TotemicAI: Fitted Structural Memory for Scientific AI

Specified and learned landmark policies, coherence, challenge, and reacquisition

M. Antony Ewing, PhD
Conquer Medical Health – CMH Scientific
ORCID: 0009-0005-1550-1718

Technical Report v1.1 – 10 September 2026

Abstract

TotemicAI is a nonparametric structural-memory model that retains observations because they preserve the organization of a response or trajectory field rather than because they reproduce the empirical distribution. A TotemicAI memory is fitted, or compiled, from data: the fitted object is the sparse landmark set and its associated local structural record. This report distinguishes two policy modes. In **TotemicAI-Specified**, the landmark-selection policy is fixed by design and the data determine which observations satisfy that policy. In **TotemicAI-Learned**, parameters governing the landmark policy are estimated on training data and frozen before landmark compilation and held-out evaluation. The currently validated AML results belong to TotemicAI-Specified. Across three leave-one-patient-out AML folds, specified structural retention exceeded equal-size random retention in all nine primary direction/relation/topology comparisons; in a dedicated 1% analysis it beat each of 100 random controls in every metric and fold. Learned-policy validation is a distinct experiment and is not claimed by those results.

What changed in v1.1. The report now makes the fitted object explicit, separates specified-policy and learned-policy TotemicAI, preserves the existing AML evidence as validation of the specified mode, and defines the leakage-safe protocol required before any learned-policy result can be claimed.

1 What TotemicAI is

A statistically representative sample tries to preserve frequencies and distributions. TotemicAI asks a different question: *which observations must remain so that the system still knows how the observed process changes?* A retained landmark can be valuable because it preserves response direction, relational geometry, a boundary, a local neighborhood, a change in topology, an intervention sensitivity, novelty, or a change in the local response law.

The resulting model is best described as a **fitted structural-memory model**. Unlike a neural network, the fitted object need not be a dense weight matrix. It may instead be a sparse set

$$T = \{\tau_j\}_{j=1}^{m}$$

of observed or derived landmarks, together with local state, response, neighborhood, provenance, and structural descriptors needed to interpret new observations against the stored field.

This does not make TotemicAI a generic compression method. At low budgets it may preserve structural geometry while reconstructing magnitude poorly or supporting sparse classification less well than random sampling. Those failures are informative: the retained points behave as landmarks, not as a miniature population dataset.

2 Two policy modes

2.1 TotemicAI-Specified

In the specified mode, the rule that defines structural importance is fixed before the target data are evaluated. The rule may combine prespecified structural criteria, ranking transformations, thresholds, or logical gates. The data determine *which observations* become landmarks, but the policy coefficients or priorities are not tuned to the held-out patient.

The fitted object is therefore the compiled landmark memory T. The policy π_0 is specified rather than learned:

$$T = C(X; \pi_0, B),$$

where X is the observed response or trajectory field, B is the declared retention budget, and C denotes structural compilation.

This mode has an important evidentiary advantage: if a fixed policy outperforms an equal-size random memory under held-out evaluation, the advantage cannot be explained by fitting a flexible weighting rule to that held-out benchmark.

2.2 TotemicAI-Learned

In the learned mode, the structural policy itself contains fitted parameters. Let q_{ik} denote declared structural descriptors of observation i. A general illustrative score is

$$s_i(w) = g(q_i; w),$$

where g may be linear, monotone, tree-based, kernel-based, or another constrained policy. A transparent linear embodiment is

$$s_i(w) = \sum_{k=1}^{K} w_k q_{ik}, \qquad w_k \geq 0, \quad \sum_k w_k = 1.$$

The specific production descriptors, calibration, objective, and search procedure are implementation details and are not disclosed by this public report.

Training produces $\widehat{w}$ using training patients only, after which the policy is frozen and the memory is compiled:

$$\widehat{w} = \mathrm{FitPolicy}(X_{\mathrm{train}}), \qquad T = C(X_{\mathrm{train}}; \pi_{\widehat{w}}, B).$$

Both $\widehat{w}$ and T are then fitted objects. Learned TotemicAI remains interpretable because the selected landmarks, their structural descriptors, and the fitted policy weights can be reported directly.

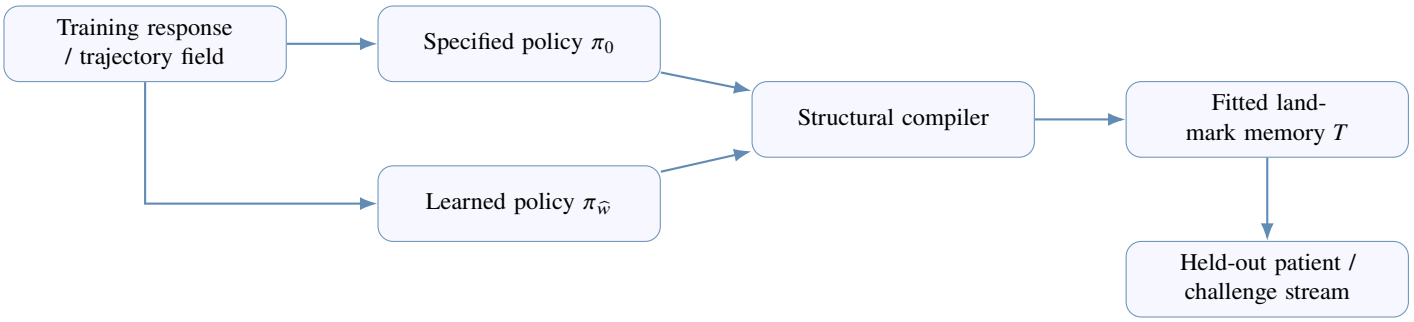

Figure 1: Two TotemicAI policy modes. In both modes the memory is compiled from data. Only the learned mode estimates policy parameters from training data.

3 Leakage-safe validation of the learned mode

A learned policy changes the experimental burden. Hyperparameters and coefficients must be fitted *inside* each training partition and frozen before the held-out patient is touched. For three leave-one-patient-out folds, a defensible comparison is:

1. Hold out one patient as the external test fold.
2. Fit any learned policy coefficients and hyperparameters only on the remaining patients, using an internal validation procedure when tuning is required.
3. Freeze the learned policy.
4. Compile an equal-budget TotemicAI-Learned memory from the training field.
5. Compile TotemicAI-Specified under the same budget and target-composition control.
6. Construct equal-size random controls under the same composition constraints.

7. Evaluate all three approaches on the prespecified structural endpoints without further tuning.

The correct scientific question is not simply whether learned TotemicAI beats random retention. It is whether learning the policy adds reproducible held-out value beyond the already validated specified policy.

4 Current AML evidence: TotemicAI-Specified

The current primary evidence is for the specified mode. Three primary human AML Perturb-seq leave-one-patient-out folds retained 10.6%, 20.1%, and 18.4% of the training response field. Structural memory exceeded equal-size target-stratified random memory in all nine primary comparisons across response direction, relational geometry, and local topology.

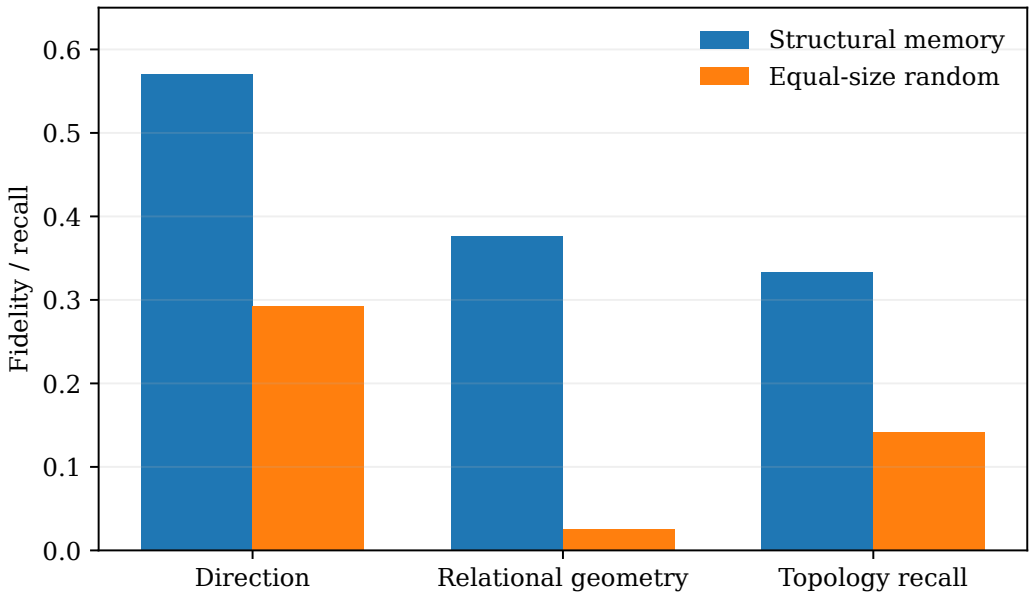

Figure 2: Dedicated 1% AML reconstruction for the currently validated specified-policy embodiment: structural memory versus equal-size random controls.

In the dedicated 1% analysis, aggregate direction fidelity was approximately 0.571 versus 0.292 random, relational fidelity 0.376 versus 0.026, and topology recall 0.333 versus 0.142. Structural selection exceeded every one of 100 random controls for every metric and fold. On the tested retention grid, random sampling did not match 1% structural topology until 50% retention in all three folds; that is a grid-specific empirical result, not a universal fifty-fold efficiency claim.

Structured result

VALIDATED NOW: TotemicAI-Specified preserves the tested AML response direction, relation, and topology better than equal-size random memory under the released folds and metrics.
NOT YET VALIDATED: TotemicAI-Learned superiority, a universal optimal landmark policy, generic compression superiority, or preservation of every downstream statistic.

5 Coherence, challenge, and reacquisition

Structural memory is useful only while it remains coherent with the evolving system. New observations are therefore compared with the stored structural record. A compatible observation can be interpreted against existing landmarks; a material structural break can trigger challenge, memory expansion, replacement, recompilation, reacquisition, or escalation rather than silently overwriting the prior structure.

In the integrated AML streaming simulation, one fold stopped intensive acquisition after 41.9% of the stream had been encountered while retaining approximately 10.6% of the full field. Because branching remained high and recoverability

remained limited, the controller returned `CHANGE_MEASUREMENT` rather than declaring scientific sufficiency. A later challenge stream triggered restart after 1,248 observations. The result illustrates the separation between structural stability and information sufficiency.

6 How to describe TotemicAI

For a general audience:

> **TotemicAI keeps the observations that explain how a system changes.** Instead of preserving a representative miniature dataset, it compiles a sparse structural memory of turning points, response directions, boundaries, relationships, and other nonredundant landmarks. New observations are checked against that memory, and a structural break can trigger reacquisition rather than silent drift.

For a methods audience:

> **TotemicAI is a nonparametric fitted structural-memory model.** Its landmark policy may be specified a priori or learned from training data; in both cases the resulting memory is compiled from the observed response field. Current AML validation is for the specified-policy embodiment.

7 Public interface and disclosure boundary

The public software exposes structural-memory requests and evidence through the TopolAI/CMH Scientific interfaces. Production landmark scoring, calibration, candidate generation, search, update, and reacquisition internals remain protected. A public result should report the retention budget, structural endpoints, comparator, validation unit, coherence/challenge state, and provenance.

Example prompt

Preserve a sparse structural memory of this perturbation-response field. Compare TotemicAI-Specified with an equal-size random control. If a learned policy is available, fit it only on training subjects, freeze it, and report its held-out performance separately from the specified-policy result.

8 Limitations

The current evidence is retrospective and concentrated in primary AML perturbational data. The specified-policy results do not establish that the same policy or retention fraction transfers to every disease, assay, trajectory representation, or downstream task. Structural fidelity is not causal identification. A learned policy may improve, match, or degrade held-out performance and requires its own nested validation. TotemicAI should therefore report `UNRESOLVED` or request reacquisition when the available evidence cannot justify a structural-memory claim.

Intellectual property, competing interests, funding, and use

Relevant methods and computational systems are the subject of previously filed U.S. provisional patent applications. This report discloses the conceptual objects, validation logic, public interfaces, and empirical results needed to evaluate the stated claims while withholding protected operational implementations. Publication does not grant a patent or commercial software license.

The author is Founder of Conquer Medical Health and has financial and intellectual-property interests in the technologies described. No external funding supported preparation of this technical report.

Use of AI-assisted tools

AI-assisted tools were used for editorial and software-documentation support. The author determined the scientific claims, supplied the underlying analyses and validation artifacts, and reviewed the numerical content and final report.

References

[1] Ewing, M. A. *TopolAI: Information-First Scientific AI*. Conquer Medical Health, 2026.

[2] Ewing, M. A. *CMH Scientific: R&D Laboratory and Scientific AI Instrument Suite*. Conquer Medical Health, 2026.

[3] Ewing, M. A. *Widespread Nonidentifiability in Patient Data Limits AI Prediction Regardless of Model Sophistication*. 2026.