Neurogeometric Embodied AI: A Topological Framework for Structured Cognition

B. Matthew Gossage
Independent Researcher
b.matthew@gossage.org

Abstract

This paper proposes *Neurogeometric Embodied AI*, a novel framework that unifies embodied cognition, topological representation learning, and temporal abstraction to advance the structural foundations of artificial intelligence. Conventional deep learning systems operate on disembodied, static data streams, producing latent spaces that are statistically rich but semantically brittle, causally shallow, and geometrically unstructured. In contrast, biological intelligence emerges from continuous sensorimotor interaction with the physical world, where perception, action, and memory co-evolve within a constrained geometric substrate.

Our central hypothesis is that embodied agents – those with proprioceptive feedback, multimodal perception, and action-conditioned learning – induce latent manifolds with topological features that reflect spatial continuity, causal entanglement, and temporal coherence. This paper outlines a methodology to analyze these emergent structures using tools from topological data analysis (TDA), including persistent homology and curvature metrics, and propose experimental protocols to compare embodied and disembodied agents across matched tasks. Recognizing the computational demands of full geometric analysis, we propose two complementary paths: one that pursues rigorous topological metrics, and another that leverages representative subsets, proxy ensembles, and simulated environments to approximate structure. To ensure the reliability of both approaches, we include a robustness evaluation that tests the stability of latent features under noise, dropout, and adversarial perturbation.

This triadic synthesis – embodiment, topology, and time – offers a path toward AI systems that do not merely scale in size, but evolve in structure. By grounding representation in the geometry of experience, *Neurogeometric Embodied AI* aims to unlock new capabilities in generalization, interpretability, and causal reasoning, with implications for autonomous systems, cognitive modeling, and neurosymbolic integration.

1 Introduction

Modern deep learning systems have achieved remarkable performance across domains, yet they remain fundamentally disembodied. These models learn from static datasets, divorced from the physical constraints and causal dynamics that shape biological intelligence. As a result, their latent representations are often brittle, uninterpretable, and poorly aligned with the structure of the real world.

In contrast, biological cognition emerges from continuous interaction with a physical environment. Organisms perceive, act, and adapt through sensorimotor feedback loops, forming representations that are geometrically constrained, causally entangled, and temporally scaffolded. Neural manifolds observed in motor cortex and hippocampal segmentation during episodic memory suggest that cognition is shaped by the geometry of experience.

We hypothesize that artificial agents exposed to similar embodied conditions will develop latent spaces with topological structure – manifolds that reflect spatial continuity, causal entanglement, and temporal abstraction. These structures can be analyzed using tools from topological data analysis (TDA), including persistent homology and curvature metrics, and framed mathematically as dynamical systems evolving over Riemannian manifolds.

This paper introduces Neurogeometric Embodied AI, a framework that integrates:

- Embodied learning via sensorimotor interaction in physics-rich environments.
- Topological representation analysis to uncover latent structure.
- Temporal scaffolding through event-triggered memory segmentation.

We propose to validate this framework through controlled experiments that contrast embodied and disembodied agents, assessing the structure of their latent representations for alignment with causal dynamics and temporally segmented experience. Our goal is to move beyond scale-driven performance toward structure-driven understanding – building agents whose intelligence emerges from the geometry of their experience. Embodiment is not merely a physical constraint but a cognitive scaffold that shapes representational structure and learning dynamics (Barrett & Stout, 2024)

2 Related Work

The development of intelligent agents capable of learning from interaction has been a central pursuit in artificial intelligence. Recent advances in reinforcement learning, robotics, and simulation environments have enabled embodied agents to acquire complex behaviors through trial and error. However, while these systems demonstrate impressive performance, they often lack structured internal representations that reflect the geometry, causality, and temporal dynamics of their experience.

2.1 Embodied AI

Embodied learning has gained traction through platforms such as DeepMind's control suite, OpenAI Gym, and robotics simulators like MuJoCo and Brax. These environments allow agents to learn through sensorimotor interaction, mimicking aspects of biological cognition. Prior work has explored proprioception, tactile feedback, and multimodal fusion to enhance policy learning and task performance. However, the focus has largely remained on behavioral outcomes and reward optimization, with limited attention to the structure of the agent's internal representations. Specifically, there has been little investigation into how embodiment shapes the topology of latent spaces – the hidden geometries that encode perception, action, and memory.

2.2 Topological Data Analysis in Machine Learning

Topological Data Analysis (TDA) offers a powerful toolkit for understanding the shape of data. Techniques such as persistent homology, Betti numbers, and curvature metrics have been applied to static datasets, generative models, and feature spaces in supervised learning (Ballester et al., 2023). Some studies have explored the topology of latent spaces in autoencoders and GANs, revealing insights into disentanglement and robustness. Yet these applications are typically divorced from embodied learning contexts. The latent spaces analyzed are formed from passive data ingestion, not from active, sensorimotor engagement with an environment. As such, the topological features observed may reflect statistical regularities rather than grounded cognitive structure.

2.3 Temporal Abstraction and Event-Based Memory

Temporal abstraction is a critical component of intelligent behavior. Models such as predictive coding, surprise-based segmentation, and hierarchical reinforcement learning have introduced mechanisms for event detection and memory formation. These approaches enable agents to compress experience, form goals, and reason over extended time horizons. However, the relationship between temporal abstraction and latent space topology remains underexplored. Few studies have examined how event-triggered memory modules influence the geometric structure of internal representations, or how temporal dynamics manifest as topological motifs such as loops, branches, and arcs.

2.4 Missing Synthesis

To our knowledge, no existing framework unifies embodied sensorimotor learning, topological representation analysis, and temporal scaffolding into a coherent system. While each component has been studied in isolation, their integration remains an open frontier. We propose *Neurogeometric Embodied AI* as a novel synthesis – a framework in which agents learn through interaction, form structured latent geometries, and scaffold cognition across time. This approach reframes intelligence not as a function of scale or reward maximization, but as an emergent property of structured experience.

3 Methods

3.1 Agent Architecture

We propose an agent architecture designed to support embodied learning, multimodal perception, and temporally segmented memory. The agent operates within a physics-rich simulation environment and receives continuous sensorimotor input, including proprioception, visual data, and tactile feedback. Action outputs are conditioned on both current observations and latent memory states, enabling goal-directed behavior and adaptive planning. The core components include:

- Transformer-based encoder for multimodal input fusion.
- Latent memory module for event-triggered segmentation and retrieval.
- Policy head trained via reinforcement learning or imitation learning, depending on task structure.

This architecture is designed to induce structured latent representations that evolve over time in response to embodied interaction.

3.2 Learning Objectives

The agent is trained to perform tasks that require spatial navigation, object manipulation, and temporal reasoning. Learning objectives include:

- Prediction of future sensory states conditioned on action sequences.
- Compression of experience into latent codes that preserve causal and temporal structure.
- Reward optimization for task completion, with optional auxiliary losses for representation smoothness and topological consistency.

These objectives are intended to encourage the emergence of latent manifolds that reflect the geometry of experience.

3.3 Topological Representation Analysis

We propose analyzing the agent's latent space using tools from topological data analysis (TDA). Specifically:

- Persistent homology will be used to identify topological invariants such as loops, branches, and voids.
- Curvature metrics (e.g., Ricci curvature, sectional curvature) will assess the smoothness and connectivity of the latent manifold.
- Dimensionality reduction techniques (e.g., UMAP, diffusion maps) will aid in visualizing latent structure and identifying topological motifs.

These analyses will be performed across training epochs and task conditions to evaluate how embodiment influences representational topology. To balance analytical fidelity with computational feasibility, we propose two complementary paths: one that applies full geometric analysis, and another that uses representative subsets and proxy metrics to approximate structure.

3.4 Temporal Scaffolding

To model temporal abstraction, we propose an event-triggered memory segmentation mechanism. This module monitors prediction error, novelty, or surprise signals to identify boundaries between meaningful episodes. Each segment is encoded into a latent vector, forming a temporally ordered scaffold of experience. We hypothesize that this scaffold will exhibit topological features such as:

- Loops, corresponding to habitual or cyclic behavior.
- Branches, reflecting decision points or counterfactual trajectories.
- Arcs, representing goal-directed sequences.

This mechanism is inspired by biological episodic memory and is designed to support long-horizon reasoning and causal inference (Ma & Li, 2023).

We also propose to evaluate the stability of these topological features under perturbation, including noise injection, dropout simulation, and adversarial inputs, as part of a broader robustness analysis.

4 Mathematical Foundations of Neurogeometric Representation

The latent spaces formed by learning agents are often treated as abstract vector spaces, optimized for task performance but rarely interrogated for structural properties. In *Neurogeometric Embodied AI*, we propose that these latent spaces can be meaningfully modeled as Riemannian manifolds – smooth, differentiable spaces equipped with a metric that reflects the agent's internal representation of its environment.

4.1 Latent Manifolds and Metric Structure

We define the agent's latent space $\mathcal{L} \subset R^n$ as a Riemannian manifold $(\mathcal{L}, g)$, where g is a metric tensor induced by the agent's encoder. This metric captures the similarity between internal states and evolves as the agent learns. Embodied interaction is expected to induce curvature in $\mathcal{L}$, shaping the manifold to reflect spatial continuity, causal entanglement, and temporal coherence.

Curvature metrics such as Ricci curvature and sectional curvature will be used to assess the smoothness and connectivity of the manifold. Regions of high curvature may correspond to decision boundaries, bottlenecks, or transitions in behavior.

4.2 Persistent Homology and Topological Invariants

To analyze the global structure of $\mathcal{L}$, we propose the use of persistent homology, a tool from topological data analysis (TDA) that captures features such as loops, branches, and voids across multiple scales. Given a filtration $\{ \mathcal{L}_\epsilon \}_{\epsilon > 0}$ of the latent space, we compute Betti numbers β_k to quantify the number of k-dimensional holes.

These topological invariants are expected to reflect cognitive motifs:

- β_1: Loops → habitual behavior or cyclic routines.
- β_0: Connected components → distinct behavioral modes or episodic memories.
- β_2: Voids → representational gaps or uncertainty regions.

Embodied agents are anticipated to exhibit richer topological structure than disembodied counterparts, shaped by the geometric constraints of sensorimotor experience. We further propose to evaluate the stability of these topological features under perturbation, using robustness metrics described in Section 8.

4.3 Dynamical Systems Perspective

The evolution of latent representations over time can be framed as a dynamical system on $\mathcal{L}$. Let $f_t\colon \mathcal{L} \to \mathcal{L}$ be the transformation induced by learning at time t. We propose to study the trajectory \{ f_t(x) \} for a given latent state x, analyzing its stability, convergence, and bifurcation behavior. This perspective enables the identification of:

- Attractors: Stable behavioral patterns.
- Limit cycles: Repeating routines or habits.
- Bifurcations: Transitions in strategy or representation.

By modeling learning as flow on a manifold, we aim to uncover the geometric dynamics of cognition.

4.4 Temporal Segmentation and Topological Evolution

We propose that temporally segmented experience – induced by event-triggered memory modules – leads to piecewise topological evolution of $\mathcal{L}$. Each episode forms a submanifold $\mathcal{L}_i$, and the sequence $\{\mathcal{L}_i\}$ traces a path through representational space. This paper proposes that:

- Transitions between episodes correspond to topological shifts (e.g., loop closure, branch formation).
- The global topology of $\bigcup_i \mathcal{L}_i$ reflects the agent's cognitive scaffold.

This mechanism parallels biological episodic memory and may support long-horizon reasoning and counterfactual inference.

5 Experimental Design

To evaluate the hypotheses proposed in *Neurogeometric Embodied AI*, we outline a set of controlled experiments designed to compare the latent representations of embodied and disembodied agents across matched tasks. These experiments are intended to test whether sensorimotor experience induces topological structure in latent space, and whether such structure confers advantages in generalization, interpretability, and causal reasoning.

5.1 Experimental Setup

We propose to deploy agents in simulated environments with varying degrees of embodiment:

- Embodied Agents: Operate in physics-rich environments with proprioception, tactile feedback, and action-conditioned learning.
- Disembodied Agents: Receive static observations or replayed sequences without direct interaction or feedback.

Tasks will include:

- Spatial navigation in procedurally generated mazes.
- Object manipulation with variable affordances.
- Temporal prediction in event-driven sequences.

All agents will be trained using identical architectures and optimization protocols, with embodiment as the only variable.

5.2 Hypotheses

We aim to test the following hypotheses:

- H1: Embodied agents will form latent spaces with richer topological structure (e.g., higher Betti numbers, more persistent features).
- H2: These topologies will align with causal dynamics and temporally segmented experience.
- H3: Structured latent spaces will improve performance on downstream tasks requiring generalization, planning, or counterfactual reasoning.

5.3 Metrics

To evaluate these hypotheses, we propose the following metrics:

- Topological Metrics:
- Betti numbers β_k via persistent homology.

- Curvature statistics (Ricci, sectional) across latent trajectories.
- Topological persistence diagrams and barcode stability.
- Behavioral Metrics:
- Task success rate and sample efficiency.
- Generalization to novel environments.
- Robustness to perturbations and occlusions, including noise injection, dropout simulation, and adversarial inputs.
- Alignment Metrics:
- Causal alignment score: correlation between latent transitions and action-conditioned predictions.
- Temporal coherence score: consistency of latent structure across episodic boundaries.

5.4 Controls and Ablations

To isolate the effects of embodiment and topology, we propose the following controls:

- Architecture Control: Identical encoders and policy heads across conditions.
- Data Control: Matched observation streams for embodied and disembodied agents.
- Ablation Studies:
- Remove event-triggered memory segmentation.
- Disable curvature regularization in latent space.
- Vary sensory modalities (e.g., vision-only vs. multimodal).

These controls will help disentangle the contributions of embodiment, topology, and temporal scaffolding.

5.5 Expected Insights

If validated, these experiments may reveal that:

- Embodied interaction induces latent geometries that reflect the structure of experience.
- Topological analysis provides interpretable signatures of cognition.
- Structured representations support more flexible and robust behavior.

These insights would support the broader claim that intelligence emerges not from scale alone, but from the geometry of grounded experience. They would also inform the design of scalable analysis pipelines, balancing full geometric fidelity with approximate methods for structure estimation.

6 Results & Discussion

6.1 Anticipated Outcomes

If the proposed experiments are conducted, we anticipate that embodied agents will exhibit latent representations with richer topological structure than their disembodied counterparts. Specifically, we expect to observe:

- Higher Betti numbers in persistent homology analyses, indicating the presence of loops, branches, and voids in latent space.
- Localized curvature in latent manifolds, corresponding to decision boundaries, behavioral transitions, and episodic segmentation.
- Temporal coherence across event-triggered memory segments, forming structured scaffolds that reflect goal-directed behavior and habitual routines.

These outcomes would support the hypothesis that sensorimotor experience induces geometric and topological structure in internal representations.

6.2 Interpretability and Cognitive Signatures

We propose that topological motifs in latent space may serve as interpretable signatures of cognition. For example:

- Loops may correspond to habitual behavior or cyclic routines.
- Branches may reflect decision points or counterfactual trajectories.
- Arcs may represent goal-directed sequences or planning horizons.

Such motifs could provide a new lens for understanding agent behavior, enabling post hoc analysis of decision-making, memory formation, and causal inference.

6.3 Implications for Generalization and Robustness

We hypothesize that agents with structured latent geometries will generalize more effectively to novel environments and exhibit greater robustness to perturbations. To evaluate these claims, we propose targeted robustness assessments described in Section 8. The underlying mechanisms may include:

- Topological continuity, which preserves semantic relationships across contexts.
- Curvature-aware representations, which encode transitions and boundaries more faithfully.
- Temporal scaffolding, which enables long-horizon reasoning and episodic recall.

These properties could enhance performance in tasks requiring planning, adaptation, and counterfactual reasoning.

6.4 Limitations and Open Questions

As this framework is currently theoretical, several limitations and open questions remain:

- Scalability: **Can topological analysis be efficiently applied to high-dimensional latent spaces in large-scale agents?**
- Biological fidelity: **To what extent do the proposed motifs align with neural structures observed in biological cognition?**
- Architectural dependence: **How sensitive are the topological features to encoder design, training regime, or sensory modality?**

These questions will guide future empirical work and refinement of the framework.

7 Figures & Visual Framework

As part of the development and communication of *Neurogeometric Embodied AI*, we propose the creation of a set of conceptual diagrams that illustrate the architecture, latent space geometry, and temporal scaffolding of the framework. These figures are not based on empirical results, but are intended to visually clarify the theoretical constructs introduced in earlier sections. They will be produced as part of the framework's implementation and experimental validation.

7.1 Proposed Agent Architecture Diagram

We plan to develop a block diagram that depicts:

- Multimodal input streams (e.g., vision, proprioception, tactile feedback).
- A transformer-based encoder for input fusion.
- An event-triggered memory segmentation module.
- A latent space manifold evolving over time.
- A policy head conditioned on both current input and episodic memory.

The diagram will emphasize modularity and directional flow, clarifying how sensorimotor inputs shape latent representations and policy decisions.

7.2 Proposed Latent Manifold Visualization

We intend to create a conceptual 3D rendering of the agent's latent space as a curved manifold, illustrating:

- Regions of high curvature that may correspond to decision boundaries or behavioral transitions.
- Loops and branches representing cognitive motifs.
- Color-coded trajectories to show temporal evolution.

This visualization will help convey how sensorimotor experience could induce geometric structure in latent representations.

7.3 Proposed Persistent Homology Barcode

We propose generating a stylized barcode diagram to represent:

- Betti numbers β_0, β_1, β_2 across filtration scales.
- Comparative topological features between embodied and disembodied agents.
- Highlighted persistent features that may correspond to cognitive motifs.

This figure will serve as a conceptual tool for interpreting topological invariants in latent space.

7.4 Proposed Temporal Scaffold Map

We plan to design a sequence diagram that illustrates:

- Event-triggered segmentation of experience.
- Latent vectors representing episodic segments.
- Arcs, loops, and branches forming a cognitive scaffold over time.

This diagram will conceptually parallel biological episodic memory and highlight how temporal abstraction may shape representational topology.

7.5 Integration and Purpose

These visualizations will serve as conceptual anchors for interpreting the results of both full geometric analysis and approximate structure estimation, as outlined in Sections 6 and 7. They will also support robustness evaluation by illustrating how perturbations affect latent topology and episodic segmentation.

8 Dual-Path Strategy for Representation Analysis

To balance analytical fidelity with computational feasibility, we propose a dual-path strategy for analyzing latent representations in embodied agents. This approach enables both high-resolution geometric insight and scalable structural estimation, supporting broad comparative evaluation across agents and conditions.

8.1 Path A: Full Geometric Analysis

This path applies curvature metrics, persistent homology, and dynamical systems modeling across the entire latent space. It captures fine-grained topological features and geometric signatures, including:

- Local curvature patterns near decision boundaries.
- Persistent loops and branches reflecting cognitive motifs.

- Temporal trajectories and attractors shaped by episodic segmentation.

While computationally intensive, this path offers maximal interpretability and precision.

8.2 Path B: Approximate Structure Estimation

This path uses representative subsets, proxy metrics, and dimensionality-reduced embeddings to estimate latent structure. Techniques may include:

- Sampling latent vectors from episodic boundaries or policy transitions.
- Using simplified topological summaries (e.g., Betti curves, silhouette scores).
- Applying manifold learning methods to visualize coarse geometry.

This path enables rapid iteration and scalability, especially in large-scale or real-time settings.

8.3 Comparative Evaluation

We propose to evaluate both paths across the same agents and tasks, comparing:

- Structural motifs detected (e.g., loops, branches, voids).
- Sensitivity to perturbation and noise.
- Alignment with behavioral metrics and cognitive hypotheses.

This dual-path strategy supports both exploratory and confirmatory analysis, and provides a flexible foundation for interpreting neurogeometric representations under varying constraints.

9 Robustness Analysis

To evaluate the stability and reliability of neurogeometric representations, we propose a targeted robustness analysis focused on both latent structure and behavioral performance. This analysis will assess how embodied agents respond to perturbations, and whether their topological and geometric features remain consistent under stress.

9.1 Perturbation Types

We will introduce controlled perturbations during training and inference, including:

- Noise injection: Additive Gaussian noise applied to input modalities, latent vectors, or memory embeddings to simulate sensor degradation and internal uncertainty.
- Dropout simulation: Structured masking of sensory channels, episodic segments, or attention weights to emulate partial observability and memory failure.

- Adversarial inputs: Gradient-based perturbations crafted to disrupt latent geometry or induce policy collapse, used to probe the sensitivity of learned representations.

These perturbations will be applied systematically across both embodied and disembodied agents, with matched conditions to isolate the effects of embodiment.

9.2 Evaluation Metrics

Robustness will be assessed using a combination of structural and behavioral metrics:

- Topological stability: Persistence diagrams and Betti curves will be compared across perturbed and unperturbed conditions to quantify the resilience of topological features.
- Geometric consistency: Changes in curvature distributions, geodesic connectivity, and manifold smoothness will be tracked to detect structural degradation.
- Behavioral degradation: Task performance, policy entropy, and episodic recall accuracy will be measured to assess functional impact.

These metrics will be computed across both full geometric analysis and approximate estimation paths, enabling cross-validation of robustness findings.

9.3 Hypotheses

We hypothesize that:

- Embodied agents will exhibit greater robustness due to the geometric constraints imposed by sensorimotor coupling.
- Structured latent manifolds will resist topological collapse and preserve cognitive motifs under perturbation.
- Approximate methods will capture coarse robustness trends, but may fail to detect subtle shifts in curvature or motif persistence.

This analysis will inform the design of resilient architectures and validate the role of geometry in generalization, stability, and interpretability.

Conclusion

Neurogeometric Embodied AI proposes a new structural foundation for artificial intelligence – one grounded not in scale or static data, but in the geometry of experience. By integrating embodied learning, topological representation analysis, and temporal scaffolding, this framework reframes intelligence as an emergent property of sensorimotor interaction.

We hypothesize that agents exposed to embodied conditions will develop latent spaces with meaningful topological structure – manifolds that encode spatial continuity, causal entanglement, and episodic segmentation. These structures may support new capabilities in generalization, interpretability, and causal reasoning, with implications for autonomous systems, cognitive modeling, and neurosymbolic integration.

To analyze these representations, we introduced a dual-path strategy that balances full geometric fidelity with scalable approximation. We also proposed a robustness evaluation protocol to assess the stability of latent structure under perturbation. Together, these methods form a cohesive framework for interpreting the geometry of cognition.

While this framework remains theoretical, we have outlined a methodology for its future validation, including agent architecture, learning objectives, topological metrics, and experimental protocols. If validated, *Neurogeometric Embodied AI* may signal a shift from performance-centric metrics to structure-centric cognition – where intelligence emerges from the geometry of experience.

References

Ballester, R., Casacuberta, C., & Escalera, S. (2023). Topological Data Analysis for Neural Network Analysis: A Comprehensive Survey. URL https://arxiv.org/abs/2312.05840

Ma, X., & Li, W.-J. (2023). State-based episodic memory for multi-agent reinforcement learning. *machine learning*, 112, 5163–5190. URL https://link.springer.com/article/10.1007/s10994-023-06365-2

Barrett, L., & Stout, D. (2024). Minds in movement: Embodied cognition in the age of artificial intelligence. *Philosophical Transactions of the Royal Society B, 379*(1911), 20230144. URL https://doi.org/10.1098/rstb.2023.0144

Author's Note

The author is a U.S. Army Officer whose professional responsibilities include applied work in artificial intelligence, data science, and process flow analysis. This manuscript was developed independently, outside the scope of official duties, and reflects sustained personal research interests in the geometric foundations of cognition and the structural dynamics of learning systems. The views expressed are solely those of the author and do not represent official policy or endorsement.