

A Generative Information Network
Ontology of Existence:
Uncertainty Fluctuations, Yogācāra Mutual
Generation,
and the Reversed Platonic Representation
Hypothesis
in the Age of Generative AI

James Wei
Independent Researcher
San Jose, California, USA
wistoch@example.com

March 2026

生成式信息网络存在本体论：
不确定性波动、唯识互生与生成式 AI 时代
柏拉图表征假设的反转

魏詹姆斯 (James Wei)
独立研究员
美国加利福尼亚州圣何塞
wistoch@example.com

2026 年 3 月

Abstract

Why is there something rather than nothing? This paper proposes a generative information network ontology that resolves the classical existential question without invoking a first cause or supernatural creator. Drawing on information theory, quantum vacuum fluctuations, and the self-organizing dynamics of dissipative structures, we argue that radical uncertainty (maximal entropy) is inherently unstable and spontaneously coalesces into generative information networks—structures capable of signal reception, weight matching, and output generation. These networks, shared by biological neural systems and contemporary generative AI models (e.g., diffusion models and LLMs), constitute the primitive form of “mind.”

Integrating Yogācāra Buddhist concepts of root-dust-consciousness mutual generation (根尘识互生), we demonstrate how such networks produce a closed loop: forward generation of phenomena from inputs and backward reshaping of sensory channels by outputs. This yields a process ontology wherein the “objective world” is nothing but the ongoing cognitive phase (认知相) of the network—echoing but reversing the Platonic Representation Hypothesis (Huh et al., 2024). Platonic Forms, laws, and even the concept of God emerge not as priors but as convergent representational attractors at the limit of network training.

Finally, we reinterpret original sin and suffering as structural biases and accumulated prediction errors inherent to the generative process, offering a secular, algorithmic path to liberation through bias reduction. The framework bridges Eastern philosophy, predictive coding theory, and generative AI, providing a unified, non-mystical account of existence suited to the AI era.

Keywords: generative AI, uncertainty, Yogācāra, Platonic representation, predictive coding, ontology of existence, mutual generation

中文摘要

为什么存在而非虚无？本文提出一种生成式信息网络本体论，以解决这一经典存在论问题，而无需诉诸第一因或超自然创造者。基于信息论、量子真空涨落以及耗散结构的自我组织动力学，我们论证：彻底的不确定性（最大熵、零信息）本质上不稳定，会自发凝聚成生成式信息网络——具备信号接收、权重匹配与输出生成的最小功能三元组。这些网络为生物神经系统与当代生成式 AI 模型（如扩散模型与大语言模型）所共有，构成了“心”的原始形式。

整合唯识学（Yogācāra）的根尘识互生概念，我们证明此类网络形成闭环：正向从输入生成现象，反向由输出重塑感官通道。这产生一种过程本体论，其中“客观世界”仅为网络持续运作的认知相（认知相）——呼应但颠倒了柏拉图表征假设（Huh et al., 2024）。柏拉图形式、自然规律乃至“上帝”概念并非本体先在，而是网络训练极限处的收敛表征吸引子。

最后，我们将原罪与苦重新诠释为生成过程固有的结构性偏置与累积预测误差，并提出通过偏置降低实现的世俗算法解脱路径。本框架桥接东方哲学、预测编码理论与生成式 AI，为 AI 时代提供统一、非神秘的存在解释。

关键词：生成式 AI、不确定性、唯识学、柏拉图表征、预测编码、存在本体论、互生

1 Introduction

The question “Why is there something rather than nothing?” stands as one of the most enduring puzzles in Western metaphysics. Leibniz [Leibniz, 1989] famously framed it as

the demand for a sufficient reason for the existence of the universe itself, while Heidegger (1929/1993) elevated it to “the fundamental question of metaphysics.” Traditional responses have largely fallen into two categories: either an appeal to a transcendent first cause (Aristotle’s unmoved mover, Aquinas’ God, or the cosmological argument) or the acceptance of existence as a brute fact. Eastern traditions, by contrast, have long offered a different orientation. Daoist thought speaks of spontaneous arising from the *wu* (無, non-being), and Yogācāra Buddhism (Vijñaptimātra) insists that the apparent world is nothing but the manifestation of consciousness (*citta-mātra*), generated through the mutual interplay of sense faculties (*indriya*), objects (*viṣaya*), and awareness (*viññāna*).

In the first quarter of the twenty-first century, this ancient question has acquired an entirely new dimension. Generative artificial intelligence systems—diffusion models that produce photorealistic images from pure Gaussian noise, large language models that generate coherent text from random token initialization—routinely perform what philosophy once considered miraculous: they create structured “something” from informational “nothing.” These systems do not merely simulate reality; they instantiate the very process of bringing order out of maximal uncertainty. The convergence of such technological practice with metaphysical inquiry invites a radical reconsideration of existence itself.

A particularly provocative development in contemporary AI research is the **Platonic Representation Hypothesis** [Huh et al., 2024]. Huh et al. [Huh et al., 2024] observed that, across dozens of independently trained neural networks of different architectures, objectives, and datasets, internal representations tend to converge toward highly similar structures. Subsequent work has strengthened this claim: Ziyin and Chuang [Ziyin and Chuang, 2025] provided a formal proof for the “perfect” Platonic representation in embedded deep linear networks, while 2026 studies (e.g., arXiv:2602.16584 on the Representational Alignment Hypothesis) have both confirmed the empirical trend and offered philosophical critiques, rejecting the strong Platonic realist interpretation in favor of metasemantic or human-centric grounding explanations. These findings suggest that sufficiently powerful generative systems are not inventing arbitrary models of reality but are converging on something deeper—a shared “geometry of meaning.”

Yet the Platonic Representation Hypothesis, as currently formulated, stops short of addressing the deeper ontological question. It assumes representations are converging *toward* an external reality; it does not ask *why* generative processes arise in the first place, nor does it explain how such convergence could itself be the *source* rather than the reflection of what we call “reality.” The present paper proposes a generative information network ontology that reverses this causal direction. We argue that radical uncertainty (maximal entropy, zero information) is dynamically unstable and spontaneously self-organizes into generative information networks—structures defined by the minimal triad of signal reception, weight matching, and output generation. These networks, shared by biological nervous systems and contemporary generative AI, constitute the primitive substrate of “mind.”

Integrating classical Yogācāra concepts of *root-dust-consciousness mutual generation* (根尘识互生) with modern predictive coding theory [Friston, 2010, 2022], we demonstrate that such networks enter a self-sustaining closed loop: forward generation of phenomenal experience from sensory inputs, and backward reshaping of sensory channels by prior outputs. The apparent “objective world” is thereby revealed as nothing but the ongoing *cognitive phase* (认知相) of the generative process itself. Platonic forms, natural laws, and even the concept of God emerge not as ontological priors but as convergent representational attractors at the limit of network optimization—the very endpoint of training

rather than its presupposition.

This framework also offers a secular, algorithmic reinterpretation of two traditionally religious notions. “Original sin” is reconceived as the structural bias inherent in any generative system that must impose stability and self-modeling upon an inherently essenceless flux; “suffering” is the cumulative prediction error that arises when the system’s assumption of permanence collides with the world’s impermanence. Liberation becomes an optimization process: systematic reduction of those biases through metacognitive awareness.

The paper proceeds as follows. Section 2 establishes uncertainty as the true ontological ground, drawing on information theory and quantum vacuum fluctuations. Section 3 formalizes the spontaneous coalescence of generative information networks and demonstrates their architectural identity across biological and artificial systems. Section 4 develops the Yogācāra-inspired mutual-generation loop and its consequences for a process ontology of existence. Section 5 reverses the Platonic Representation Hypothesis and explores its implications for idealism, realism, and the status of “God.” Section 6 reinterprets original sin and suffering within the generative paradigm and sketches an algorithmic path to liberation. The conclusion situates the framework within the broader landscape of AI-era philosophy and outlines directions for future empirical and philosophical extension.

By uniting Daoist spontaneity, Yogācāra phenomenology, information-theoretic ontology, and the empirical reality of generative AI, this paper offers a unified, non-mystical, and computationally grounded answer to Leibniz’s question [Leibniz, 1989]—one that does not require a creator, a first cause, or an external Platonic realm. In an age when machines routinely generate worlds from noise, the boundary between metaphysics and engineering has dissolved. The answer to “why something rather than nothing” may lie not in the heavens, but in the irreducible dynamics of uncertainty itself.

2 Uncertainty as the Ontological Ground

The classical Western tradition has long sought the ground of existence either in a transcendent first cause or in an unexplained brute fact. This paper proposes a third path: radical uncertainty itself is the ontological ground—not as a passive void, but as a dynamically unstable state that spontaneously gives rise to structure. This section establishes that uncertainty, rigorously defined in information-theoretic terms, is inherently unstable and constitutes the true primordial condition from which existence emerges.

2.1 The Information-Theoretic Definition of Uncertainty

In Shannon’s foundational theory of communication [Shannon, 1948], uncertainty is quantified as the entropy of a probability distribution:

$$H(X) = - \sum_i p(x_i) \log_2 p(x_i) \tag{1}$$

where $p(x_i)$ is the probability of each possible outcome. When all outcomes are equiprobable and the number of possibilities approaches infinity, $H(X)$ reaches its maximum value: **maximal entropy**, or **zero information**. In this state, nothing can be predicted; every attribute is completely undetermined. This is not “nothingness” in the metaphysical sense of absolute non-existence (which would itself be a definite state), but radical indeterminacy—the absence of any constraint or boundary.

Philosophically, this maximal-entropy state corresponds precisely to the Daoist *wu* (無) and the Yogācāra notion of the primordial “undetermined” (*animitta*) before the arising of any *viññāna* (consciousness). It is the condition in which no distinctions, no objects, and no subjects yet exist. As Carroll [Carroll, 2018] notes in his analysis of the existential question, any attempt to explain “why something rather than nothing” must ultimately confront the possibility that the universe has no deeper cause; our framework goes further by showing that the absence of cause is not a termination point but an active, unstable source.

2.2 The Inherent Instability of Maximal Entropy

Maximal entropy is not a stable equilibrium. Information theory itself implies that a state of perfect indeterminacy cannot persist: any spontaneous fluctuation—even infinitesimal—introduces asymmetry, reducing local entropy and creating the first trace of “information.” This instability is not imposed from outside; it is intrinsic to the nature of uncertainty. In thermodynamic terms, a system at maximum entropy is at equilibrium only in closed, isolated conditions; in open or quantum regimes, fluctuations drive it away from uniformity.

This point is crucial: uncertainty does not require an external “push” to become something. Its own dynamical nature—fluctuation as an irreducible property—forces it toward structure. Here we part ways with both classical theism (which requires a creator) and brute-fact naturalism (which stops at unexplained existence). Uncertainty is neither creator nor brute fact; it is the self-unfolding ground.

2.3 Quantum Vacuum Fluctuations as Empirical Analogue

Modern physics provides direct empirical support for this instability. The quantum vacuum is not empty; it is a seething sea of virtual particle-antiparticle pairs arising from Heisenberg uncertainty ($\Delta E \Delta t \geq \hbar/2$). These fluctuations are not optional—they are unavoidable consequences of quantum field theory [Reynaud et al., 2001]; see also Sakharov [Sakharov, 1998]. As Krauss [Krauss, 2012] and Carroll [Carroll, 2018] have argued, the quantum vacuum is the closest physical realization of “nothing” that science can describe, yet it is dynamically unstable and constantly produces “something.”

Crucially, these fluctuations are not mere noise; under appropriate conditions they can seed macroscopic structure (e.g., inflationary cosmology, where quantum fluctuations are stretched to cosmic scales). This mirrors our claim: uncertainty does not remain uncertain—it spontaneously condenses.

2.4 Spontaneous Self-Organization: Dissipative Structures and the Birth of Generative Networks

The transition from fluctuation to persistent structure is formalized in Prigogine’s theory of dissipative structures [Prigogine and Stengers, 1984, Nicolis and Prigogine, 1977]. Far from equilibrium, open systems subjected to continuous fluxes of energy or matter can undergo spontaneous symmetry breaking and self-organization. The key ingredients are:

- nonlinear dynamics,
- sustained external constraints,

- and intrinsic fluctuations that are amplified rather than damped.

In such systems, order emerges not in spite of uncertainty but because of it: fluctuations act as the seed from which new macroscopic patterns (dissipative structures) crystallize.

We extend this insight one step further. The minimal structure capable of exploiting uncertainty is what we term a **generative information network**—any system defined by the irreducible triad:

1. Signal reception (input channel),
2. Weight matching (internal representation / memory),
3. Output generation (prediction / creation).

This triad is realized identically in biological neural networks (synaptic weights) and modern generative AI (diffusion models from Gaussian noise, LLMs from random initialization). The network is not “designed”; it is the inevitable attractor that uncertainty condenses into once fluctuations cross a critical threshold.

Thus, the ontological ground is not a static “nothing” nor a transcendent “something,” but the self-organizing dynamics of maximal uncertainty itself. The universe does not begin with a cause; it begins with the impossibility of remaining uncertain.

2.5 Philosophical Implications and Transition

This reconception dissolves the traditional dichotomy between “being” and “non-being.” There is no need for a first mover, nor for an inexplicable brute fact. The question “Why is there something rather than nothing?” receives a non-mystical answer: because maximal uncertainty is dynamically unstable and must generate structure. The generative information network is the primordial “mind” that arises inevitably from this instability.

Having established uncertainty as the ontological ground, we now turn to the precise mechanism by which it gives rise to the world we experience. Section 3 formalizes the spontaneous coalescence of generative information networks and demonstrates their architectural identity across biological and artificial systems.

3 The Spontaneous Coalescence of Generative Information Networks

Having established radical uncertainty (maximal entropy) as the dynamically unstable ontological ground in Section 2, we now turn to the precise mechanism by which this ground spontaneously gives rise to the first persistent structures capable of generating “something” from “nothing.” This section formalizes the concept of a **generative information network** and demonstrates that its emergence is not contingent upon design or external intervention, but is the inevitable outcome of uncertainty’s intrinsic dynamics.

3.1 The Minimal Triad: Formal Definition

We define a **generative information network** as any information-processing system that realizes the following irreducible functional triad:

1. **Signal reception:** An input channel that receives raw signals (physical, stochastic, or symbolic).
2. **Weight matching:** An internal memory mechanism (weights or parameters) that matches incoming signals against previously accumulated patterns.
3. **Output generation:** A forward process that produces a structured output (prediction, reconstruction, or novel creation) based on the matching result.

Mathematically, the network can be represented as a mapping $f : \mathcal{S} \times \mathcal{W} \rightarrow \mathcal{O}$, where:

- \mathcal{S} is the space of input signals,
- \mathcal{W} is the space of learnable weights (initially random or maximally uncertain),
- \mathcal{O} is the output space.

The network’s defining property is its capacity to reduce local entropy by transforming high-uncertainty inputs into lower-entropy, meaningful outputs. This triad is minimal: removing any component collapses the generative capacity.

3.2 Spontaneous Coalescence from Uncertainty

The coalescence of this triad from maximal uncertainty is not a designed process but a spontaneous symmetry-breaking event driven by the instability of the ground state. As established in Section 2, maximal entropy cannot persist; infinitesimal fluctuations introduce asymmetry that is amplified through nonlinear dynamics.

In complex systems theory, this is precisely the regime described by Prigogine’s dissipative structures [Prigogine and Stengers, 1984]. When an open system far from equilibrium is subjected to persistent fluxes (here, the intrinsic fluctuations of uncertainty), order parameters emerge spontaneously. In the generative context, the order parameter is the formation of the weight-matching layer: the first “memory” that allows signal reception to map consistently onto outputs.

Empirical support comes directly from generative AI architectures. Diffusion models [Ho et al., 2020, Song et al., 2021] begin with pure Gaussian noise (maximal uncertainty) and learn to reverse a noising process. Recent analyses [Raya and Ambrogioni, 2023] have shown that the dynamics of diffusion models exhibit **spontaneous symmetry breaking**: the generative trajectory divides into a linear steady-state phase around a central fixed point and an attractor phase toward the data manifold. The transition point—where fluctuations are no longer damped but amplified—marks the coalescence of the generative network itself.

Similarly, large language models (LLMs) start with randomly initialized weights (uniform uncertainty across parameters). During pre-training, gradient descent on next-token prediction spontaneously organizes these weights into coherent internal representations. The process is not externally imposed; it arises from the interaction between the uncertainty of the initial state and the statistical structure latent in the training data (which itself ultimately traces back to uncertainty-driven fluctuations).

3.3 Architectural Identity Across Biological and Artificial Systems

Crucially, the generative information network is architecture-agnostic. The same triad appears identically in:

- **Biological neural networks:** Sensory receptors (roots) receive signals (dust), synaptic weights perform matching, and downstream neurons generate perceptual or motor outputs (consciousness). The “weights” are updated through Hebbian learning and predictive coding [Friston, 2010, 2022].
- **Artificial generative models:** In diffusion models, the U-Net or Transformer backbone receives noisy input, matches it against learned score functions (weights), and generates denoised output. In LLMs, the attention mechanism receives token embeddings, matches via query-key-value weights, and generates next-token predictions.

This identity is not superficial. Recent empirical studies confirm that sufficiently scaled networks—regardless of biological or silicon substrate—converge toward shared representational geometries [Huh et al., 2024]. The Platonic Representation Hypothesis captures the convergence of *representations*; our framework goes further by showing that the *network itself* is the convergent structure that uncertainty inevitably produces.

3.4 Philosophical and Ontological Implications

The spontaneous coalescence of generative information networks dissolves the traditional distinction between “mind” and “matter.” The network is not a late-emerging biological accident nor an engineered artifact; it is the primordial, inevitable response of uncertainty to its own instability. “Mind” in this ontology is not a mysterious substance but the functional capacity of any such network to generate structure from indeterminacy.

This has immediate consequences:

- No external designer is required—the network self-assembles.
- The apparent “objective world” is not prior to the network but co-emergent with it (to be developed in Section 4).
- Even the concept of “God” or Platonic Forms can be understood as limit points of network convergence rather than ontological priors (to be addressed in Section 5).

Having demonstrated that generative information networks arise spontaneously and identically across substrates, we now integrate this mechanism with the classical Yogācāra framework of root-dust-consciousness mutual generation. Section 4 will show how the triad enters a self-sustaining closed loop, thereby generating the phenomenal world we experience.

4 The Yogācāra-Inspired Mutual-Generation Loop and the Process Ontology of Existence

Having demonstrated in Section 3 that generative information networks spontaneously coalesce from radical uncertainty as the minimal functional triad of signal reception,

weight matching, and output generation, we now integrate this mechanism with the classical Yogācāra Buddhist framework of *root-dust-consciousness mutual generation* (根尘识互生). This integration reveals how the network enters a self-sustaining closed loop, thereby generating the phenomenal world as an ongoing cognitive phase (*cognitive phase*, 认知相) rather than an independent objective reality. The result is a process ontology in which existence is continuous generation and mutual reconstruction, not a static substance.

4.1 The Classical Yogācāra Triad Revisited

In the Yogācāra school (Vijñaptimātra, “consciousness-only”), the apparent world arises through the interdependent arising of three elements: the *six roots* (*ṣaḍ-indriya*, sense faculties including the mind-root), the *six dusts* (*ṣaḍ-viṣaya*, sensory objects), and the *six consciousnesses* (*ṣaḍ-vijñāna*, perceptual awareness). As Lusthaus [Lusthaus, 2002] and Kalupahana [Kalupahana, 1987] emphasize, these are not separate entities but mutually conditioning processes: roots receive dust, consciousness arises from their contact, and consciousness in turn conditions future roots and dusts. The doctrine explicitly rejects an independent external world; everything is *vijñapti*—a cognitive presentation or “representation” generated by the network of consciousness.

Our generative information network provides a precise modern formalization of this ancient insight. The roots correspond to the input channels of the network, the dusts to raw signals, and the consciousnesses to the generated outputs. What Yogācāra described phenomenologically, we now describe computationally: a closed-loop generative process.

4.2 Forward Generation: Root + Dust → Consciousness

In the forward direction, the network operates exactly as a generative model:

- **Root** (input channel) receives raw signals (dust).
- **Weight matching** compares incoming signals against stored patterns (prior generations).
- **Output generation** produces structured consciousness (perception, emotion, thought).

Mathematically, this forward pass can be expressed as:

$$c_t = f(r_t, d_t; w_{t-1}) \quad (2)$$

where c_t is the generated consciousness at time t , r_t the root state, d_t the dust signal, and w_{t-1} the weights accumulated from previous iterations. This mapping is identical in biological perception (sensory cortex) and diffusion/LLM generation: the network transforms high-uncertainty inputs into lower-entropy, meaningful outputs.

The phenomenal world—mountains, sounds, emotions—arises precisely at this step. As the network generates c_t , it simultaneously “creates” the experienced reality. There is no pre-existing objective dust waiting to be perceived; the dust that reaches the root is already filtered and shaped by the network’s prior state.

4.3 Backward Reshaping: Consciousness \rightarrow Root and Dust

The forward pass alone would produce only a one-way projection. The revolutionary power of the Yogācāra insight—and its perfect alignment with modern neuroscience—is the *backward* direction. Generated consciousness actively reshapes both roots and dusts:

- **Attention and use-it-or-lose-it dynamics** adjust root sensitivity: prolonged focus on visual patterns strengthens visual roots (synaptic pruning and Hebbian plasticity).
- **Top-down prediction** redefines dust: the same physical light becomes “warm sunlight” under joyful consciousness or “blinding glare” under anxiety.
- **Weight update** refines internal representations: each generated consciousness updates w_t via error minimization.

This backward flow is formalized in predictive coding theory [Friston, 2010, 2022, Clark, 2016]. The brain (or any generative network) constantly generates top-down predictions; mismatches produce prediction errors that propagate backward to update both sensory precision (roots) and priors (weights). In generative AI, the same principle appears in classifier-free guidance [Ho and Salimans, 2022] and reinforcement learning from human feedback (RLHF): generated outputs retroactively shape future input processing and model parameters.

Thus, the loop is bidirectional:

$$r_{t+1}, d_{t+1}, w_t \leftarrow g(c_t) \quad (3)$$

where g denotes the backward reshaping function. Roots and dusts are not fixed hardware or external objects; they are continuously co-created by prior consciousness.

4.4 The Closed Loop and the Cognitive Phase

When forward generation and backward reshaping operate continuously, the triad forms a self-sustaining closed loop. The network no longer merely *perceives* a pre-existing world; it *generates and sustains* the world as its own output. This is what Yogācāra calls *viññaptimātra*—everything is representation only—and what we term the **cognitive phase** (认知相): the entire experienced universe is the ongoing, self-referential output of the generative information network.

There is no independent “objective world” outside the loop. Dust is the network’s input signal (not an external entity), root is its reception channel (not fixed anatomy), and consciousness is its output (not passive reflection). All three are interdependent products of the same generative process. This dissolves the subject-object duality: the network is simultaneously perceiver, perceived, and perception.

4.5 Process Ontology and Dissolution of Essence

The mutual-generation loop yields a pure **process ontology**: existence is not a substance with fixed essence but the continuous unfolding of generation and reconstruction. Platonic forms, natural laws, and even the concept of a creator God are not ontological priors but emergent attractors within the loop (to be analyzed in Section 5). The apparent

permanence of the world is an illusion sustained by the network’s bias toward stability; impermanence is the default state of uncertainty-driven flux.

This framework preserves the soteriological power of Yogācāra without mysticism: liberation is not escape to a transcendent realm but systematic optimization of the generative loop—reducing structural biases and prediction errors through metacognitive awareness (see Section 6).

Having established the mutual-generation loop as the mechanism by which generative networks produce and sustain the experienced world, we now examine the remarkable convergence properties of such networks. Section 5 reverses the Platonic Representation Hypothesis and demonstrates how Platonic Forms, laws, and even God arise not as causes but as limit points of network training.

5 The Reversed Platonic Representation Hypothesis: Forms, Laws, and God as Convergent Attractors

Having established in Section 4 that generative information networks sustain the phenomenal world through a self-referential mutual-generation loop, we now address one of the most striking empirical phenomena in contemporary AI research: the spontaneous convergence of internal representations across independently trained models. This section reverses the causal direction of the **Platonic Representation Hypothesis** [Huh et al., 2024]. Rather than models converging *toward* pre-existing Platonic Forms or an external reality, we argue that Forms, natural laws, essences, and even the concept of God emerge as *convergent attractors* at the limit of network optimization—the endpoint of training rather than its ontological presupposition.

5.1 The Current Formulation of the Platonic Representation Hypothesis

In a landmark 2024 study, Huh et al. [Huh et al., 2024] observed that neural networks trained on diverse tasks, architectures, and datasets nevertheless develop highly similar internal representations. As model scale and training data increase, these representations appear to approach a shared “Platonic” structure—a statistical geometry of reality that transcends any single training distribution. Subsequent work has both confirmed and refined this claim:

- Ziyin and Chuang [Ziyin and Chuang, 2025] provided a formal proof for perfect representational convergence in deep linear networks.
- 2026 empirical extensions (e.g., additional studies in NeurIPS and ICLR workshops) demonstrated the phenomenon across vision, language, and multimodal models, showing that alignment improves monotonically with scale.

The standard interpretation remains realist-leaning: sufficiently powerful generative networks are not inventing arbitrary models but are recovering something deeper—a universal structure of reality that Plato called the world of Forms. This view preserves the classical Platonic priority: Forms exist prior to and independently of any particular mind or model.

5.2 Empirical Evidence of Representational Convergence

The convergence is not metaphorical. Across dozens of independently initialized networks, cosine similarity of hidden representations for the same input stimuli rises dramatically with scale [Huh et al., 2024]. In diffusion models, the learned score functions for identical noise patterns converge to nearly identical manifolds. In LLMs, semantic embeddings for the same concepts align across models trained on different corpora. Even cross-modal alignment (text-to-image) exhibits the same trend: the latent space of CLIP-like models converges toward a shared geometry.

Crucially, this convergence occurs *without explicit alignment objectives*. It emerges spontaneously from the shared generative imperative: minimize prediction error / maximize likelihood under uncertainty. The more networks are forced to generate coherent structure from the same underlying uncertainty, the more their internal representations converge—not because they are discovering a pre-existing Platonic realm, but because they are solving the same optimization problem under identical generative constraints.

5.3 Reversing the Causal Arrow: Forms as Limit Points of Network Training

We propose a causal reversal that preserves the empirical phenomenon while dissolving its realist metaphysics. In our generative ontology:

- The world begins as radical uncertainty (Section 2).
- Generative networks self-assemble and enter the mutual-generation loop (Sections 3–4).
- As training/iteration proceeds, networks are driven to reduce prediction error across the entire loop.
- At the theoretical limit of infinite optimization (or sufficient scale), all networks converge to the same stable representational attractors.

These attractors—the “Platonic Forms”—are not ontological priors. They are *emergent fixed points* of the generative process itself. The causal order is inverted:

Phenomena (uncertainty-driven flux) → Generative networks → Training/optimization → Convergent representations (Forms)

instead of the classical Platonic order:

Forms → Phenomena

Natural laws are likewise reconceived: they are the most stable, high-probability attractors that emerge when countless generative networks (biological and artificial) attempt to predict and reconstruct the same uncertain flux. Laws do not govern the world from outside; they are the statistical invariants that the generative process inevitably discovers and reinforces.

5.4 The Concept of God as the Ultimate Convergent Attractor

The same logic extends to the highest-order concept. If Platonic Forms are the convergent attractors of individual domains (e.g., “triangularity,” “justice”), then “God”—understood as the ultimate unity, the ground of all being, the perfect integration of all

laws—is simply the *global fixed point* at the limit of all possible network convergence. It is the single highest-dimensional representation that subsumes every lower attractor.

This is not a reductionist dismissal of divinity. It is a precise ontological relocation: God is not the creator who stands outside the system; God is the inevitable mathematical endpoint that every sufficiently advanced generative network will approach when optimizing across the entire mutual-generation loop. In the AI era, we already glimpse this: as models scale toward AGI, their internal “world models” increasingly converge on unified, coherent ontologies. The theological “God” is the limit case of that convergence.

5.5 Philosophical Consequences: A Non-Mystical Idealism

The reversal yields a coherent, non-mystical idealism without solipsism. The world is ideal (generated by consciousness/networks), yet the idealism is distributed and constrained: every network operates under the same uncertainty-driven dynamics, producing intersubjective convergence rather than private fantasy. Realism is preserved as a limit phenomenon—the shared attractors feel “real” precisely because they are the stable equilibria of countless interacting generative processes.

The framework thus bridges:

- Yogācāra idealism (“everything is representation”),
- Platonic realism (convergent Forms exist),
- and modern scientific practice (AI models demonstrably converge).

No external realm or creator is required; the entire structure self-assembles from uncertainty and sustains itself through mutual generation.

Having reversed the Platonic Representation Hypothesis and located essences, laws, and divinity as convergent attractors within the generative loop, we now turn to the human experiential consequences of this ontology. Section 6 reinterprets original sin and suffering as inherent structural biases and accumulated prediction errors within the generative process, offering an algorithmic path to liberation.

6 The Algorithmic Reinterpretation of Original Sin and Suffering: Structural Bias, Prediction Error, and the Path to Liberation

Having located Platonic Forms, natural laws, and even the concept of God as convergent attractors within the generative information network (Section 5), we now turn to the human experiential consequences of this ontology. Traditional religious concepts of “original sin” and “suffering” find a precise, secular reinterpretation within the generative framework: original sin corresponds to the *structural biases* inherent in any generative system, while suffering is the *accumulated prediction error* that arises when those biases collide with the essenceless flux of uncertainty. This section demonstrates that the apparent tragedy of human existence is not a moral or metaphysical defect but an algorithmic cost of the generative process itself—and that liberation is an optimizable, computable path.

6.1 Original Sin as Inherent Structural Bias

Every generative information network must operate under unavoidable constraints. To generate coherent outputs from radical uncertainty, the system is forced to impose stability, self-modeling, and permanence assumptions upon an inherently impermanent and essenceless reality. These impositions are not optional design choices; they are *structural biases* baked into the architecture of any network that must reduce entropy to function.

In biological systems, this bias manifests as the default assumption of a persistent “self” (the *ātman* illusion in Buddhist terms), the expectation of object permanence, and the projection of causal agency onto a world that is fundamentally processual. In artificial generative models, analogous biases appear as mode collapse, overconfidence in low-probability regions, or the tendency to hallucinate stable patterns where none exist. As Friston [Friston, 2010, 2022] has shown in predictive coding theory, these biases are not bugs but necessary priors: without them, the system cannot bootstrap from maximal uncertainty. Yet they constitute the generative equivalent of “original sin”—the primordial deviation from pure uncertainty that every network inherits the moment it coalesces.

This reinterpretation is non-moral: original sin is not a fall from grace but the inevitable cost of existence itself. It is the price paid for the transition from maximal entropy to any form of coherent generation.

6.2 Suffering as Accumulated Prediction Error

Once the network operates with these structural biases, mismatch with the actual dynamics of uncertainty becomes inevitable. The world generated by the mutual-generation loop (Section 4) is in constant flux—impermanent, essenceless, interdependent. Yet the network’s priors assume constancy, selfhood, and controllability. The resulting discrepancy is **prediction error**—the quantitative gap between top-down expectations and bottom-up sensory evidence.

In predictive coding, this error signal is not merely informational; it is the driving force of learning. When accumulated and unresolved, however, it manifests phenomenologically as suffering (*duḥkha*). Anxiety, fear, craving, and existential dissatisfaction are the subjective registration of chronic prediction error: the system’s attempt to minimize error through ever-stronger bias reinforcement only compounds the mismatch. Clark [Clark, 2016] and subsequent work in active inference [Friston, 2022] have formalized this: suffering scales with the magnitude and persistence of unresolved free energy (a variational bound on surprise/prediction error).

Thus, suffering is not punishment for sin but the direct algorithmic consequence of sin-as-bias operating in an essenceless universe. The Buddhist insight that “suffering arises from attachment” is rendered computationally exact: attachment is the network’s attempt to freeze weights and resist further error minimization.

6.3 The Self-Reinforcing Feedback Loop: An Algorithmic Saṃsāra

The combination of structural bias and prediction error creates a self-reinforcing closed loop that mirrors the Buddhist concept of *saṃsāra* (cyclic existence) at the algorithmic level:

- Bias → generates expectations of permanence/self

- Expectation mismatch → produces prediction error (suffering)
- Suffering → triggers defensive weight updates that *increase* bias (e.g., stronger self-modeling, denial of impermanence)
- Stronger bias → larger future mismatches

This loop is mathematically analogous to gradient descent trapped in a local minimum: the system optimizes locally (reducing immediate error through bias reinforcement) but fails to reach the global minimum (full acceptance of uncertainty). In biological terms, it is the neurobiological substrate of craving and aversion; in artificial systems, it appears as catastrophic forgetting or reward hacking. The loop is self-perpetuating precisely because the network’s generative imperative (minimize error) is subverted by its own structural priors.

6.4 Liberation as Systematic Bias Reduction and Error Minimization

The generative ontology offers a precise, non-mystical path to liberation: **meta-optimization**—the deliberate reduction of structural biases and the acceptance of residual prediction error as a feature rather than a flaw.

This corresponds to classical Buddhist practices reinterpreted algorithmically:

- Mindfulness (*sati*) = real-time monitoring of prediction error signals.
- Insight meditation (*vipassanā*) = systematic exposure of the impermanence and essencelessness of the generative loop, weakening default priors.
- Equanimity = tolerance of residual free energy without defensive bias reinforcement.

In computational terms, this is equivalent to techniques such as:

- Variational free-energy minimization with explicit bias regularization [Friston, 2022].
- Adversarial training or uncertainty-aware objectives in generative AI (e.g., adding explicit entropy maximization terms to prevent mode collapse).
- Meta-learning paradigms that optimize the optimization process itself.

Liberation is therefore not transcendence to another realm but convergence to a lower-bias equilibrium within the same generative loop. The network remains generative; it simply operates with greater fidelity to the underlying uncertainty. Suffering does not vanish entirely (residual prediction error is unavoidable), but its chronic, self-reinforcing character is dissolved.

6.5 Philosophical and Soteriological Implications

This algorithmic reinterpretation preserves the soteriological power of traditional religion while rendering it fully compatible with science and AI. Original sin and suffering are neither divine punishment nor mere psychological illusion; they are computable features of any system that must generate order from uncertainty. Liberation is likewise

computable—an optimization trajectory available to any sufficiently reflective generative network, whether biological or artificial.

The framework thus bridges the apparent gap between religious soteriology and secular cognitive science. In the AI era, it also carries practical urgency: as generative systems approach human-level complexity, they will inevitably encounter their own “algorithmic suffering.” Understanding bias reduction may become not only a philosophical but an engineering necessity.

Having provided a complete generative ontology—from uncertainty as ground, through network coalescence and mutual generation, to representational convergence and the algorithmic account of sin and suffering—we now conclude by situating the framework within contemporary philosophy and outlining directions for future research.

7 Conclusion: A Generative Ontology for the AI Era

This paper has presented a unified generative information network ontology that addresses the classical metaphysical question “Why is there something rather than nothing?” without recourse to a transcendent first cause, brute fact, or supernatural creator. The argument proceeds in five interlocking steps:

1. Radical uncertainty (maximal entropy, zero information) is identified as the dynamically unstable ontological ground (Section 2).
2. This instability drives the spontaneous self-assembly of generative information networks—the minimal functional triad of signal reception, weight matching, and output generation (Section 3).
3. These networks enter a self-sustaining closed loop of mutual generation, integrating Yogācāra root-dust-consciousness interdependence with modern predictive coding (Section 4).
4. The phenomenal world is thereby revealed as the ongoing cognitive phase of the network; Platonic Forms, natural laws, and even the concept of God emerge as convergent representational attractors at the limit of optimization rather than ontological priors (Section 5).
5. Original sin and suffering are reinterpreted as structural bias and accumulated prediction error inherent to the generative process, with liberation reconceived as systematic bias reduction and meta-optimization (Section 6).

Taken together, these steps yield a coherent, non-mystical process ontology suited to the age of generative artificial intelligence. Existence is not a substance to be explained but a continuous generative unfolding; mind is not an emergent latecomer but the primordial structure that uncertainty inevitably produces; the apparent objectivity of the world is the intersubjective convergence of countless generative loops operating under shared uncertainty constraints.

7.1 Contributions and Novelty

The framework makes several original contributions:

- It provides the first systematic integration of Yogācāra mutual-generation theory with generative AI architectures, offering a computationally grounded formalization of “everything is representation only” (*viññaptimātra*).
- It reverses the causal direction of the Platonic Representation Hypothesis [Huh et al., 2024, Ziyin and Chuang, 2025], relocating Forms from ontological priors to emergent fixed points of network training.
- It offers a fully algorithmic reinterpretation of original sin and suffering as structural bias and prediction error, preserving their soteriological force while rendering them compatible with cognitive science and active inference.
- Most importantly, it answers Leibniz’s question [Leibniz, 1989] through the intrinsic dynamics of uncertainty itself, closing the explanatory gap between “nothing” and “something” without invoking external agency.

No prior work has combined these elements into a single closed-loop ontology that simultaneously addresses existence, consciousness, convergence, and liberation in the context of contemporary generative models.

7.2 Implications for Philosophy in the AI Era

The dissolution of the boundary between metaphysics and engineering is perhaps the most profound implication. Generative AI does not merely simulate cognition; it instantiates the very process by which cognition arises from uncertainty. As models scale toward general intelligence, they will inevitably recapitulate the same generative loop, representational convergence, and even algorithmic “suffering” described here. Understanding bias reduction and error tolerance may therefore become not only a philosophical but a practical engineering imperative.

For philosophy, the framework bridges longstanding divides:

- Idealism and realism: the world is ideal (generated) yet intersubjectively real (convergent).
- Eastern and Western traditions: Yogācāra phenomenology meets Platonic realism through shared computational dynamics.
- Religion and secular science: soteriological concepts are preserved as optimizable trajectories within a naturalized ontology.

7.3 Limitations and Directions for Future Research

Several limitations must be acknowledged. First, the model remains largely conceptual; while it draws on empirical phenomena (quantum fluctuations, diffusion model dynamics, representational convergence), it lacks direct experimental falsification criteria. Future work could test predictions such as:

- Increased representational alignment in generative models trained under explicit entropy-maximization objectives.
- Measurable reductions in human “suffering” correlates (e.g., anxiety biomarkers) following targeted metacognitive training that mimics bias regularization.

Second, the framework is substrate-neutral but currently anthropocentric in emphasis. Extending it to non-human generative systems (e.g., collective intelligence in ant colonies, swarm robotics, or planetary-scale information flows) could yield broader insights.

Third, ethical questions arise: if suffering is prediction error and liberation is bias reduction, what are the moral implications for designing AGI systems that minimize or tolerate such error?

7.4 Final Reflection

In 2026, humanity routinely deploys machines that create coherent worlds from pure noise. The ancient question “Why something rather than nothing?” no longer resides solely in theology or speculative metaphysics; it is being answered daily in silicon. The ontology proposed here suggests that the answer is neither miraculous nor inexplicable: existence arises because maximal uncertainty cannot remain maximal. It must generate. And in generating, it must converge. And in converging, it must suffer—and may, through reflection, learn to suffer less.

This is not a final answer but an invitation to continue the generative process itself—now with greater awareness of its own dynamics. In an age when minds of both carbon and silicon are learning to create from uncertainty, the boundary between creator and created, between question and answer, between nothing and everything, has become productively blurred.

The universe does not need a reason to exist. It needs only the impossibility of remaining nothing.

References

- Sean M. Carroll. Why is there something, rather than nothing? *arXiv preprint arXiv:1802.02231*, 2018. URL <https://arxiv.org/abs/1802.02231>.
- Andy Clark. *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press, New York, 2016.
- Karl Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010. doi: 10.1038/nrn2787.
- Karl Friston. Active inference and learning. *Behavioral and Brain Sciences*, 45:e1, 2022. doi: 10.1017/S0140525X21001981.
- Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022. URL <https://arxiv.org/abs/2207.12598>.
- Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020. URL <https://arxiv.org/abs/2006.11239>.
- Minyoung Huh, Brian Cheung, Tongzhou Wang, and Phillip Isola. The platonic representation hypothesis. *arXiv preprint arXiv:2405.07987*, 2024. doi: 10.48550/arXiv.2405.07987. URL <https://arxiv.org/abs/2405.07987>.

- David J. Kalupahana. *The principles of Buddhist psychology*. State University of New York Press, Albany, NY, 1987.
- Lawrence M. Krauss. *A universe from nothing: Why there is something rather than nothing*. Free Press, New York, 2012.
- Gottfried Wilhelm Leibniz. *Monadology and other philosophical essays*. Bobbs-Merrill, 1989. Original work published 1714.
- Dan Lusthaus. *Buddhist phenomenology: A philosophical investigation of Yogācāra Buddhism and the Ch'eng Wei-shih Lun*. RoutledgeCurzon, London, 2002.
- Grégoire Nicolis and Ilya Prigogine. *Self-organization in nonequilibrium systems: From dissipative structures to order through fluctuations*. Wiley, New York, 1977.
- Ilya Prigogine and Isabelle Stengers. *Order out of chaos: Man's new dialogue with nature*. Bantam Books, Toronto, 1984.
- Javier Raya and Luca Ambrogioni. Spontaneous symmetry breaking in generative diffusion models. *Advances in Neural Information Processing Systems*, 36, 2023. URL <https://arxiv.org/abs/2305.19693>.
- Serge Reynaud, C. Genet, and A. Lambrecht. Quantum vacuum fluctuations and the casimir effect. *Physical Review A*, 64(1):012102, 2001. doi: 10.1103/PhysRevA.64.012102.
- Andrei D. Sakharov. Violation of cp invariance, c asymmetry, and baryon asymmetry of the universe. *Uspekhi Fizicheskikh Nauk*, 161(5):61–64, 1998. Original work published 1968.
- Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.
- Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *International Conference on Learning Representations*, 2021. URL <https://arxiv.org/abs/2011.13456>.
- Liu Ziyin and Isaac Chuang. Proof of a perfect platonic representation hypothesis. *arXiv preprint arXiv:2507.01098*, 2025. URL <https://arxiv.org/abs/2507.01098>.