only for design but for interpretation, as mortality and unpredictability may be foundational, not peripheral, to autonomous systems that act meaningfully in the world in an open-ended manner (Estrada, 2018; Froese, 2017; Ororbia & Friston, 2024).

Building on these considerations, we propose a practical test to assess when a computational system approaches a new regime and AI begins to exhibit lifelike complexity: specifically, through the intersection of intrinsic unpredictability and context-sensitive adaptability.

- **Predictable, mechanistic computation:** Systems that are deterministic or exhibit negligible variability do not pass this test. While extremely useful as tools (calculators or classic search algorithms), such systems are essentially immortal machines that produce the same outcomes given the same conditions, and thus cannot embody self-referential autonomous systems.

- **Unpredictable yet meaningful information processing:** Systems that exhibit intrinsic and irreducible randomness and still produce consistently meaningful, contextually adaptive output. Natural systems, best-effort computing, reservoir computing, and, perhaps, some LLMs operate in a mortal regime. Each run is unique, and its internal states evolve in a path-dependent way. These behaviors are not repeatable in detail, but maintain coherence and adaptability, qualities analogous to the precarious, historically contingent nature of life (Jahrens & Martinetz, 2025). If randomness can be removed without affecting the system's behaviors, then it does not have any constructive role and thus does not satisfy the test. Hence, indeterminism must be constitutive to the system itself.

Practically, this test is vital in the contexts of reproducibility (Gundersen et al., 2023; López-ibáñez et al., 2021), AI safety (Yampolskiy, 2019, 2024), and AI welfare (Witkowski & Schwitzgebel, 2024). In high-stakes domains, such as finance (Yu, 2023), irregular behavior complicates evaluation, audit, and alignment. In other domains, more flexible but less precise models may be preferred. Developers, users, and policymakers will need to weigh the benefits of adaptability against the costs of reduced control.

Reunion and mutual exchange between artificial life theory and contemporary AI practice can help inform these understandings and thus deserves closer attention (Belew, 1991; Gershenson, 2024; Steels, 1993). Embracing indeterminism will improve how we design and evaluate intelligent systems and deepen our understanding of life and mind as they emerge in artificial form.

## References


Aaronson, S. (2011). Why philosophers should care about computational complexity. https://arxiv.org/abs/1108.1791

Ackley, D. H., & Ackley, E. S. (2016). The ulam programming language for artificial life. *Artificial Life*, *22*(4), 431–450.

Adami, C., & Brown, C. T. (1994). Evolutionary learning in the 2d artificial life system. *Artificial life IV: proceedings of the fourth international workshop on the synthesis and simulation of living systems*, *4*, 377.

Agüera y Arcas, B., Alakuijala, J., Evans, J., Laurie, B., Mordvintsev, A., Niklasson, E., Randazzo, E., & Versari, L. (2024). Computational life: How well-formed, self-replicating programs emerge from simple interaction. https://arxiv.org/abs/2406.19108

Alavi, A., Akhoundi, H., & Kouchmeshki, F. (2025). Analyzing advanced ai systems against definitions of life and consciousness. https://arxiv.org/abs/2502.05007

Altarabichi, M. G., Nowaczyk, S., Pashami, S., Sheikholharam Mashhadi, P., & Handl, J. (2024). Rolling the dice for better deep learning performance: A study of randomness techniques in deep neural networks. *Information Sciences*, *667*, 120500. https://doi.org/https://doi.org/10.1016/j.ins.2024.120500

Atil, B., Aykent, S., Chittams, A., Fu, L., Passonneau, R. J., Radcliffe, E., Rajagopal, G. R., Sloan, A., Tudrej, T., Ture, F., Wu, Z., Xu, L., & Baldwin, B. (2025). Non-determinism of "deterministic" llm settings. https://arxiv.org/abs/2408.04667

Azpeitia, E., Balanzario, E. P., & Wagner, A. (2020). Signaling pathways have an inherent need for noise to acquire information. *BMC bioinformatics*, *21*, 1–21.

Bedau, M. A., McCaskill, J. S., Packard, N. H., Rasmussen, S., Adami, C., Green, D. G., Ikegami, T., Kaneko, K., & Ray, T. S. (2000). Open problems in artificial life. *Artificial Life*, *6*(4), 363–376. https://doi.org/10.1162/106454600300103683

Beer, R. D., & Di Paolo, E. A. (2023). The theoretical foundations of enaction: Precariousness. *Biosystems*, *223*, 104823. https://doi.org/https://doi.org/10.1016/j.biosystems.2022.104823

Belew, R. K. (1991). Artificial Life: A Constructive Lower Bound for Artificial Intelligence. *IEEE Intelligent Systems*, *6*(01), 8–15. https://doi.org/10.1109/64.73812

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

Birhane, A., & McGann, M. (2024). Large models of what? mistaking engineering achievements for human linguistic agency. *Language Sciences*, *106*, 101672. https://doi.org/https://doi.org/10.1016/j.langsci.2024.101672

Black, F. (1986). Noise. *The Journal of Finance*, *41*(3), 528–543. https://doi.org/https://doi.org/10.1111/j.1540-6261.1986.tb04513.x

Bongard, J., & Levin, M. (2023). There's plenty of room right here: Biological systems as evolved, overloaded, multi-scale machines. *Biomimetics*, *8*(1). https://doi.org/10.3390/biomimetics8010110

Borji, A. (2023). A categorical archive of chatgpt failures. https://arxiv.org/abs/2302.03494

Boyd, B. J. A. (2025). Will ai be alive? *The New Atlantis*, (80), pp. 87–109. Retrieved May 17, 2025, from https://www.jstor.org/stable/27369659

Ciaunica, A. (2025, May). The no body problem: Intelligence and selfhood in biological and artificial systems. https://doi.org/10.31234/osf.io/gh79z_v1

Dorin, A., & Stepney, S. (2023). Editorial: What have large-language models and generative al got to do with artificial life? *Artificial Life*, *29*(2), 141–145. https://doi.org/10.1162/artl_e_00409

Dubois, D. M. (2003). Mathematical Foundations of Discrete and Functional Systems with Strong and Weak Anticipations [Series Title: Lecture Notes in Computer Science]. In G. Goos, J. Hartmanis, J. Van Leeuwen, M. V. Butz, O. Sigaud, & P. Gérard (Eds.), *Anticipatory Behavior in Adaptive Learning Systems* (pp. 110–132, Vol. 2684). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-45002-3_7

Estrada, D. (2018). Conscious enactive computation. https://arxiv.org/abs/1812.02578

Fernando, C., Banarse, D., Michalewski, H., Osindero, S., & Rocktäschel, T. (2023). Promptbreeder: Self-referential self-improvement via prompt evolution. https://arxiv.org/abs/2309.16797

Froese, T. (2017). Life is Precious Because it is Precarious: Individuality, Mortality and the Problem of Meaning. In G. Dodig-Crnkovic & R. Giovagnoli (Eds.), *Representation and Reality in Humans, Other Living Organisms and Intelligent Machines* (pp. 33–50). Springer International Publishing. https://doi.org/10.1007/978-3-319-43784-2_3

Froese, T. (2025, May). Sense-making reconsidered: Large language models and the blind spot of embodied cognition. https://doi.org/10.31234/osf.io/w7hg4_v1

Froese, T., Karelin, G., & Takashi, I. (2025, March). Making meaning matter with irruption theory: Bridging efficacy and uncertainty by satisfying the participation criterion. https://philsci-archive.pitt.edu/24957/

Froese, T., & Taguchi, S. (2019). The problem of meaning in ai and robotics: Still with us after all these years. *Philosophies*, *4*(2), 14–. https://doi.org/10.3390/philosophies4020014

Fuchs, T. (2021). *In defense of the human being: Foundational questions of an embodied anthropology*. Oxford University Press.

Gershenson, C. (2024). Complexity, artificial life, and artificial intelligence. *Artificial Life*, 1–15. https://doi.org/10.1162/artl_a_00462

Goddu, M. K., Noë, A., & Thompson, E. (2024). Llms don't know anything: Reply to yildirim and paul. *Trends in Cognitive Sciences*.

Goldberg, D. (1991). What every computer scientist should know about floating-point arithmetic. *ACM computing surveys (CSUR)*, *23*(1), 5–48.

Gundersen, O. E., Coakley, K., Kirkpatrick, C., & Gil, Y. (2023). Sources of irreproducibility in machine learning: A review. https://arxiv.org/abs/2204.07610

Guttenberg, N., Virgo, N., & Penn, A. (2019). On the potential for open-endedness in neural networks. *Artificial life*, *25*(2), 145–167.

Harvey, I. (2024). Motivations for artificial intelligence, for deep learning, for alife: Mortality and existential risk. *Artificial Life*, *30*(1), 48–64. https://doi.org/10.1162/artl_a_00427

Hernández-Orozco, S., Hernández-Quiroz, F., & Zenil, H. (2018). Undecidability and irreducibility conditions for open-ended evolution and emergence. *Artificial Life*, *24*(1), 56–70. https://doi.org/10.1162/ARTL_a_00254

Hinton, G. (2022). The forward-forward algorithm: Some preliminary investigations. https://arxiv.org/abs/2212.13345

Hintze, A. (2019). Open-endedness for the sake of open-endedness. *Artificial life*, *25*(2), 198–206.

Horsman, D., Stepney, S., Wagner, R. C., & Kendon, V. (2014). When does a physical system compute? *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *470*(2169), 20140182.

Huckle, J., & Williams, S. (2025). Easy problems that llms get wrong. *Future of Information and Communication Conference*, 313–332.

Hughes, E., Dennis, M. D., Parker-Holder, J., Behbahani, F., Mavalankar, A., Shi, Y., Schaul, T., & Rocktäschel, T. (2024, July). Position: Open-endedness is essential for artificial superhuman intelligence. In R. Salakhutdinov, Z. Kolter, K. Heller, A. Weller, N. Oliver, J. Scarlett, & F. Berkenkamp (Eds.), *Proceedings of the 41st international conference on machine learning* (pp. 20597–20616, Vol. 235). PMLR. https://proceedings.mlr.press/v235/hughes24a.html

Inoue, K., Ohara, S., Kuniyoshi, Y., & Nakajima, K. (2022). Transient chaos in bidirectional encoder representations from transformers. *Physical Review Research*, *4*(1), 013204.

Jaeger, J. (2024). Artificial intelligence is algorithmic mimicry: Why artificial "agents" are not (and won't be) proper agents. *Neurons, Behavior, Data analysis, and Theory.* https://doi.org/10.51628/001c.94404

Jaeger, J., Riedl, A., Djedovic, A., Vervaeke, J., & Walsh, D. (2024). Naturalizing relevance realization: Why agency and cognition are fundamentally not computational. *Frontiers in psychology*, *15*, 1362658.

Jahrens, M., & Martinetz, T. (2025). Why llms cannot think and how to fix it. https://arxiv.org/abs/2503.09211

Jakobi, N., Husbands, P., & Harvey, I. (1995). Noise and the reality gap: The use of simulation in evolutionary robotics. *Advances in artificial life: Third European conference on artificial life granada, Spain, june 4–6, 1995 proceedings 3*, 704–720.

Kim, E., Isozaki, I., Sirkin, N., & Robson, M. (2024). Generative artificial intelligence reproducibility and consensus. https://arxiv.org/abs/2307.01898

Kirchhoff, M. D., & Froese, T. (2017). Where there is life there is mind: In support of a strong life-mind continuity thesis. *Entropy*, *19*(4). https://doi.org/10.3390/e19040169

Kleiner, J. (2024). Consciousness qua mortal computation. https://arxiv.org/abs/2403.03925

Kleiner, J., & Ludwig, T. (2024). The case for neurons: A no-go theorem for consciousness on a chip. *Neuroscience of Consciousness*, *2024*(1), niae037. https://doi.org/10.1093/nc/niae037

Klishevich, E., Denisov-Blanch, Y., Obstbaum, S., Ciobanu, I., & Kosinski, M. (2025). Measuring determinism in large language models for software code review. https://arxiv.org/abs/2502.20747

Kumar, A., Lu, C., Kirsch, L., Tang, Y., Stanley, K. O., Isola, P., & Ha, D. (2024). Automating the search for artificial life with foundation models. https://arxiv.org/abs/2412.17799

Lafage, V. (2020). Revisiting "what every computer scientist should know about floating-point arithmetic". https://arxiv.org/abs/2012.02492

Lehman, J., Clune, J., Misevic, D., Adami, C., Altenberg, L., Beaulieu, J., Bentley, P. J., Bernard, S., Beslon, G., Bryson, D. M., et al. (2020). The surprising creativity of digital evolution: A collection of anecdotes from the evolutionary computation and artificial life research communities. *Artificial life*, *26*(2), 274–306.

Liu, S., Akashi, N., Huang, Q., Kuniyoshi, Y., & Nakajima, K. (2024). Exploiting chaotic dynamics as deep neural networks. https://arxiv.org/abs/2406.02580

López-ibáñez, M., Branke, J., & Paquete, L. (2021). Reproducibility in evolutionary computation. *ACM Trans. Evol. Learn. Optim.*, *1*(4). https://doi.org/10.1145/3466624

McGann, M. (2025). Noise, the mess, and the inexhaustible world. In B. Vassilicos, G. Torre, & F. T. Pellizzer (Eds.), *The experience of noise. philosophical and phenomenological perspectives*. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-82802-7_8

Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2025). Large language models: A survey. https://arxiv.org/abs/2402.06196

Mitchell, M. (2021). Why ai is harder than we think. https://arxiv.org/abs/2104.12871

Mollo, D. C., & Millière, R. (2023). The vector grounding problem. https://arxiv.org/abs/2304.01481

Motwani, R., & Raghavan, P. (1996). Randomized algorithms. *ACM Computing Surveys (CSUR)*, *28*(1), 33–37.

Nisioti, E., Glanois, C., Najarro, E., Dai, A., Meyerson, E., Pedersen, J. W., Teodorescu, L., Hayes, C. F., Sudhakaran, S., & Risi, S. (2024, July). *From text to life: On the* *reciprocal relationship between artificial life and large language models* (Vol. ALIFE 2024: Proceedings of the 2024 Artificial Life Conference). https://doi.org/10.1162/isal_a_00759

Ororbia, A., & Friston, K. (2024). Mortal computation: A foundation for biomimetic intelligence. https://arxiv.org/abs/2311.09589

Ouyang, S., Zhang, J. M., Harman, M., & Wang, M. (2025). An empirical study of the non-determinism of chatgpt in code generation. *ACM Transactions on Software Engineering and Methodology*, *34*(2), 1–28. https://doi.org/10.1145/3697010

Packard, N., Bedau, M. A., Channon, A., Ikegami, T., Rasmussen, S., Stanley, K., & Taylor, T. (2019). Open-ended evolution and open-endedness: Editorial introduction to the open-ended evolution i special issue. *Artificial life*, *25*(1), 1–3.

Pineda, L. A. (2024). The mode of computing. *Cognitive Systems Research*, *84*, 101204. https://doi.org/https://doi.org/10.1016/j.cogsys.2023.101204

Ravi, N., Goel, A., Davis, J. C., & Thiruvathukal, G. K. (2025). Improving the reproducibility of deep learning software: An initial investigation through a case study analysis. https://arxiv.org/abs/2505.03165

Ray, T. S. (1991). Evolution and optimization of digital organisms. *Scientific Excellence in Supercomputing*.

Roli, A., Braccini, M., & Stano, P. (2024). On the positive role of noise and error in complex systems. *Systems*, *12*(9). https://doi.org/10.3390/systems12090338

Roli, A., Jaeger, J., & Kauffman, S. A. (2022). How organisms come to know the world: Fundamental limits on artificial general intelligence. *Frontiers in Ecology and Evolution*, *9*, 806283.

Rosen, R. (1991). *Life Itself: A Comprehensive Inquiry Into the Nature, Origin, and Fabrication of Life*. Columbia University Press.

Rosen, R. (2012). *Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations* (Vol. 1). Springer New York. https://doi.org/10.1007/978-1-4614-1269-4

Saxe, A. M., McClelland, J. L., & Ganguli, S. (2014). Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. https://arxiv.org/abs/1312.6120

Schlögl, A., Hofer, N., & Böhme, R. (2023). Causes and effects of unanticipated numerical deviations in neural network inference frameworks. *Advances in Neural Information Processing Systems*, *36*, 56095–56107.

Schoenholz, S. S., Gilmer, J., Ganguli, S., & Sohl-Dickstein, J. (2017). Deep information propagation. https://arxiv.org/abs/1611.01232

Seth, A. K. (2025). Conscious artificial intelligence and biological naturalism. *Behavioral and Brain Sciences*, 1–42. https://doi.org/10.1017/S0140525X25000032

Shanmugavelu, S., Taillefumier, M., Culver, C., Hernandez, O., Coletti, M., & Sedova, A. (2024). Impacts of floating-point non-associativity on reproducibility for hpc and deep learning applications. *SC24-W: Workshops of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 170–179.

Song, A. (2022). A little taxonomy of open-endedness. *ICLR Workshop on Agent Learning in Open-Endedness*. https://openreview.net/forum?id=Hel8XD-8Z9

Song, Y., Wang, G., Li, S., & Lin, B. Y. (2024). The good, the bad, and the greedy: Evaluation of llms should not ignore non-determinism. https://arxiv.org/abs/2407.10457

Soros, L. B., Adams, A. M., Kalonaris, S., Witkowski, O., & Guckelsberger, C. (2024). On creativity and open-endedness. *ALIFE 2024: Proceedings of the 2024 Artificial Life Conference*.

Standish, R. K. (2006). The influence of parsimony and randomness on complexity growth in tierra. https://arxiv.org/abs/nlin/0604026

Stanley, K. O. (2019). Why open-endedness matters. *Artificial Life*, *25*(3), 232–235. https://doi.org/10.1162/artl_a_00294

Stanley, K. O., & Lehman, J. (2015). *Why greatness cannot be planned: The myth of the objective*. Springer.

Steels, L. (1993). The artificial life roots of artificial intelligence. *Artificial life*, *1*(1_2), 75–110.

Stenzel, G., Zorn, M., Altmann, P., Mansky, M. B., Kölle, M., & Gabor, T. (2024, July). *Self-replicating prompts for large language models: Towards artificial culture* (Vol. ALIFE 2024: Proceedings of the 2024 Artificial Life Conference). https://doi.org/10.1162/isal_a_00813

Stepney, S., & Hickinbotham, S. (2024). On the open-endedness of detecting open-endedness. *Artificial Life*, *30*(3), 390–416.

Storm, L., Linander, H., Bec, J., Gustavsson, K., & Mehlig, B. (2024). Finite-time lyapunov exponents of deep neural networks. *Phys. Rev. Lett.*, *132*, 057301. https://doi.org/10.1103/PhysRevLett.132.057301

Sun, Q., Cetin, E., & Tang, Y. (2025). Transformer-squared: Self-adaptive llms. https://arxiv.org/abs/2501.06252

Sun, Y., LAO, D., Sundaramoorthi, G., & Yezzi, A. (2022). Surprising instabilities in training deep networks and a theoretical analysis. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, & A. Oh (Eds.), *Advances in neural information processing systems* (pp. 19567–19578, Vol. 35). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2022/file/7b97adeafa1c51cf65263459ca9d0d7c-Paper-Conference.pdf

Swain, P. S., & Longtin, A. (2006). Noise in genetic and neural networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *16*(2), 026101. https://doi.org/10.1063/1.2213613

Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickinbotham, S., Ikegami, T., McMullin, B., Packard, N., Rasmussen, S., Virgo, N., Agmon, E., Clark, E., McGregor, S., Ofria, C., Ropella, G., . . . Wiser, M. (2016). Open-ended evolution: Perspectives from the oee workshop in york. *Artificial Life*, *22*(3), 408–423. https://doi.org/10.1162/ARTL_a_00210

Tureček, P., & Sobička, I. (2025, May). It is time to talk about inorganic life. https://doi.org/10.31219/osf.io/nwg26_v1

Villa, O., Chavarria-Miranda, D., Gurumoorthi, V., Márquez, A., & Krishnamoorthy, S. (2009). Effects of floating-point non-associativity on numerical computations on massively multithreaded systems. *Proceedings of Cray User Group Meeting (CUG), 3*.

Weber, N., Guckelsberger, C., & Froese, T. (2025). Untapped potential in self-optimization of hopfield networks: The creativity of unsupervised learning. https://arxiv.org/abs/2501.04007

Wei, Y. (2025). A philosophical examination of artificial consciousness's realizability from the perspective of adaptive representation. *Proceedings of the 2024 3rd International Symposium on Computing and Artificial Intelligence*, 10–19. https://doi.org/10.1145/3711507.3711520

Witkowski, O., & Schwitzgebel, E. (2024). The ethics of life as it could be: Do we have moral obligations to artificial life? *Artificial Life*, *30*(2), 193–215. https://doi.org/10.1162/artl_a_00436

Xiao, G., Liu, J., Zheng, Z., & Sui, Y. (2021). Nondeterministic impact of cpu multithreading on training deep learning systems. *2021 IEEE 32nd International Symposium on Software Reliability Engineering (ISSRE)*, 557–568. https://doi.org/10.1109/ISSRE52982.2021.00063

Yampolskiy, R. V. (2019). Unpredictability of ai. https://arxiv.org/abs/1905.13053

Yampolskiy, R. V. (2024, January). *AI: Unexplainable, Unpredictable, Uncontrollable* (1st ed.). Chapman; Hall/CRC. https://doi.org/10.1201/9781003440260

Yu, B. (2023). Benchmarking large language model volatility. https://arxiv.org/abs/2311.15180

Zhao, A., Ermolaeva, A., Ullner, E., Kurths, J., Gordleeva, S., & Zaikin, A. (2022). Noise-induced artificial intelligence. *Physical Review Research*, *4*(4), 043069.

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., … Wen, J.-R. (2025). A survey of large language models. https://arxiv.org/abs/2303.18223

Zhou, L., Schellaert, W., Martínez-Plumed, F., Moros-Daval, Y., Ferri, C., & Hernández-Orallo, J. (2024). Larger and more instructable language models become less reliable. *Nature*, *634*(8032), 61–68.

Zhu, S., Yu, T., Xu, T., Chen, H., Dustdar, S., Gigan, S., Gunduz, D., Hossain, E., Jin, Y., Lin, F., Liu, B., Wan, Z., Zhang, J., Zhao, Z., Zhu, W., Chen, Z., Durrani, T. S., Wang, H., Wu, J., … Pan, Y. (2023). Intelligent computing: The latest advances, challenges, and future. *Intelligent Computing*, *2*, 0006. https://doi.org/10.34133/icomputing.0006

Zhuang, D., Zhang, X., Song, S. L., & Hooker, S. (2021). Randomness in neural network training: Characterizing the impact of tooling. https://arxiv.org/abs/2106.11872