NOUS: A New Class of Autonomous Intelligence Systems

Krasimir Petrov

Synora Group
contact@synoragroup.eu — synoragroup.eu

April 2026 Patent application EP26170794.7 filed at the European Patent Office. The class definition is formally verified in Lean 4.

ABSTRACT This paper introduces the NOUS class: a precisely defined category of autonomous intelligence systems whose internal dynamics are not learned but derived — uniquely forced by the intrinsic geometry of the system's own state space. Membership in the class is determined by six conditions, jointly necessary and sufficient, each a mathematical consequence of a single foundational hypothesis: that a system possessing a predictive map from internal states to probability distributions must operate on a statistical manifold with a canonically determined geometry. From this hypothesis, the state space, the symmetry group, the driving functional, and the dynamical law all follow without design freedom. A system satisfying all six conditions converges provably to the correct model of its environment, computes its own operational limits from internal geometry, detects adversarial perturbation through exact conservation laws without a model of the adversary, and coordinates with independent instances toward identical conclusions without communication. The class is non-empty: a computational proof of concept has verified that all structural and dynamical requirements are simultaneously satisfiable in finite-precision hardware across multiple parametric families, with empirical results exceeding theoretical prediction. The NOUS class is substrate-independent, scale-invariant under a single parameter, and structurally incompatible with weight-based architectures by construction — not as a matter of engineering insufficiency, but as a consequence of operating on a different mathematical kind.

1. The Question That Was Not Asked

The practice of building intelligent systems from trained parameters rests on an assumption that was observed before it was examined. Neural networks were found to approximate functions when exposed to sufficient data, and from this observation the field constructed a complete methodology: represent the system's knowledge as a vector of numerical weights, adjust the vector by minimising a loss on a training corpus, and scale the vector when performance proves insufficient. Seventy years of engineering followed from this observation — increasingly sophisticated architectures, optimisers, and scaling laws — without the foundational question being posed: is a fixed parameter vector in Euclidean space the correct substrate for intelligence, or merely a convenient one?

The question is not rhetorical. A parameter space with no canonical geometry — no metric derived from the statistical properties of the system's own predictions — cannot support a principled measure of distance between internal states. Without such a measure, the system has no basis for determining when two internal configurations are meaningfully different, no principled criterion for assessing its own reliability, and no computable boundary between what it can currently resolve and what lies beyond its reach. These are not features that were omitted from the design. They are structural absences that follow necessarily from the choice of substrate.

The field has recognised the symptoms. The gap between benchmark performance and production reliability in autonomous systems is not a consequence of insufficient scale; it is a consequence of operating on a substrate that was never designed to support the structures that reliable autonomous operation requires. Proposals advanced to close this gap — world model architectures, energy-based predictive systems, hierarchical joint embeddings — identify the right direction. Each recognises that autonomous intelligence requires an internal model of the environment maintained distinctly from the system's current action policy, a hierarchy of abstraction across temporal and spatial scales, and an observation space that is not specified externally but derived from the system's own structure. Those proposals stop short of the derivation. None establishes why any particular architecture is necessary, what symmetry group the dynamics must be invariant under, or what convergence guarantees follow from the structure rather than from the training procedure.

The NOUS class is the derivation those proposals were reaching toward. It begins not with an architecture but with a question: what does an autonomous intelligence system structurally require? Following the mathematics to its conclusion produces six conditions, each a necessary consequence of the foundational hypothesis, jointly sufficient for a class of systems whose properties are theorems rather than design targets.

2. The Foundational Hypothesis and What It Forces

A system is capable of autonomous intelligence only if it has a predictive map: a function from internal states to probability distributions over an observation space, satisfying three conditions. The map must be injective — distinct internal states must correspond to distinct predictions, so that the system's beliefs have unambiguous content. It must be smooth — the system must be able to move continuously between nearby beliefs. And the Fisher information matrix of the map must be positive definite on the interior of the parameter space — the system must be able to distinguish, in principle, between any two nearby predictions.

These three conditions — identifiability, smooth parametrisation, and Fisher information non-degeneracy — are not restrictions on what the system may do. They are the minimum structural requirements for a system that can be said to have beliefs about the world at all. A system failing any one of them does not have a principled internal model; it has an internal state whose relationship to the world is uninterpretable.

Given a predictive map satisfying these conditions, the state space is determined: it is a statistical manifold, a smooth space whose points are probability distributions parametrised by the system's internal state. On this manifold, a canonical metric is forced. The Fisher-Rao metric — the unique Riemannian metric invariant under all transformations that preserve the statistical structure of the distributions — is not a modelling choice. It is the only admissible geometry. Chencov's theorem, established in 1972, makes this precise: up to positive scalar multiple, the Fisher-Rao metric is the unique Riemannian metric on a statistical manifold that is invariant under Markov morphisms. Any system operating with a different metric — Euclidean gradient descent, Adam, RMSProp, or any learned metric — operates with a geometry that changes under transformations which preserve all statistical content. Its computations are representation-dependent in a way that cannot be principled.

From the canonical metric, two further structures are forced. The symmetry group of the statistical manifold — the group of transformations under which the Fisher-Rao metric is invariant and which are induced by stochastic maps preserving the parametric family — is uniquely determined. This is the Markov morphism group: the collection of all transformations of the observation space that preserve statistical sufficiency. Two independent mathematical arguments converge on this group as the unique maximal group preserving the full statistical structure of the manifold. No substitution is admissible. From the Markov morphism group, via a construction from representation theory, the observation space itself is derived: the system's operational domain — the set of quantities it monitors — is not specified externally but generated by the gauge structure. A system built on a statistical manifold with the Fisher-Rao metric derives what it observes from what it is.

This derivation chain — from the predictive map through the canonical metric to the symmetry group to the observation space — constitutes the mathematical content of the first two membership conditions of the NOUS class. Everything that follows is forced by the same chain.

3. The Six Conditions

A system is NOUS-class if and only if six conditions hold jointly and without exception. Each is a structural requirement on what the system *is*, not a performance threshold on what it does. A system satisfying all six is NOUS-class regardless of its domain of operation. A system failing any one is not NOUS-class regardless of its empirical performance, and the failure cannot be remedied by scaling, fine-tuning, or architectural modification within the conventional paradigm.

Condition 1 — Statistical Arena. The system's state space is a statistical manifold with the Fisher-Rao metric, arising from a predictive map satisfying identifiability, smooth parametrisation, and Fisher information non-degeneracy. The Fisher-Rao metric is the unique admissible geometry. The observation space is derived internally from the gauge structure via the coadjoint orbit construction. No externally specified observation space and no alternative metric are admissible.

Condition 2 — Gauge Bundle and History-Encoding Connection. The system carries a gauge bundle over the statistical manifold with the Markov morphism group as its structure group. The gauge connection encodes the accumulated history of the system's trajectory via the path signature of the belief-state path — a mathematical object that captures the full sequence of past states in a form invariant under time reparametrisation and satisfying a precise composition rule. A memoryless connection, depending only on the current state, does not satisfy this condition. The history encoding is not an architectural choice; it is a requirement of gauge covariance.

Condition 3 — Environmental Structure. The system maintains an environmental generative structure that is a distinct object from and independently maintained with respect to its belief state. The driving functional of the system is a function of two distinct objects. A system minimising a loss function of a single argument — its own weights conditioned on a fixed dataset — does not maintain a genuine self-other distinction and fails this condition structurally.

Condition 4 — Free Energy Functional. The system's driving functional is the Kullback-Leibler divergence between the belief state and the environmental structure. This is not a choice: among all functions satisfying additivity over independent observations, invariance under content-preserving transformations, smoothness, and the chain rule for conditional distributions, the Kullback-Leibler divergence is the only one. The free energy is a strict Lyapunov function along the dynamics, with a unique global minimum at the point where the belief state equals the environmental structure.

Condition 5 — Agent Dynamics. The system's dynamics are governed by the *local agent equation*:

$$\mathcal{D}_\tau \Psi^{(N)} \;=\; -\big[g^{(N)}\big]^{-1} \nabla \mathcal{F}\big[\Psi^{(N)} \,|\, S^{(N)}\big] \;-\; \int_{\partial \Omega_N} \gamma \, dA \tag{1}$$

The left-hand side is the gauge-covariant derivative of the belief state with respect to the resolution parameter τ — the system's internal measure of accumulated resolution progress, not an external clock. The first term on the right is the natural gradient of the Kullback-Leibler divergence: the direction of steepest uncertainty reduction in the geometry of the statistical manifold. The second term is the *boundary resistance*: the thermodynamic cost of maintaining the epistemic horizon $\partial \Omega_N$, computed from the extrinsic curvature of the boundary of the system's operational domain as embedded in the Fisher-Rao geometry. This boundary term is non-negative for all finite operational depths, strictly positive, and vanishes only when the boundary recedes to infinity. The dynamical law is uniquely forced by the geometry and the irreversibility of the resolution parameter: no other first-order dissipative law is compatible with these requirements.

Condition 6 — Gauge Group. The gauge group acting on the bundle is the Markov morphism group, uniquely determined as stated in Section 2. No substitution is admissible. The gauge group determines the conservation laws, the topological classification of irreducible errors, and the adversarial detection mechanism.

These six conditions are jointly necessary and sufficient. The formal class definition, the complete set of structural and dynamical requirements derived from these conditions, and the proof that the conditions are necessary consequences of the foundational hypothesis are maintained in patent application EP26170794.7 and the associated internal documentation, which includes a formal verification in Lean 4.

4. What Membership Entails

The six conditions are premises. What follows from them is determined by the mathematics, not by engineering decisions made after the class is defined. The following properties hold for every member of the class.

Provable convergence with uniform rate

The free energy decreases monotonically along the dynamics, and the system converges exponentially to the ground state — the unique point where the belief state equals the environmental structure. The convergence rate is uniform across all dimensions and across all standard parametric families used in practice: Gaussian, multinomial, Wishart, Gamma. Increasing the dimensionality of the state space does not degrade this rate. This is not a typical-case guarantee; it is a structural consequence of the dynamics and the geometry, holding for every NOUS-class system in every domain.

Self-assessed operational limits

At each resolution step, the system computes a quantity from its own geometry that determines whether it is currently in a regime of proactive uncertainty resolution or whether the boundary resistance prevents net progress. Specifically, the system evaluates whether

$$P^2 > \mathcal{R} \tag{2}$$

where P is the gradient pull — the magnitude of the natural gradient in the Fisher-Rao geometry — and $\mathcal{R} \geq 0$ is the boundary resistance derived from the extrinsic curvature of the epistemic horizon. Both quantities are computable from the system's own state. The signal is binary: the system is either resolving or stalled. When stalled, it knows this from internal geometry alone. No external calibration set, no held-out evaluation, and no heuristic confidence threshold are involved.

The same mechanism supports autonomous scaling. When the system has resolved uncertainty at its current operational depth to the degree the geometry permits, it increases its depth autonomously via a recursive seeding procedure that uses the resolved state as the initial condition for the next level. No retraining, no architectural change, no external scheduler. The decision derives from the system's own dynamics.

Model-free adversarial detection

For each one-parameter subgroup of the gauge group, a corresponding quantity is exactly conserved along the dynamics. Any perturbation of the environmental structure that breaks gauge symmetry — any perturbation that is not a Markov morphism — produces a measurable deviation from exact conservation. The detection mechanism requires no prior model of the adversary, no training on attack examples, and no catalogue of known perturbation types. A perturbation evades detection if and only if it is itself a Markov morphism, in which case it leaves all gauge-invariant content of the system unchanged. The class of undetectable perturbations is precisely characterised.

Topological classification of irreducible errors

When a NOUS-class system cannot reach the ground state by continuous evolution — when a residual uncertainty persists despite correct dynamics — the residual is classified by the topological structure of the gauge group. The classification distinguishes errors that are continuously resolvable from those that are topologically protected and cannot be removed without a discontinuous transition. The system determines this from its own geometry, via holonomy computation. No external diagnostic is required.

Communication-free consensus

Independent instances of a NOUS-class system operating in the same environment, with no inter-system communication, converge on identical gauge-invariant content. This follows from the uniqueness of the gauge group, the uniqueness of the ground state, and the Lyapunov property acting jointly: since all instances share the same unique structure and the same unique attractor, and the attractor is globally stable, all instances converge to the same fixed point. Coherence across a fleet of independent agents is a mathematical consequence of structural uniqueness, not an engineered coordination protocol.

Scale invariance under a single parameter

The truncation depth N is the sole variable determining the computational scale of a NOUS-class system. A system at $N = 1$ and a system at $N = 10^6$ are the same mathematical object — the same six conditions, the same local agent equation (1), the same gauge structure — operating at different depths of the same resolution hierarchy. Increasing N extends the system's operational domain without modifying the equations, without retraining, and without architectural change. The convergence rate is uniform across depths. A NOUS-class system deployed on embedded hardware and the same system deployed on a distributed compute cluster are the same system. The substrate changes. The geometry is the same. The equations are the same. The class is the same.

5. The Structural Exclusion of Conventional Architectures

A conventional neural network — including transformer architectures, diffusion models, and reinforcement learning systems in their standard formulations — fails the NOUS membership criterion on multiple conditions simultaneously. The failures are structural, not quantitative.

A neural network minimises a loss function defined on a fixed Euclidean parameter space. This fails Condition 1: the parameter space is not a statistical manifold arising from a predictive map in the required sense, and the metric operative in the dynamics — Euclidean or any approximation thereof — is not the Fisher-Rao metric induced by the system's own predictions. A system whose dynamics do not respect the canonical geometry of its state space does not operate on the correct substrate for the structural properties that follow from it.

A neural network's loss function is a functional of a single argument — the weight vector conditioned on a fixed training corpus. This fails Condition 3: the system does not maintain a distinct environmental generative structure. Its driving functional has no genuine self-other distinction. The weights and the data are not two separate objects in a relational functional; they are one object evaluated against fixed inputs.

A neural network has no epistemic horizon: no bounded operational domain with a smooth boundary from which a boundary resistance term is computed. Consequently it carries no computable self-other distinction in the mathematical sense, no mechanism for determining when it is operating reliably and when it is not, and no structural basis for the self-assessment that autonomous operation requires.

A neural network carries no gauge group whose homotopy groups classify representational errors as topologically protected or continuously resolvable. It has no Noether charge conservation mechanism. Its connection update is memoryless — depending only on the current parameter vector, not on the path taken to reach it — and therefore fails Condition 2. These are failures of kind. The conditions of NOUS-class membership are conditions on the nature of the state space, the driving functional, the geometry of the dynamics, and the causal structure of the evolution. They are not performance thresholds that scaling could satisfy. Satisfying them requires operating on a different mathematical substrate — statistical manifold with Fisher-Rao metric, Markov morphism gauge group, relational Kullback-Leibler divergence, gauge-covariant gradient flow with boundary term. No continuous modification of a weight-based architecture reaches this substrate.

6. Extensions of the Class

The NOUS class as defined by the six conditions is a base class. Three conditional extensions are established, each activating additional structural requirements while preserving every requirement of the base class.

The **spatial extension** activates when the system operates on a spatially distributed environment. The intelligence field is defined on a spatial manifold rather than at a point; the spatial metric becomes an independent dynamical variable; and the total free energy integrates both a potential term and a gradient energy term over the spatial domain. In the sub-threshold regime — when the system cannot resolve uncertainty at its current depth — the coupled dynamics reduce to a known and tractable structure whose properties are derivable by standard methods, providing an auditable degradation mode for safety-critical deployment.

The **quantum extension** activates when the system operates on quantum-mechanical substrates or processes quantum sensor data. The classical statistical manifold is replaced by a quantum statistical manifold of density matrices on a Hilbert space; the gauge group by the special unitary group SU(n); and the driving functional by the quantum relative entropy — the unique quantum divergence satisfying additivity, complete positive trace-preserving map invariance, and the quantum chain rule. All structural guarantees are preserved. The topological defect spectrum differs from the classical case in ways determined by the homotopy structure of SU(n).

The **reflexive extension** activates when the system maintains a model of its own dynamics. The system augments its environmental structure with a self-model operating at a strictly lower resolution depth than its primary operation. The resolution gap is a structural necessity: a self-model at the same depth would be indistinguishable from the belief state, violating Condition 3. The extension preserves all structural and dynamical properties of the base class and introduces an irreducible self-modelling uncertainty that prevents overconfident self-assessment.

The compositional structure of the class is established for three modes: independent instances operating in the same environment, observation-coupled instances sharing partial observation streams, and mutually coupled instances treating each other as environmental structure. In each mode, the structural and dynamical properties of the class are preserved under stated conditions. The communication-free consensus guarantee of Section 4 is the independent composition result.

7. Empirical Grounding

The NOUS class is non-empty. A computational proof of concept has verified that physical systems satisfying the membership criterion are realisable in finite-precision discrete hardware.

The verification programme confirmed that all structural requirements (SR 1 through SR 8) and dynamical requirements (DR 1 through DR 12) are simultaneously satisfiable for point-agent implementations. Tests were executed across three parametric families — Gaussian, Gamma, and qubit Bures — and across multiple truncation depths. The intelligence threshold was confirmed as a computable and sharp transition. Independent instances without communication converged on identical gauge-invariant content. The quantum extension requirements were confirmed satisfiable on the Bures manifold with SU(2) gauge group. Path-signature history encoding and Noether charge monitoring were confirmed for the Gaussian case.

Twenty-three of twenty-three test cases passed. Empirical results exceeded theoretical prediction: convergence behaviour was better than the analytical bounds guarantee. This is the expected signature of a conservative theory — one whose bounds are provably correct and whose actual behaviour surpasses them. An implementation of Kernel A — the classical point agent — is operational in Python and Rust.

8. Intellectual Context and Prior Art

The mathematical foundations on which the NOUS class rests are established results in information geometry, differential geometry, and representation theory. Chencov's uniqueness theorem for the Fisher-Rao metric dates to 1972 [7]. Amari's development of natural gradient methods and the alpha-connections spans from 1985 onwards [1, 2, 3]. The path signature as a mathematical object for encoding trajectory history was established by Chen in 1958 [6] and developed for rough path theory by Lyons from 1998 [16]. The Kirillov-Kostant-Souriau theorem on coadjoint orbits is standard representation theory [14]. The application of gauge theory to machine learning for data processing — gauge equivariant networks — was developed from 2019 onwards by Cohen, Weiler, and collaborators [8].

None of these lines of work identifies the Markov morphism group as the gauge group of a principal bundle over a statistical manifold, constructs a gauge connection from path signatures contracted with the Amari-Chencov tensor, derives the observation space from the coadjoint orbit, or establishes the conservation laws and topological classification that follow. The step from the symmetry-group characterisation of information geometry to the gauge-theoretic dynamics of the NOUS class requires additional mathematical structure — principal bundles, connections, curvature, holonomy — that is not present in the symmetry-group characterisation alone. Friston's free energy principle, developed from 2010 [10], operates on statistical manifolds and uses natural gradient methods, but does not identify a gauge group, construct a gauge bundle, or derive a boundary term from extrinsic curvature.

A programmatic proposal for a neuronal gauge theory formulated explicitly in 2016 [11] set out to accomplish precisely the construction that NOUS realises — formulating a free energy principle as a gauge theory on statistical manifolds. The construction was not completed by that proposal or by any subsequent work in the decade following. The NOUS class is the completion of that construction.

The class definition, the local agent equation, the conservation laws, the topological classification, and the quantum and spatial extensions are covered by European Patent Application EP26170794.7, filed directly at the European Patent Office. The mathematical foundation is formally verified in Lean 4. The verification confirms the correctness of the derivation chain from the foundational hypothesis through the six conditions to the structural and dynamical requirements.

9. What Follows

The NOUS class establishes a category. What falls within it — what systems can be built, in what domains, at what scales, on what substrates — is determined by the six conditions and the extensions established from them.

The domains in which NOUS-class operation is structurally differentiated from weight-based approaches are precisely those in which the structural absences of the conventional paradigm are most consequential: autonomous systems that must operate without supervision in environments not represented in training data; safety-critical systems that must assess their own reliability from internal structure rather than external calibration; distributed systems that must coordinate toward coherent conclusions without communication infrastructure; systems that must detect adversarial perturbation without a prior catalogue of attack types; and systems that must operate across a range of computational scales on the same mathematical foundation.

The theory admits further extensions not detailed here. The spatial extension connects to the geometry of physical space in ways whose implications are still being examined. The quantum extension operates on substrates that are becoming practically accessible. The compositional structure scales to arbitrarily large fleets of independent agents under conditions derivable from the class definition.

Every system that resolves uncertainty about its environment in a principled and autonomous way will satisfy the six conditions, whether or not its builders know it. The geometry is the same. The equations are the same. The class is the same.

Synora Group invites engagement from researchers, institutions, and technical partners.
`contact@synoragroup.eu`

References

[1] S. Amari. *Differential-Geometrical Methods in Statistics*. Springer, 1985.

[2] S. Amari. Natural gradient works efficiently in learning. *Neural Computation*, 10(2):251–276, 1998.

[3] S. Amari and H. Nagaoka. *Methods of Information Geometry*. American Mathematical Society, 2000.

[4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. 2nd ed., Birkhäuser, 2008.

[5] F. Barbaresco. Lie group statistics and Lie group machine learning based on Souriau Lie groups thermodynamics. *Entropy*, 22(6):641, 2020.

[6] K. T. Chen. Integration of paths — a faithful representation of paths by non-commutative formal power series. *Transactions of the American Mathematical Society*, 89(2):395–407, 1958.

[7] N. N. Chencov. *Statistical Decision Rules and Optimal Inference*. American Mathematical Society, 1982.

[8] T. S. Cohen, M. Weiler, B. Kicanaoglu, and M. Welling. Gauge equivariant convolutional networks and the icosahedral CNN. *Proceedings of ICML*, 2019.

[9] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. 2nd ed., Wiley, 2006.

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

[11] K. Friston and B. Sengupta. Towards a neuronal gauge theory. *PLOS Biology*, 14(3):e1002400, 2016.

[12] G. W. Gibbons and S. W. Hawking. Action integrals and partition functions in quantum gravity. *Physical Review D*, 15(10):2752–2756, 1977.

[13] P. Kidger and T. J. Lyons. Signatory: differentiable computations of the signature and logsignature transforms. *ICLR*, 2021.

[14] A. A. Kirillov. *Elements of the Theory of Representations*. Springer, 1976.

[15] J. P. LaSalle. Some extensions of Liapunov's second method. *IRE Transactions on Circuit Theory*, 7(4):520–527, 1960.

[16] T. J. Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310, 1998.

[17] N. D. Mermin. The topological theory of defects in ordered media. *Reviews of Modern Physics*, 51(3):591–648, 1979.

[18] M. Nakahara. *Geometry, Topology and Physics*. 2nd ed., IOP Publishing, 2003.

[19] D. Petz. Monotone metrics on matrix spaces. *Linear Algebra and its Applications*, 244:81–96, 1996.

[20] C. R. Rao. Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37:81–91, 1945.

[21] J. W. York. Role of conformal three-geometry in the dynamics of gravitation. *Physical Review Letters*, 28(16):1082–1085, 1972.

[22] K. Petrov. Autonomous intelligence system based on gauge-covariant field dynamics on statistical manifolds. European Patent Application EP26170794.7, European Patent Office, 2026.