ChatGPT

Edge Research Directions for a Continuous Math-First Decision System

What the two Essentia essays actually argue

The two essays you linked are published by Essentia Foundation [1] and are written as philosophical–theoretical prompts rather than as conventional research surveys. They are best read as proposing *design metaphors* (boundary, selection, homeostasis, structural realism/idealism) that can be translated into formal objectives and architectures—especially if your goal is an always-on system that turns “math” into actionable representations and decisions. [2]

In “The cell membrane as the ‘missing link’ for the evolution of consciousness,” John Torday [3] argues that (i) quantum mechanics is salient to consciousness, but (ii) the crucial “missing link” is the *cell*, specifically the *cell membrane*, which creates a separation between organism and environment and thereby enables selective “assimilation or mirroring” of cosmic/quantum properties into differentiated organismic consciousness. [4]

A second, equally central theme in that essay is that evolution is fundamentally driven by *serial homeostatic control of energy*; symbiogenesis (in the broad sense of incorporations/mergers over evolutionary history) is described as serving that homeostatic energy control. This is used to motivate a picture in which organism–environment interactions, mediated by membranes, constitute a kind of boundary-conditioned “logic” linking physiology, cosmic structure, and (in the author’s own framing) consciousness. [5]

Torday also explicitly connects “assimilation” to *mathematics*: he claims that in assimilating environmental factors symbiogenically, organisms have “passively acquired the mathematics inherent in the Cosmos,” with references to ideas like the “Golden Ratio,” “fine-tuning,” and Fibonacci-like sequences. Whether or not you accept the metaphysical leap, the engineering-relevant content is: **boundary + selective exchange + energy/homeostasis + pattern regularities spanning scales.** [6]

The second essay, “Why mathematics works: The mind-reality connection,” claims that mathematics often develops *without direct empirical motivation* and later fits physical reality with surprising precision; it proposes that the most plausible explanation is that nature is itself the expression of “mind-like structures” that are also present in human intellect. This is positioned as an “analytic idealist” resolution meant to avoid a “causal gap” between abstract mathematical objects (as in Platonism) and physical reality. [7]

That essay is explicitly reacting to the long-running “unreasonable effectiveness” puzzle associated with Eugene Wigner [8] and to standard philosophy-of-math positions; it leans toward a metaphysics where “structure” is primary and minds participate in it. Regardless of metaphysics, one *computationally actionable* translation is: **treat mathematics not as a post-hoc fit, but as an internal structural prior and a compression target (discover the simplest, most reusable structure that explains diverse phenomena).** [9]

Boundary formalisms that map unusually well to your energy-driven architecture

A striking point of contact between Torday's "cell membrane" emphasis and modern theoretical neuroscience/AI is the idea that a *living boundary* can be formalized as a probabilistic conditional-independence shield—often discussed via "Markov blankets"—that separates internal states from external states while permitting structured exchange through sensory/active channels. This is one mainstream way to make "membrane-like" separation mathematically operational, without requiring speculative quantum-to-consciousness commitments. [10]

The best-known umbrella theory in this neighborhood is Karl Friston [11] 's Free Energy Principle (FEP), which frames perception/action/learning as variational free-energy minimization under generative models. Importantly for your project: *free energy* here is a rigorous objective functional, and the "boundary" theme shows up repeatedly (organisms maintain their organization by remaining within a limited region of state space, which the theory connects to probabilistic boundaries). [12]

Active inference extends this into decision-making by minimizing *expected* free energy (EFE), which decomposes into terms often interpreted as goal-seeking (extrinsic value) plus information-seeking (intrinsic epistemic value). This decomposition is one of the most direct "math-first" bridges between: **energy functional + constraints/preferences + continual uncertainty reduction**—which matches your repo's intention to separate experimental/uncertain patterns from deterministic/foundational ones. [13]

Two "edge research" directions arise immediately from this mapping:

First, treat your Level boundaries (L0–L3) and cross-level coupling as *explicit conditional-independence interfaces* rather than only as circuit analogies. This matters because it gives you a principled way to decide what information is allowed to cross levels and what must be summarized, which is exactly what Markov-blanket style formalisms are for. [14]

Second, reinterpret your Lyapunov-style stability energy as part of a broader "control-as-inference" family. In ML, maximum-entropy control and related formulations show formal equivalences between optimal control/RL and probabilistic inference, which provides another mathematically grounded bridge between "making decisions" and "minimizing an energy." [15]

Equation-first knowledge representations that go beyond symbolic equivalence

Your current pipeline already does something rare and valuable: it is **equation-centered**, not just text-centered. The frontier work that most cleanly improves this is the growing "math information retrieval / math embedding" literature: learning vector representations for equations and formulas that preserve both (a) symbolic structure and (b) semantic context. [16]

One influential approach ("equation embeddings") learns equation representations from surrounding words, explicitly targeting the fact that equations are nearly unique strings and cannot be treated like ordinary repeated tokens. These embeddings can capture semantic similarity among equations even when their surface forms differ—exactly what you need to reduce false negatives when you're trying to find cross-domain recurrence. [17]

In parallel, formula-embedding approaches for retrieval (e.g., Tangent-CFT) represent formulas through tree-structured encodings (appearance/content trees) plus embedding machinery, improving similarity search in large corpora. The practical punchline for your system is: **don't rely solely on algebraic canonicalization and SymPy equivalence; add a learned similarity layer that can propose "likely-equivalent or likely-analogous" candidates for deeper symbolic checking.** [18]

A closely related "edge" idea—highly compatible with your HDV mapping—is to treat your function/equation library as a Vector Symbolic / hyperdimensional code where binding/bundling operations encode compositional structure. Hyperdimensional computing argues that high-dimensional random vectors support robust compositional representations and similarity-based retrieval, often with strong noise tolerance. This gives a mathematically motivated alternative to "dimension assignment" heuristics: your HDV space can become a *structured algebra* rather than an index. [19]

Finally, the best "always-on parsing" upgrade is to move upstream: instead of scraping arbitrary HTML/PDF layouts, preferentially ingest sources that already expose MathML/structured markup. arXiv [20] now offers HTML for some papers via LaTeXML [21] conversions, and the ar5iv [22] project provides large-scale HTML5+MathML conversions. Compared with raw PDF scraping, this can substantially reduce parsing brittleness and increase the fidelity of equation extraction. [23]

Where PDFs remain unavoidable, the most widely used open-source pipeline for structured scientific PDF extraction is GROBID [24], which targets TEI/XML outputs for scientific publications. For high-accuracy equation OCR from images/PDF snippets, commercial tools like Mathpix [25] exist, though they change your privacy/cost profile. [26]

Turning your function basis into a discovery-and-falsification engine

Right now, your system discovers candidate universals by building a function basis library and looking for cross-domain recurrence. The research frontier that most directly strengthens "meaningful results" is: **make every promoted pattern survive at least one automated falsification loop**, not just a frequency threshold. [27]

Two families of methods are especially compatible with your goals:

Symbolic regression and equation discovery methods search for *compact* interpretable equations explaining data. AI Feynman is a well-known example that combines neural fitting with "physics-inspired" heuristics to recover ground-truth analytic forms on benchmark sets. [28]

In a more "engineering" mode, PySR [29] targets practical symbolic regression with configurable operator sets and a search backend optimized for performance. For a CPU-only always-on system, this matters because you can schedule symbolic-regression jobs as background validators for your candidate universal functions, producing human-reviewable hypotheses rather than only internal embeddings. [30]

Sparse-regression approaches like SINDy focus on discovering governing equations for dynamical systems under a sparsity assumption ("only a few terms matter"). If your library already enumerates candidate basis functions, SINDy-like selection becomes a principled way to score which subsets are actually explanatory in a dynamical setting, and it naturally aligns with MDL-style parsimony. [31]

To make this *continuous*, you also need a policy for choosing what to measure/ingest next. Experiment design and identifiability research emphasizes that parameters can remain practically unidentifiable even in large models, and that "optimal design" based on Fisher information can fail when models are sloppy or misspecified—meaning your promotion criteria should include misspecification checks, not only shrinking error bars. [32]

A powerful "edge" upgrade is to attach modern SciML validators to the patterns you discover. Physics-informed neural networks (PINNs) constrain neural models with known differential equations and can be used for both forward solution and inverse identification; universal differential equations (UDEs) embed learnable components into differential equation models to discover missing physics while retaining mechanistic scaffolding. [33]

Operator learning (neural operators, including Fourier neural operators) generalizes function approximation to *mappings between function spaces*, often yielding resolution-invariant surrogates for PDE families. For your "universal patterns across domains" hypothesis, this is relevant because many universals are better expressed as operators (transforms, integral operators, flow maps) rather than as isolated scalar functions. [34]

A concrete way to implement your dual-geometry idea using current research

Your Layer 9 concept—splitting learning into a statistical "Fisher information manifold" and a deterministic "isometric regularization manifold"—has strong precedents in contemporary optimization and representation learning, and there are now practical approximations that work without GPUs.

On the Fisher-information side, natural gradient methods treat parameter space as a Riemannian manifold with the Fisher information matrix as the metric; updates are scaled by curvature, improving conditioning relative to Euclidean gradients. Modern treatments emphasize when and why natural gradient works and how it relates to invariances of the parameterization. [35]

However, explicitly inverting the Fisher matrix is usually intractable at scale, so approximate methods are the practical path. K-FAC approximates curvature (often described via the Fisher) in a way that is efficiently invertible per layer via Kronecker factorization, making it one of the most operationally successful natural-gradient approximations. [36]

This gives you a direct, research-grounded implementation strategy for your "experimental manifold": maintain (approximate) Fisher blocks for the parameters that instantiate a candidate pattern, and use them for (a) parameter-importance ranking, (b) uncertainty surrogates, and (c) scheduling: query/ingest data that most increases information where it matters. [37]

On the “isometric regularization” side, the deep learning literature has converged on a related notion: training is dramatically easier when the network stays close to *dynamical isometry* (singular values of the input-output Jacobian concentrated near 1), because gradients neither explode nor vanish. Work on dynamical isometry gives both theory and practices (notably orthogonal constraints/initializations) that concretely instantiate your “distance preservation” goal. [38]

If your IRMF is intended to preserve distances and support robust transfer, then Lipschitz/orthogonality control mechanisms become exact “edge research” upgrades:

Parseval networks explicitly constrain layer Lipschitz constants (roughly, sensitivity to perturbations) by keeping weights close to Parseval tight frames, which are generalizations of orthogonal matrices. This offers a concrete regularizer for “near-isometry” in deep compositions. [39]

Spectral normalization is a computationally light way to control the operator norm of weight matrices, often used to stabilize training (classically in GAN discriminators, but the tool is general). In your setting it’s a knob for bounding distortion in learned transforms—useful if you want foundational patterns to remain stable as they compose. [40]

If “foundational patterns” must also satisfy conservation laws, then geometric/energy-conserving network classes supply a stricter test than generic smoothness: Hamiltonian neural networks and Lagrangian neural networks explicitly encode conservation behavior in the model class, and symplectic architectures target long-horizon stability of learned dynamics. These models give you a rigorous operationalization of your planned “conservation law test” for promotion to IRMF. [41]

Decision-making under constraints: where the cutting edge actually is

Your stated end-goal is not only to discover patterns, but to train a deep network that makes decisions “from a mathematical basis” while meeting constraints and optimizing an energy function. Three adjacent research lines are especially relevant and can be combined into a single coherent decision layer.

Constrained reinforcement learning offers explicit algorithms for learning policies under cost constraints (CMDP-style). Constrained Policy Optimization (CPO) is a canonical result here because it targets near-constraint satisfaction during learning via a trust-region style update and Lagrangian structure. [42]

Safe RL surveys emphasize that many practical approaches reduce to either (i) constrained optimization in policy space, or (ii) filtering/projection of actions to satisfy constraints, sometimes step-wise. This mirrors your desire for “meeting certain constraints” continuously, not just at convergence. [43]

Differentiable optimization layers let you *embed constraints as optimization problems inside networks*, and then backpropagate through the solution map. With CVXPY [44] plus cvxpylayers [45], you can express a disciplined convex program representing your constraints (or a relaxation of them), solve it in the forward pass, and differentiate through it. This is one of the most direct ways to guarantee constraint satisfaction by construction when your constraints are convex (or can be convexified). [46]

Energy-based learning offers an even more fundamental framing: inference and decision-making are formulated as minimizing an energy over configurations (with observed variables clamped), and learning shapes the energy landscape so that "good" configurations have low energy. This aligns strongly with your architecture's "energy function" language and with the active-inference/control-as-inference bridges. Yann LeCun [47]'s tutorial on energy-based models is still one of the clearest references for how inference-as-energy-minimization generalizes probabilistic models and supports structured outputs. [48]

Finally, maximum-entropy RL and "RL/control as probabilistic inference" make the equivalence explicit: optimizing policies can be written as inference under certain model classes, and algorithms like Soft Actor-Critic instantiate this with an objective that combines reward maximization and entropy maximization. This becomes a practical template for "decision from math" where your "energy" corresponds to (negative) log posterior or free-energy-like quantities, and uncertainty/entropy are first-class. [49]

The most "pioneer" synthesis, in the spirit of your repo and the Essentia prompts, is to treat your always-on system as a *closed loop*:

the ingestion/equation library proposes candidate structures; the dual-geometry layer classifies and stabilizes them; symbolic/SciML validators attempt to falsify or compress them; and the decision layer chooses the next actions (which papers to ingest, which domains to probe, which candidates to refine) by minimizing an expected-free-energy / control-as-inference objective while keeping constraints satisfied via embedded optimization layers or safe-RL style projections. [50]

---

[1] [39] Parseval Networks: Improving Robustness to Adversarial Examples
https://arxiv.org/abs/1704.08847?utm_source=chatgpt.com

[2] [4] [5] [6] [20] [24] The cell membrane as the 'missing link' for the evolution of consciousness | Essentia Foundation
https://www.essentiafoundation.org/the-cell-membrane-as-the-missing-link-for-the-evolution-of-consciousness/reading/

[3] [34] Neural Operator: Learning Maps Between Function Spaces ...
https://jmlr.org/papers/v24/21-1524.html?utm_source=chatgpt.com

[7] Why mathematics works: The mind-reality connection | Essentia Foundation
https://www.essentiafoundation.org/why-mathematics-works-the-mind-reality-connection/reading/

[8] [10] [14] The Markov Blanket Trick: On the Scope of the Free Energy ...
https://philsci-archive.pitt.edu/18843/1/The%20Markov%20Blanket%20Trick.pdf?utm_source=chatgpt.com

[9] [25] [29] THE UNREASONABLE EFFECTIVENSS OF ...
https://webhomes.maths.ed.ac.uk/~v1ranick/papers/wigner.pdf?utm_source=chatgpt.com

[11] [30] Interpretable Machine Learning for Science with PySR and ...
https://arxiv.org/abs/2305.01582?utm_source=chatgpt.com

[12] The free-energy principle: a unified brain theory?
https://pubmed.ncbi.nlm.nih.gov/20068583/?utm_source=chatgpt.com

[13] Whence the Expected Free Energy? | Neural Computation
https://direct.mit.edu/neco/article/33/2/447/95645/Whence-the-Expected-Free-Energy?utm_source=chatgpt.com

15 22 47 49 Reinforcement Learning and Control as Probabilistic ...
https://arxiv.org/abs/1805.00909?utm_source=chatgpt.com

16 17 44 Equation Embeddings
https://arxiv.org/abs/1803.09123?utm_source=chatgpt.com

18 Tangent-CFT: An Embedding Model for Mathematical Formulas
https://clgiles.ist.psu.edu/pubs/IDTIR2019.pdf?utm_source=chatgpt.com

19 Hyperdimensional Computing:
https://redwood.berkeley.edu/wp-content/uploads/2018/01/kanerva2009hyperdimensional.pdf?utm_source=chatgpt.com

21 28 AI Feynman: A physics-inspired method for symbolic ...
https://www.science.org/doi/10.1126/sciadv.aay2631?utm_source=chatgpt.com

23 Accessibility update: arXiv now offers papers in HTML format
https://blog.arxiv.org/2023/12/21/accessibility-update-arxiv-now-offers-papers-in-html-format/?utm_source=chatgpt.com

26 GROBID Documentation - Introduction
https://grobid.readthedocs.io/en/latest/Introduction/?utm_source=chatgpt.com

27 31 Discovering governing equations from data by sparse ...
https://www.pnas.org/doi/10.1073/pnas.1517384113?utm_source=chatgpt.com

32 The Limitations of Model-Based Experimental Design and ...
https://pmc.ncbi.nlm.nih.gov/articles/PMC5140062/?utm_source=chatgpt.com

33 Physics-informed neural networks: A deep learning ...
https://www.sciencedirect.com/science/article/pii/S0021999118307125?utm_source=chatgpt.com

35 New Insights and Perspectives on the Natural Gradient ...
https://jmlr.org/papers/volume21/17-678/17-678.pdf?utm_source=chatgpt.com

36 37 Optimizing Neural Networks with Kronecker-factored Approximate Curvature
https://arxiv.org/abs/1503.05671?utm_source=chatgpt.com

38 [1711.04735] Resurrecting the sigmoid in deep learning ...
https://arxiv.org/abs/1711.04735?utm_source=chatgpt.com

40 Spectral Normalization for Generative Adversarial Networks
https://arxiv.org/abs/1802.05957?utm_source=chatgpt.com

41 45 [1906.01563] Hamiltonian Neural Networks
https://arxiv.org/abs/1906.01563?utm_source=chatgpt.com

42 [1705.10528] Constrained Policy Optimization
https://arxiv.org/abs/1705.10528?utm_source=chatgpt.com

43 State-wise Safe Reinforcement Learning: A Survey
https://www.ijcai.org/proceedings/2023/0763.pdf?utm_source=chatgpt.com

46 Differentiable Convex Optimization Layers
https://arxiv.org/abs/1910.12430?utm_source=chatgpt.com

48 A Tutorial on Energy-Based Learning
https://yann.lecun.com/exdb/publis/pdf/lecun-06.pdf?utm_source=chatgpt.com

50 Generalised free energy and active inference - PMC - NIH
https://pmc.ncbi.nlm.nih.gov/articles/PMC6848054/?utm_source=chatgpt.com