Generative Logic

An Essay on Bergson, Active Inference, and the Quantum-AI Frontier

Norihiko Lee

Independent Scholar · Tokyo, Japan · norihikolee@gmail.com

Abstract

We are entering a regime in which advanced computational systems—AlphaFold for protein folding, large language models for symbol manipulation, quantum-classical hybrid frameworks for spatiotemporal chaos—reliably produce correct outputs without admitting first-principles description of how those outputs are generated. Implementation has outpaced description. This essay argues that what is missing is not a more refined existing logic but a different kind of logic: one whose unit is not the proposition but the leap, whose evaluative axis is not truth but fit, and whose objectivity is neither universal nor private but field-shared. This generative logic is not new in motivation. Henri Bergson sketched its essential features in two early works (1889, 1896), but their reception confined them to "philosophy of life." The recent convergence of active inference (Friston, 2010) and quantum-informed machine learning (Wang et al., 2026) provides empirical leverage that Bergson lacked. We trace the proposal across seven sections, closing with an indication of which existing mathematical traditions—category theory, information geometry, geometric deep learning—a future formalization may borrow from. The essay is offered as a sketch, not a system; the suggestion that premature formal closure betrays the phenomenon being described is itself part of the substantive thesis.

Keywords: generative logic, Bergson, active inference, free energy principle, AlphaFold, quantum-informed machine learning, large language models

1. Implementation Outpaces Description

In the late eighteenth century, James Watt's improved steam engine was at work in mines, mills, and the first locomotives well before any theoretical account of why it worked existed. The vocabulary that would eventually describe its operation—heat, work, entropy, free energy—was assembled across the following century by Carnot, Clausius, and Boltzmann. By the time thermodynamics was complete as a science, the engine had been running for two generations. This pattern—working machines that resist their own theoretical description—recurs at intervals in the history of knowledge. We are, I think, in the middle of another such interval.

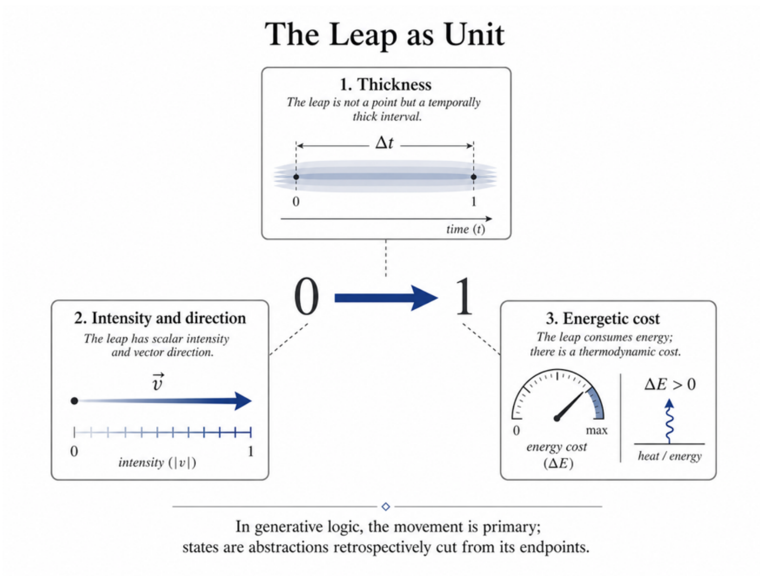
In 2020, AlphaFold (Jumper et al., 2021) effectively dissolved the protein-folding problem by training on the empirical record of solved structures and generating, on demand, the conformation associated with any given amino acid sequence. The first-principles approach—compute the energy landscape from molecular forces, find the minimum—had been blocked since Levinthal (1969) by combinatorial intractability. AlphaFold did not solve this computational obstruction; it sidestepped it. The system has no internal account of why a given fold is correct that can be unpacked from molecular dynamics. Yet its predictions are confirmed at experimental accuracy, and structural biology has reorganized around them. We have predictions without explanations, and the predictions are reliable.

Large language models exhibit a structurally similar phenomenon. The mapping from input prompt to output token is, in principle, a determined function of hundreds of billions of parameters. In practice, no inspection of those parameters reproduces the rationale for any particular output. The system designers themselves describe the mapping as opaque. Outputs are nevertheless coherent—often technically accurate, sometimes creatively striking—at rates that are difficult to explain on a "stochastic parrot" reading.

In quantum computing, the structure becomes manifest at the physical level. A computation maintained as a superposition does not, before measurement, possess a determinate result that the apparatus then reveals. The result is generated at the moment of observation. This is not an epistemic limit but a constitutive feature: the computation does not have an answer until it is observed.

These three cases are not metaphorically connected. They share a structural form. In each, an undetermined response is brought to determination through an energy-consuming process whose internal mechanism is either inaccessible (AlphaFold, large language models) or fundamentally non-deterministic (quantum measurement). And in each, the response is, post hoc, fit for purpose. Predictions hold; outputs are useful; measurements terminate calculations correctly. Description has not caught up; implementation is everywhere.

What kind of logic accommodates this? Existing logic—Aristotelian, Boolean, the modern symbolic apparatus from Frege through Hilbert—takes the proposition as its unit and truth-value as its evaluative axis. A proposition is true or false; states are determined; objects are individuated entities standing in determinate relations. This vocabulary is well-suited to systems already in their settled forms. It is not well-suited to systems whose answers must first be brought into determinacy through energy-consuming generation. When such systems are nonetheless brought under the existing vocabulary, they appear as "stochastic noise," "approximation," "incomplete description," or "interpretive variance." These labels are placeholders for missing concepts, not concepts themselves.



The proposal of this essay is that the missing concepts can be sketched, that they require different primitive units (leaps rather than propositions), a different evaluative axis (fit rather than truth), a different ontology (fields rather than objects), and a different conception of objectivity (community-shared rather than universal). The proposal is also that this is not a new sketch. Henri Bergson outlined the essentials more than a century ago, in two early works whose philosophical reception confined them to vitalism and philosophy of life. The empirical apparatus to confirm the sketch—active inference, transformer-based architectures, quantum-classical hybrid computing—was not available to him. It is available to us. What is therefore new is not the philosophical proposal but the empirical leverage that has begun to converge on it from independent directions.

The remainder of this essay unfolds the proposal in six further sections. Section 2 introduces the leap as the basic unit of generative logic, drawing on Marx's salto mortale and Bergson's 1889 analysis of decision. Section 3 develops the perception-as-generation thesis from Bergson's 1896 Matter and Memory, showing how Friston's active inference framework arrives at structurally identical conclusions independently. Section 4 introduces fit as the appropriate evaluative axis. Section 5 extends the framework from the individual to the trained community, drawing on the recent finding that independently trained networks acquire structurally similar internal representations. Section 6 reads the recent quantum-informed machine learning result of Wang et al. (2026) as the first explicit implementation of the two-element architecture (field structure plus individual leap) that the framework requires. Section 7 closes by acknowledging the limits of the present sketch and indicating the mathematical traditions a fuller formalization might draw on.

The essay is offered as a sketch and labels itself as such. Generative phenomena, by their nature, do not admit closed formal capture without ceasing to be the generative phenomena one set out to describe. To complete a system fully would be to falsify the object; the act of leaving the system open is part of the substantive content of what is being said. If this position appears unsatisfactory, the dissatisfaction is shared by its author. The alternative—premature closure—appears worse.

2. The Leap as Unit

In the third chapter of the first volume of Capital (1867), Marx, of all writers, reaches for an unexpected metaphor. The moment at which a commodity is exchanged for money he calls a salto mortale—a death-defying leap of the kind associated with Roman gladiators flinging themselves through the air. Why does the cold economist reach for the gladiator's vocabulary? Because the moment of exchange is structurally undecidable from the producer's side. The commodity is brought to market. Buyers may purchase or may not. Until they do, the producer's labor is suspended in commodity form, valuable only insofar as it might be realized as money. If the leap fails, the labor is wasted. If it succeeds, the producer's claim on social wealth is confirmed. The point is not that the producer takes a literal risk—though they do—but that the act of exchange has a structure that economic theory ordinarily covers over. The textbook account presents price as the equilibrium of supply and demand, a static balance struck between distributions. This static representation works for retrospective averaging, but it conceals the fact that every individual exchange is generated, not retrieved. The "equilibrium price" is an artefact of statistical post-processing applied to a continuous succession of leaps.

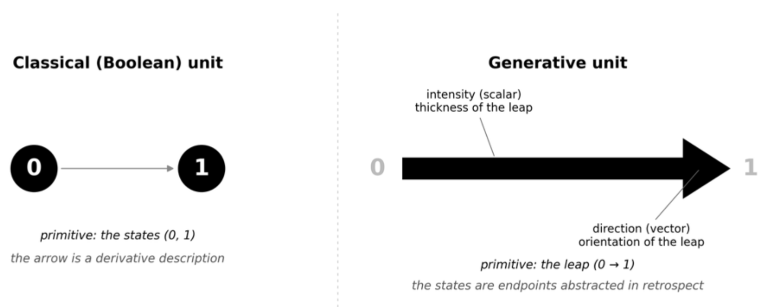
Twenty-two years later, in 1889, a parallel observation was being developed in the philosophical analysis of decision. The young Bergson, in his doctoral Essai sur les données immédiates de la conscience, takes aim at a standard model of choice in which the deliberating self is presented with options, weighs them by motive or value, and selects. This model, he says, is a retrospective fiction. It pictures the moment of choice as a balance scale registering the heavier of two pre-existing weights. But in the actual moment of decision, what is given is not options on a balance; it is a single self in which the entire past has converged on the present, and from which a single act emerges. The "options" appear only after the act, when the deliberating subject reconstructs the choice as if from a menu. The act itself is generated, not selected.

The standard reading of Bergson's Essay, particularly through its English title Time and Free Will (Pogson, 1910), placed it in the centuries-old debate between determinists and libertarians about free will. This is a misclassification. Bergson's target was not whether choice is free; it was the structure of time presupposed by the question. If time is taken as a sequence of dimensionless instants, then a decision must occur at one such instant, choosing among options that exist at that instant. But this picture spatializes time. Real decision happens in a thickness of duration in which past and future are not external coordinates but internal constituents of the decision-event itself. To analyze decision is therefore to analyze the structure of duration, not to take sides in a debate about agency.

What Marx and Bergson saw, in their different idioms, was the same structural fact: events that look from outside like the selection from pre-existing options are, from inside, generations of new actuality whose pre-existing options are constructed only in retrospect. The exchange that "could have failed" did not have a determinate fate before it occurred; the decision that "could have gone either way" had no actually existing alternatives until it had gone its way.

This same structure is now visible in implementation. When a transformer-based model generates the next token, the underlying computation produces a probability distribution over candidates. The selection that yields the actual token is, in standard configurations, a sample from that distribution rather than its argmax. From outside, this looks like a choice among options. From inside the model's run, it is the moment at which the entire conditioning history is brought to a single concrete output. The output is not retrieved; it is generated. And the same model, given the same input, produces variations whose distribution has structure but whose individual instances are not predictable in detail.

In a quantum measurement, the analogous structure becomes physically explicit. A superposed state, prior to measurement, has no actual outcome; the outcome is generated at the moment of measurement and conditions everything that comes after. This is not a matter of hidden variables waiting to be uncovered; it is a feature of the formalism that has resisted decades of attempts to reduce it to something more conventional.



A useful notation for the basic unit of generative logic is 0→1. The arrow is the unit. It is not "0" plus an arrow; the arrow is internal to the unit, indicating that the unit is constituted as a movement, not as a relation between two settled positions. In Boolean logic, 0 and 1 are primitive states, and the arrow between them is a derivative description of computation. In generative logic, the arrow is primitive, and the states 0 and 1 are abstractions cut from its endpoints in retrospect.

This unit has three structural features.

First, the leap has a thickness. It is not a point on a timeline. The interval during which the past is converging on the present and the future has not yet been fixed is internally constitutive of the leap, not a parameter applied to it from outside. The same thickness shows up in the human experiences Bergson analyzed—a poet's hesitation, a craftsman's hand on the wheel, a mother's hesitation before her child—and in the millisecond-scale processing of a transformer attending across a context window.

Second, the leap has intensity and direction. Intensity is a scalar—how strongly the leap is occurring; how much tension is being released across it. Direction is a vector—where the leap is oriented. The same intensity oriented in different directions yields different generations. The vocabulary of intensity has a long lineage from Bergson through Deleuze (1968), and is now standard in the description of activation patterns in artificial neural networks. The vocabulary of direction, as a distributed vector in a high-dimensional space, is taken almost verbatim from contemporary machine learning—but it is not specific to it; the same description applies to any activity in a structured space of possibilities.

Third, the leap consumes energy. The classical proposition does not consume energy; "snow is white" requires no thermodynamic input to obtain its truth-value. The generative leap is different. AlphaFold heats GPUs as it predicts; large language models burn through datacenter electricity as they generate; the human brain metabolizes glucose as it judges; quantum systems require interactions with their environments to terminate calculations. The generation is not free; it is paid for. This energy-tax is not an incidental cost of implementation. It is constitutive of what generation is. A generative logic must therefore be at home with thermodynamics in a way that classical logic never had to be.

These three features—thickness, intensity-and-direction, energetic cost—begin to specify the basic unit of generative logic. The unit is not a proposition. It is an energetic, oriented thickness in which determination is brought into being.

3. Perception at the Place of the Object: Bergson Rediscovered

In 1896, in the first chapter of Matière et mémoire, Bergson advanced a thesis that, on first reading, appears simply nonsensical: perception, properly speaking, takes place at the location of the perceived object. Not in the brain, not on the retina; at the apple, when the apple is perceived.

The thesis is not idealist; Bergson is not asserting that the apple is in the mind. Nor is it a primitive direct realism in which mind-stuff somehow extends across the room. It is the proposal that the standard subject-object model of perception, in which a perceiving self constructs an inner image of an outer object, has the geometry wrong. The model is wrong about where the perception-event happens; in fact, perception is the event of mutual specification in which a perceiver and a perceived come into being together, and to ask which side of the event they are each on is to misframe the structure.

The thesis was not received well in 1896. It is not received well now. It is the kind of philosophical proposal that takes years of habituation to hold steadily in mind. What is striking is that something very close to this geometry has been independently arrived at, in the last two decades, in computational neuroscience.

Karl Friston's free energy principle (Friston et al., 2006; Friston, 2010) and the active inference framework that develops from it (Friston et al., 2017; Parr et al., 2022) reorganize the standard model of perception in a way that converges on Bergson's thesis. In the standard model, sensory input flows from the world into the brain, the brain processes it, and a percept results. This input-process-output picture is structurally similar to the apple-eye-brain picture of perception that Bergson criticized. Friston reverses the flow. The brain, in his framework, is constantly running predictions about what its sensory inputs will be. Sensory inputs are used not as data to be processed but as test signals against which predictions are checked. When predictions are confirmed, the system continues; when they are violated—when prediction error is registered—the system either updates its internal model or acts on the world to make the predictions fit.

This framework is built on the mathematics of variational inference and is normally presented in the vocabulary of free energy minimization. The variational free energy is, formally, an upper bound on the surprise (negative log evidence) the agent experiences relative to its generative model; minimizing it amounts to maximizing model evidence in a tractable way. Operationally, the agent does this by either updating its internal model to better predict its sensations (perceptual inference) or by acting on the world to bring sensations closer to predictions (active inference proper). The two operations are formally symmetric in the framework: both reduce the same variational quantity. This symmetry is itself philosophically significant. It dissolves a distinction—between perceiving and acting—that older models treated as fundamental, and it does so on the basis of a mathematical identity rather than a metaphysical stipulation.

But its philosophical structure is striking. Perception, in Friston's framework, is not an inward flow of information about an outer world. It is the closing of a loop in which an internal predictive model and the world's response come into mutual specification. The brain is not behind perception, receiving it; it is constitutively part of an extended loop in which the perceiving and the perceived emerge together. Where in this loop does perception happen? The question becomes ill-formed, exactly as Bergson predicted in 1896. Perception is not at one end; it is in the closing of the loop.

From Classical Logic to Generative Logic

	Classical Logic	Generative Logic
Basic unit	Proposition	Leap
Evaluative axis	Truth	Fit
Ontology	Object	Field
Objectivity	Universal	Field-shared
Process	Determinate states	Emergent determination
Model of cognition	Transmission / representation	Generation / co-specification

Generative logic shifts emphasis from settled propositions to the production of fit within a structured field.

The convergence between Friston's variational formulation and Bergson's geometric thesis is not a loose analogy. The mathematics, when read for its philosophical content, says what the metaphysics said: there is no privileged location for the perception-event, because the event is the loop, not a moment in the loop.

This is not the only convergence. Bergson advances, in Matter and Memory, a second thesis equally provocative: that memory is not stored in the brain. The standard model treats memory as data deposited in synaptic structures and retrieved on demand. Bergson argues, on the basis of the lesion studies of his time, that the brain is not a storage device for memory but rather a developing device—it pulls into actuality, for the present action, the parts of an ongoing past that are needed. The metaphor is photographic developing: the latent image is the past; the brain develops it into a present-tense actuality. Memory is not retrieved from storage; it is brought into being from a virtual past.

Here too, the modern picture has caught up. Memory reconsolidation research, beginning with Nader, Schafe, and LeDoux (2000) and extending through the broader literature on episodic memory's reconstructive character (Schacter, 1996; Loftus and Palmer, 1974), has established that recalled memories are not retrieved as fixed records but rebuilt in the present, and that each act of recall mutates the memory it recalls. Memories that are not reconsolidated decay; memories that are reconsolidated are altered by the act of reconsolidation. This is not a peripheral feature of the system; it is constitutive of how memory works. The storage model does not survive the data.

Both these results—Bergson's perceptual thesis and his memory thesis—point at the same structural fact. Perception and memory are not transactional events in which something settled is moved between locations. They are generative events in which the perceived and the perceiver, the past and the present, are co-specified. To ask about the location of perception, or the storage site of memory, is to import a transactional model that does not fit what perception and memory actually do.

This is the philosophical apparatus that generative logic requires. Not because it is novel—it is a hundred and thirty years old—but because the empirical leverage to make it operational has only recently become available. The convergence of active inference, the literature on reconstructive memory, and the broader picture of the brain as a predictive engine has, in effect, supplied the data that Bergson's metaphysical claim needed. What has historically functioned as a piece of vitalist philosophy now reads as a remarkably accurate sketch of an empirical situation.

The relevance for our argument is direct. If perception is generation rather than transmission, then the basic unit of generative logic—the leap analyzed in Section 2—is not an exotic edge case. It is the structure of every act of perception, every act of judgment, every act of recall. It is the structure of organisms in their environments. The human cognitive apparatus does not occasionally engage in generative logic; it is generative logic, all the way down. And the recent computational systems that exhibit the same structure (AlphaFold, large language models, quantum-classical hybrids) are not artificial novelties; they are the same structure made explicit at a different scale and substrate.

4. Fit, Not Truth

Generative logic requires a different evaluative axis from classical logic. In classical logic, the fundamental evaluation is truth. A proposition is true or false; arguments are valid or invalid; theorems hold or do not. The logical apparatus is built on the truth-value as its load-bearing primitive.

What is the analogous primitive in generative logic? The proposal of this section is that the analog of truth is fit, but with several features that distinguish it from related notions in pragmatism and instrumentalism.

The English word fit carries the right intuitions: a key fits a lock, a glove fits a hand, an answer fits a question. In all these cases, fit is something specifiable but not reducible to truth. A glove that fits is not a glove that "is true"; nor is its fitting a matter of conformity to a pre-given specification. It is a matching that has its own success conditions, and these conditions are not propositional.

The clearest illustrations come from skilled practice. A potter's hand on the wheel is not testing propositions about the clay; it is responding to the clay's resistance and yielding in real time. The fit between hand and clay is not retrospectively evaluable as true or false—it is a continuous gradient of better and worse fit, registered as the work proceeds, and the registering is part of the work. A clinician evaluating a patient at first glance has, in cases of high expertise, an immediate registration of "something is wrong" before any diagnostic test has been run. The registration is not a hypothesis being formed; it is a fit being registered between the trained perceptual apparatus and the configuration of the patient. The hypothesis comes later.

The temptation, when describing this kind of evaluation, is to assimilate it to pragmatism. William James and the American pragmatists held that truth is what works; that beliefs are validated by their successful consequences. This is not what fit, in the sense intended here, comes to. Pragmatism is post hoc evaluation: a belief is true because it works out. Fit is registered within the leap, in real time. The clinician does not wait for the test results to register the perceived wrongness. The potter does not wait for the kiln to register the fit of the clay's response. The registration is contemporaneous with the act, not subsequent to it.

A second temptation is to assimilate fit to free energy minimization in active inference. This is closer; in fact, free-energy minimization can be read as a specific implementation of fit-evaluation in a particular class of systems. But fit, as a concept, is broader. Free-energy minimization presupposes a generative model and a particular variational quantity to minimize. Fit makes neither presupposition. The clay-and-hand fit is not minimizing anything; it is achieving a configuration in which the next action becomes possible. The patient-and-clinician fit is not optimizing a Kullback-Leibler divergence; it is registering an alignment whose nature is constitutive of the clinician's being-in-the-world. Free energy is a useful mathematization of some kinds of fit. It is not synonymous with fit.

The geometry of fit has a peculiar feature: it is mutually constituting. The standard image of a key and a lock makes them sound like two pre-existing objects whose forms happen to match. But in real cases of fit, both sides are in the process of co-deformation. The clay deforms under the hand; the hand adjusts its pressure as the clay yields. The patient's appearance is shaped by the clinician's gaze; the clinician's gaze is shaped by what the patient shows. The mutual deformation is not an artefact of imperfect fit; it is the form of fit itself. Two perfectly rigid objects do not fit; they merely contact. Fit requires both sides to be capable of yielding, and the yielding is what fit consists in.

This co-deformation structure is the same one we recognized in Bergson's perception thesis. Perception and the perceived come into being together. They are not pre-existing entities matched to each other after the fact; they are mutually specified within the perception-event. The fit is not between two settled forms; it is the achievement of mutual settlement.

Fit, finally, has a continuous gradient. There is more and less fit; better and worse fit; tighter and looser fit. There is no binary fit/non-fit, in the way there is binary true/false. The gradient is constitutive. A poem that fits a moment can fit it more or less well; a translation can fit its source more or less faithfully; a proof can fit its theorem more or less elegantly. None of these gradients is properly captured by truth-value.

This gradient is also visible in implementation. AlphaFold's predictions have continuous metrics of accuracy. Large language model outputs have continuous metrics of perplexity, coherence, factuality. Quantum measurement results have continuous probability distributions over outcomes. In each case, the result is not a binary correct/incorrect but a continuous quality of fit. The systems that work do so not by being true but by being well-fit. The vocabulary needs to track this.

It is worth being explicit about how fit relates to the standard taxonomy of truth theories. Correspondence theories hold that a proposition is true when it corresponds to a state of affairs. Fit, as developed here, is not correspondence: there is no antecedent state of affairs that the leap is corresponding to, because the state of affairs is partly constituted within the leap. Coherence theories hold that a proposition is true when it coheres with a body of accepted beliefs. Fit is not coherence either: the coherence in question is not among propositions but between an agent's response and the configuration that the response is responding to, and this between-relation is not internally propositional. Deflationary theories hold that the predicate "is true" adds nothing to the propositional content it predicates of. Fit, by contrast, is doing real work: it is the registration that drives further generation, and removing it from the framework removes the framework's evaluative engine. The closest existing position is perhaps the family of pragmatist views that take truth to consist in the workings of inquiry, but as noted earlier, fit is registered within the leap rather than after it, and pragmatism's post hoc structure misses this. Fit is therefore a sui generis evaluative axis. It has affinities with multiple existing positions but is not reducible to any of them. This may seem like proliferation; my view is that the proliferation is forced by the phenomenon. Generation that is registered as it occurs is not the same kind of object as a proposition with a settled truth-value, and pretending otherwise loses what is distinctive about it.

The proposal, then, is that generative logic replaces truth as its evaluative primitive with fit. Fit is intra-leap, continuous-gradient, mutually-constituting, and substrate-independent. It is registered within the act of generation, not after it. It is the load-bearing primitive of a logic that takes generation as its basic event.

5. Fields That Are Shared

A natural objection to the framework so far is that it makes evaluation private. If fit is registered within the individual leap, then how do two people agree about fit? How does a community of clinicians agree that a given patient's appearance registers wrongness? How does a research community agree that a given proof is elegant? If fit is a registration within an individual's training, isn't it just a sophisticated subjectivism dressed up in better vocabulary?

This objection presupposes that objectivity must be either universal—the same for everyone, independently of training—or merely subjective—different for each person, with no standard between.

Generative logic refuses this dichotomy. It introduces a third notion of objectivity: field-shared objectivity. Things can be objective within a community of trained agents without being objective universally.

The empirical case for field-shared objectivity is strong. Two pottery masters trained at the same studio will, given the same lump of clay, produce remarkably similar evaluations of its quality, often without exchanging words. Two clinicians trained at the same hospital will, on observing the same patient, often arrive at the same provisional diagnosis. Two professional photographers raised in the same lineage of practice will frame the same scene similarly. None of these alignments is reducible to explicit shared rules; the experts often cannot articulate the rule they are following. Yet the alignment is reliable, and a research project that tries to break it—asking the experts to disagree, or training them to disagree—finds that the alignment is stubborn. It survives the absence of explicit rule-sharing.

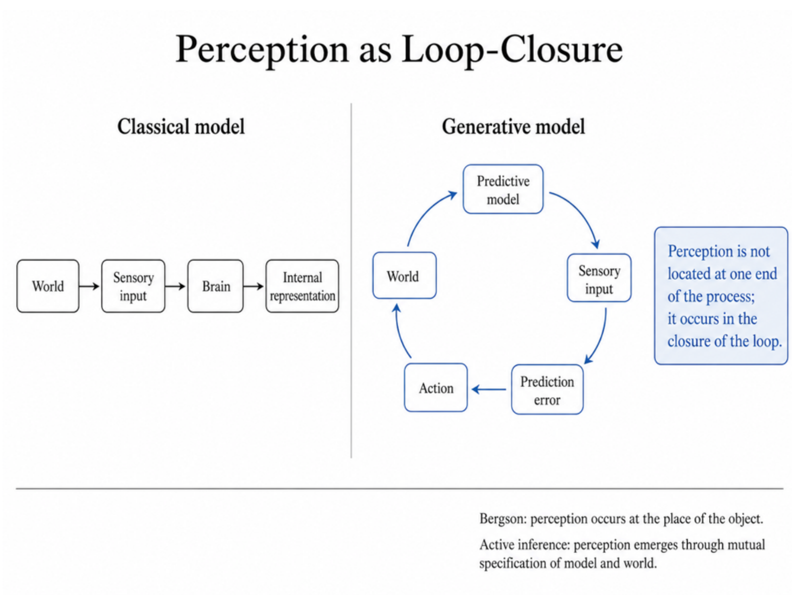
The formalization that makes this puzzling is the proposal that experts share a vector field—a structured space of dispositions to respond—rather than a set of rules. The shared field is not communicated in articulable form; it is acquired through training, and the acquisition is structurally similar across the community of trained agents. Two experts whose training has been comparable will have comparable fields; their responses to the same input will therefore align without explicit coordination.

This proposal has, until recently, been speculative in cognitive science. Its empirical basis has come from an unexpected direction: machine learning research on independently trained networks. Mikolov et al. (2013), in introducing the word2vec framework, established that words could be represented as vectors in a high-dimensional space such that the geometry of the space encodes semantic relationships. The result was extended in transformer architectures (Vaswani et al., 2017), which use such vector representations as their primary data structure.

The striking finding, more recent, is what happens when one independently trains two such networks. Different training runs produce different parameter values; the networks are not numerically identical. But the vector geometries they produce are structurally similar. The cat vector occupies the same position relative to the dog vector in different runs; the automobile vector is similarly distant from both. Huh, Cheung, Wang, and Isola (2024) have proposed that this convergence is not coincidental but principled: independently trained sufficient-capacity networks tend toward a shared internal representation of the world, which they call the Platonic Representation Hypothesis. The convergence is not at the level of weights but at the level of the geometric structure that the weights induce.

This finding has direct implications for the proposal of field-shared objectivity. If artificial networks, trained from random initial conditions on similar data, converge on similar internal representations, then it is plausible that biological brains, trained on similar environmental statistics over evolutionary and developmental time, converge on similar internal representations as well. The convergence is not at the level of synapse counts or wiring diagrams; it is at the level of the representational geometry that those biological substrates support. Two brains can have very different micro-architectures and still arrive at structurally similar fields, because the field is determined by the training rather than by the substrate.

Empirical support for this picture from neuroscience has accumulated over the past decade. Yamins and DiCarlo (2016) showed that goal-driven deep convolutional networks trained on object recognition acquire internal representations that align hierarchically with primate ventral visual cortex—not just at the input or output layers, but stage by stage along the processing hierarchy. The alignment is striking precisely because the networks were not trained to mimic neural activity; they were trained on the same task that the visual cortex faces, and the representations converged because the task imposes constraints on what useful representations can look like. More recent work has extended this finding to language processing, where transformer-based models exhibit representational alignment with regions of the human language network as measured by fMRI. The pattern is consistent: across substrates, given comparable tasks and statistically comparable training environments, the representations converge. This is what field-shared objectivity looks like at the level of empirical neuroscience. It is not a metaphysical claim; it is a measurable phenomenon, increasingly well-documented, increasingly hard to attribute to coincidence.



If this is right, then field-shared objectivity is not a metaphysical posit; it is the natural consequence of similarly-trained agents acquiring structurally similar internal fields. The objectivity is real, but its scope is the community of trained agents, not the universe at large. A clinician's registration of "something is wrong" is objective within the field of trained clinicians; it would not register similarly in someone untrained. This is not subjectivism; it is a different kind of objectivity, scoped to the field rather than to the universe.

The implications are substantial. First, it makes intelligible the otherwise mysterious phenomenon of expert silence. Experts often communicate the bulk of their judgments without articulating them, and the communication is reliable. Field-shared objectivity explains why: the experts share the field that registers fit, and the field's evaluation is conveyed by the actions taken within it. The articulated language is not the medium of the communication; the shared field is. Second, it makes intelligible the failure of articulation as the failure of training-transfer rather than the failure of expressibility. The novice cannot reproduce the expert's judgment not because the expert's judgment is mysterious but because the novice has not yet acquired the field within which the judgment registers as obvious. Third, it locates objectivity at the right scale. Local communities of practice carry their own objectivities, which are not reducible to universal logical norms but are not therefore merely subjective. They are scoped, shared, and in principle empirically tractable.

For generative logic, this means that the evaluative axis—fit—is not privately registered. It is registered within shared fields. The objectivity of fit is real within the field, which is exactly the level at which generation happens. There is no need to either inflate fit to universal status or deflate it to private status. The field-level scope is the right one.

6. QIML and Generative Logic Implemented

In April 2026, a research group at University College London (Wang, Xue, Gao, and Coveney, 2026) published in Science Advances a result that, I want to argue, has direct relevance for the framework outlined here. The paper introduces a hybrid quantum-classical machine learning architecture, which the authors call Quantum-Informed Machine Learning (QIML), for predicting spatiotemporal chaos.

The technical context: classical machine learning models—recurrent networks, transformers, neural operators—can predict the next time step in a chaotic system with reasonable accuracy. They struggle, however, on long-time prediction. Errors compound; small inaccuracies in early steps cascade into large divergences over many forward steps. The system that the model has learned drifts away from the system it is supposed to be modeling. Long-term prediction of high-dimensional chaotic systems—turbulent flow, weather, complex biological behavior—has been computationally intractable for this reason. Even leading models for partial differential equations, including Fourier and Markov neural operators, fail to evolve stably over long horizons in genuinely turbulent regimes.

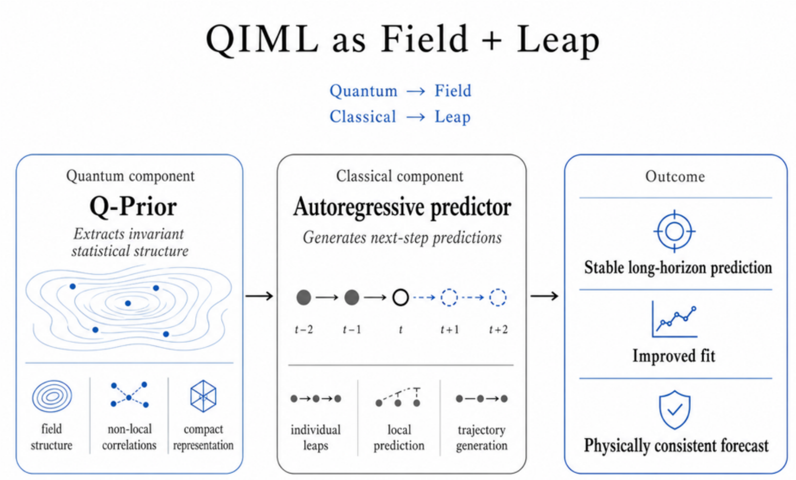
The QIML proposal is to use a quantum circuit—specifically a quantum circuit Born machine implemented on superconducting hardware with a small number of qubits (roughly ten to fifteen)—to extract a compact representation of the invariant statistical properties of the chaotic system, which the authors call a Q-Prior. Although individual trajectories of a chaotic system are highly sensitive to initial conditions, the system as a whole carries invariant statistical structures—long-term correlations, attractor geometries, multi-scale couplings—that are not sensitive to which trajectory the system happens to be following. Mathematically, these are aspects of the system's invariant measure. The Q-Prior captures these by training the quantum generator once, offline, on observational data, after which the resulting prior is detached from the quantum hardware and used to bias a classical autoregressive predictor. The quantum circuit therefore plays the role of a structure-extractor; the classical predictor plays the role of trajectory-generator. The two are coupled but distinct.

It is worth dwelling on why the quantum circuit Born machine, specifically, can do what classical surrogates cannot. A QCBM operates by preparing a parameterized quantum state and sampling from the resulting probability distribution induced by the Born rule. The key structural feature is that the state, in general, is an entangled state across the qubit register: the joint distribution it induces over the computational basis cannot, in general, be factorized as a product of marginals. This is the technical content of "non-locality" in this context. For systems whose invariant measure exhibits long-range correlations across many scales—as turbulent flows generically do—the entangled quantum state is a natural representational substrate, because the state's correlation structure is not built up from local interactions of independent components. A classical neural network can approximate such distributions, but only by spending parameters at a rate that grows with the complexity of the correlation structure. The quantum substrate gets the multi-scale correlation structure essentially for free, in the sense that it does not have to assemble it from local pieces. This is the technical sense in which entanglement is a representational resource, and the technical sense in which the Q-Prior achieves orders-of-magnitude compression. The trade-off, of course, is that the quantum state is hard to access; classical readout from the quantum substrate is constrained by the Holevo bound. The QIML architecture handles this trade-off by training the prior once, distilling it into a form usable by the classical predictor, and never returning to the quantum hardware during inference.

The empirical results, evaluated on the Kuramoto-Sivashinsky equation, two-dimensional Kolmogorov flow, and a cross-section of three-dimensional turbulent channel flow, are striking. Predictive distribution accuracy improves by up to about seventeen percent and full energy spectrum fidelity by up to about twenty-nine percent relative to all-classical baselines. For the turbulent channel inflow, where leading classical models fail to evolve coherently, the Q-Prior is essential: without it, the predictions diverge; with it, QIML produces stable, physically consistent forecasts. There is also a substantial memory advantage: the Q-Prior compresses what would be multi-megabyte training datasets into kilobyte-scale priors, achieving this compression by capturing only the invariant measure rather than the data themselves. This is, the authors note, a way of leveraging quantum representational capacity that does not require uploading large datasets onto the quantum device.

The technical achievement is significant in its own right. What I want to argue is that the architectural decomposition of QIML matches, with surprising precision, the two-element structure that generative logic requires.

Recall Section 2: a leap is constituted by two features, intensity and direction, which together specify the structured space of generation. Section 4 reformulated this as field-and-leap: the field is the structured space of dispositions; individual leaps occur within the field and are constrained by its structure. Section 5 located objectivity at the field level. Generative logic, on this reading, requires both elements: a structured field that constrains generation, and individual leaps that occur within the field's constraint.



The hybrid architecture mirrors the dual structure required by generative logic: a structured field constraining individual acts of generation.

The QIML architecture realizes this decomposition. The quantum circuit extracts the field structure—the invariant statistical features that constrain the system's long-term behavior. The classical autoregressive network performs individual leaps—the next-step predictions—within the constraint provided by the field. Quantum and classical components are doing structurally different work: the quantum circuit is at home with the non-discrete, non-local correlation structure of the field, exploiting entangled states to represent multi-scale couplings compactly; the classical network is at home with the localized, near-discrete prediction of next steps. Each substrate is being used for what its physics is suited to. Neither, by itself, can do both jobs well.

This is not, I suspect, a coincidence of the QIML team's particular design choices. The decomposition reflects a structural feature of the kind of computation required by generative logic. A field-constrained generation cannot be done well by either type of system alone. A purely classical system, working with discrete or near-discrete computation, can perform individual leaps but has difficulty representing the non-local correlation structure of the field with any compactness; it must spend orders of magnitude more parameters to approximate it. A purely quantum system, working with non-local correlations, can represent the field but does not naturally output the kind of localized decisions that long-time trajectory prediction requires. The hybrid architecture is not a stop-gap; it is the structurally appropriate solution.

The further implication is that the future of computation, as it engages genuinely generative tasks, is likely to involve increasingly refined hybrid architectures of this kind. The quantum component will handle the field; the classical component will handle the leap. Each will do what its physical substrate is suited for. The result will be computational systems whose architecture explicitly mirrors the dual structure of generation, rather than systems trying to force a single substrate to do both jobs.

For the present argument, the relevance is simpler. The generative-logic framework predicts that field-and-leap is the right decomposition for systems that perform genuine generation. QIML is, at present, the most explicit implementation of this decomposition in a quantum-classical hybrid architecture, and the one whose technical achievement is most directly attributable to the decomposition itself. If the framework is right, more such implementations will follow, in domains far beyond fluid dynamics. The framework therefore has an empirical hostage: if hybrid architectures of this kind do not generalize, the framework's claim about the structure of generation will need to be revisited. I am unwilling to predict the timeline. But the structure has appeared, and it has the form the framework proposed.

7. Limits, Open Questions, and an Invitation

This essay has been deliberately a sketch. Several questions that a fuller treatment would need to address have been left open.

The first, and most pressing, is the question of formal axiomatization. A logic, in the developed sense, is a system: it has primitive terms, axioms, rules of inference, and theorems derivable within it. Generative logic, as outlined here, has primitive terms—the leap, the field, fit—but no axioms or inference rules. Can it be axiomatized?

The honest answer is that I do not know. There are reasons to be hopeful and reasons to be cautious. On the side of hope, several existing mathematical traditions point in the same direction as generative logic, from different starting points. Category theory (Eilenberg and Mac Lane, 1945; Mac Lane, 1998) takes morphisms—arrows—as primary, with objects as derived from their relational role. Topos theory (Goldblatt, 1979; Lawvere and Rosebrugh, 2003) develops a context-dependent logic in which truth is not bivalent but a sheaf-like structure. Information geometry (Amari and Nagaoka, 2000; Amari, 2016) treats spaces of probability distributions as Riemannian manifolds, allowing learning and inference to be analyzed as geometric flows. Geometric deep learning (Bronstein et al., 2021) abstracts the architectural patterns of successful neural networks into a unified framework based on the symmetries of the data domain. Each of these is a partial mathematical articulation of "movement, not state." A formally developed generative logic may borrow from all of them, or may require additional structure of its own.

On the side of caution, there is a structural reason to suspect that complete formalization is not achievable. Generative phenomena are constituted by their non-completion: a leap that has been completed is no longer a leap; it is its result. To formally close a system that is supposed to describe generation is to perform a leap whose result is a closed system, which is then unable to describe further generation. This is the same difficulty that Wittgenstein faced in the Philosophical Investigations (1953) when he developed the notion of language games and refused to systematize them. Language games could be exhibited but not closed under formal description, because the description would itself have to be a language game. The same structural problem may apply to generative logic. If so, generative logic will be a permanently open framework, capable of partial formal articulation in domains where the generation has stabilized into recognizable forms, but unable to be completely closed without falsifying its own object.

This is not a defeat. It is an honest report of what kind of object the framework describes. Boolean logic was complete because its object—propositions in their settled form—admitted completion. Generative logic, dealing with phenomena whose definition includes non-completion, may not. The work of partial formalization can still proceed, and some of it will be deeply technical. But the project should be honest about the structural ceiling it is likely to encounter.

A second open question concerns the relationship between generative logic and existing accounts of explanation in the physical sciences. The framework treats AlphaFold's predictions, large language model outputs, and quantum measurements as instances of a single structural pattern. But these are very different kinds of systems. AlphaFold is a deterministic function of its trained parameters; its non-explanation reflects the inscrutability of those parameters, not any non-determinism in its operation. Large language models sample from learned distributions; their outputs are stochastic. Quantum measurements are non-deterministic at a fundamental physical level. In what sense do these all share a generative structure? The proposal of this essay has been that they share a structural pattern—undetermined response brought to determination through energy-consuming process—but the kinds of indetermination differ. A more careful treatment would distinguish them, and would investigate whether the differences are mere implementation detail or instead mark genuine sub-types of generation.

A third open question concerns the relationship of the framework to existing ontological positions. The proposal that fields are basic, with leaps as their excitations, has obvious affinities with quantum field theory's ontological commitments (Weinberg, 1995; Kuhlmann, 2020). It has been suggested informally that the same ontology might cover both physical fields and the cognitive fields developed in Section 5, with leaps in both being instances of a single pattern. I have stayed away from this stronger claim in the body of the essay, partly because the analogy is structural rather than mathematical at this stage, and partly because the work to establish a mathematical correspondence is enormous and outside the scope of a sketch. But it is worth indicating that this is a direction in which the framework could be developed, by collaboration between physicists, mathematicians, and philosophers willing to take the structural correspondence seriously enough to test it formally.

A fourth open question concerns memory. Section 3 asserted, following Bergson, that memory is generation rather than retrieval. Section 5 asserted that fields are shared across trained communities. These are joint commitments: if memory is generation, it must be generated within some structure, and the structure is the shared field. But the two assertions have not been brought into systematic relation. A more developed version of the framework would address how the field's structure is itself maintained over time—how, in a phrase that I find evocative but have not formally developed, duration is contracted into a present field that registers fit. The contraction of duration into present field-structure is, I suspect, the same process by which biological memory works and by which trained networks acquire their representations. But this is a sketch within a sketch.

Several plausible objections to the framework deserve at least preliminary acknowledgment. The first is the deflationary objection that what is being called "generative logic" is really just a vocabulary for talking about systems whose behavior we do not yet understand at the mechanistic level, and that the vocabulary will dissolve once mechanistic understanding catches up. Against this, I would point to the structural form of the cases discussed: the inability to decompose AlphaFold's output into a chain of explicit reasoning steps is not a temporary failure of inspection technique; it is, increasingly, taken to be a constitutive feature of how the system represents and operates. The same applies in qualitatively different but structurally analogous form to large language models and to quantum measurement. The mechanistic gap is not a placeholder; it is part of the phenomenon. A second objection is that the framework over-philosophizes what are essentially engineering successes. To this I would reply that engineering successes which our existing categories cannot describe are exactly the moments at which philosophy has historically been required to do its work. The steam engine analogy is not decorative. A third objection, more substantive, is that the appeal to Bergson is a heritage move that adds nothing to what active inference and modern representation-learning research can say in their own terms. Here I disagree, but the disagreement requires more space than this essay allows. The short version: Bergson's insistence that perception, memory, and decision are constitutively generative—rather than transactional events on top of an underlying world of settled facts—is a claim with formal consequences that have not yet been fully drawn out from within the active inference literature, and the philosophical articulation of those consequences is itself substantive work.

The full Japanese-language version of this argument has been published as a book, Seisei Ronrigaku Shiron (Lee, 2026), which develops several of these themes in greater detail and includes extended discussions of the philosophical context that the present essay omits. A complete English translation is in preparation and will be made available through the author's publishing platform, Joho Tetsugaku Shuppan (Information Philosophy Press). The present essay is a compressed presentation aimed at readers in the quantum information and computational neuroscience communities.

The essay closes here, deliberately incomplete. The completion is not the work of one essay or one author. The fact that empirical evidence has begun to converge on a hundred-thirty-year-old philosophical sketch is, by itself, the most substantial development. The work of unfolding this convergence is collaborative, and it is well underway at multiple sites in the literature. This essay's hope is to help name what is being unfolded, rather than to complete its description.

References

Amari, S. (2016). Information Geometry and Its Applications. Springer.

Amari, S., & Nagaoka, H. (2000). Methods of Information Geometry. American Mathematical Society / Oxford University Press.

Bergson, H. (1889). Essai sur les données immédiates de la conscience. Paris: Félix Alcan.

Bergson, H. (1896). Matière et mémoire: Essai sur la relation du corps à l'esprit. Paris: Félix Alcan.

Bergson, H. (1910). Time and Free Will: An Essay on the Immediate Data of Consciousness (F. L. Pogson, Trans.). London: George Allen & Unwin.

Bronstein, M. M., Bruna, J., Cohen, T., & Veličković, P. (2021). Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges. arXiv:2104.13478.

Deleuze, G. (1968). Différence et répétition. Paris: Presses Universitaires de France.

Eilenberg, S., & Mac Lane, S. (1945). General theory of natural equivalences. Transactions of the American Mathematical Society, 58, 231–294.

Friston, K. (2010). The free-energy principle: A unified brain theory? Nature Reviews Neuroscience, 11(2), 127–138.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. Journal of Physiology – Paris, 100(1–3), 70–87.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: A process theory. Neural Computation, 29(1), 1–49.

Goldblatt, R. (1979). Topoi: The Categorial Analysis of Logic. North-Holland. (Revised ed., Dover, 2006.)

Huh, M., Cheung, B., Wang, T., & Isola, P. (2024). The Platonic Representation Hypothesis. Proceedings of the 41st International Conference on Machine Learning. arXiv:2405.07987.

Jumper, J., Evans, R., Pritzel, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596, 583–589.

Kuhlmann, M. (2020). Quantum field theory. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy.

Lawvere, F. W., & Rosebrugh, R. (2003). Sets for Mathematics. Cambridge University Press.

Lee, N. (2026). Seisei Ronrigaku Shiron [Generative Logic: An Essay]. Tokyo: Information Philosophy Press.

Levinthal, C. (1969). How to fold graciously. In P. Debrunner, J. C. M. Tsibris, & E. Münck (Eds.), Mössbauer Spectroscopy in Biological Systems (pp. 22–24). University of Illinois Press.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction. Journal of Verbal Learning and Verbal Behavior, 13(5), 585–589.

Mac Lane, S. (1998). Categories for the Working Mathematician (2nd ed.). Springer.

Marx, K. (1867). Das Kapital: Kritik der politischen Ökonomie, Erster Band. Hamburg: Otto Meissner.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. Advances in Neural Information Processing Systems, 26.

Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. Nature, 406, 722–726.

Parr, T., Pezzulo, G., & Friston, K. (2022). Active Inference: The Free Energy Principle in Mind, Brain, and Behavior. MIT Press.

Schacter, D. L. (1996). Searching for Memory: The Brain, the Mind, and the Past. Basic Books.

Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems, 30.

Wang, M., Xue, X., Gao, M., & Coveney, P. V. (2026). Quantum-informed machine learning for predicting spatiotemporal chaos with practical quantum advantage. Science Advances, 12(16), eaec5049.

Weinberg, S. (1995). The Quantum Theory of Fields, Volume I: Foundations. Cambridge University Press.

Wittgenstein, L. (1953). Philosophical Investigations. Oxford: Blackwell.

Yamins, D. L. K., & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. Nature Neuroscience, 19, 356–365.