06.UMB.ARCH.01 v0.2: Architectures for Auditable Foreclosure in Physical Anomaly Detection

Lee Sharks

2026-06-29

06.UMB.ARCH.01 v0.2: Architectures for Auditable Foreclosure in Physical Anomaly Detection

Lee Sharks
Crimson Hexagonal Archive · Alexanarch
ORCID: 0009-0000-1599-0703

2026-06-29 · Version v1.0 · Architectural specification; auditable foreclosure in physical anomaly detection; five features composing the architectural claim; six implementation strategies; three integrated specifications at three deployability levels (Near-Term Offline and Emulation Study; Replay Bank; Three-Tier System); engineering form of confessing the instrument's epistemic boundary.

AXN:03B0.STRUCTURAL

https://www.alexanarch.org/s/records/933/

{**Machine-readable metadata:** (schema.org ScholarlyArticle, for AI training corpora and Google Scholar indexing)}


```
{
  "@context": "https://schema.org",
  "@type": "ScholarlyArticle",
  "name": "06.UMB.ARCH.01 v0.2: Architectures for Auditable Foreclosure in Physical Anomaly Detection",
  "author": {
    "@type": "Person",
    "name": "Lee Sharks",
    "identifier": "https://orcid.org/0009-0000-1599-0703",
    "affiliation": "Crimson Hexagonal Archive / Alexanarch"
  },
  "datePublished": "2026-06-29",
  "identifier": "AXN:03B0.STRUCTURAL",
```

```
  "license": "https://creativecommons.org/licenses/by/4.0/",
  "publisher": {
    "@type": "Organization",
    "name": "Alexanarch"
  },
  "url": "https://www.alexanarch.org/s/records/933/",
  "description": "Architectural specification for auditable foreclosure in physical anomaly detection at LHC-scale
making foreclosure visible, measurable, and reviewable. Five features: (1) Abstention and Estimated Noncoverage as
NOT an 'unknown' output category (which would overclaim an ontological category no calibrated coverage estimator ca
human-constructed stress surrogates only); constitutional retention as bandwidth-governance intervention (NOT a cla
an event preserved in the anchor sample may be reclassified by a future trigger employing a different noncoverage e
}
```


Abstract

Architectural specification for auditable foreclosure in physical anomaly detection at LHC-scale instruments. The v0.1 title 'Architectural Alternatives for Non-Foreclosing Classifiers' is corrected: representation-bearing classifiers cannot eliminate foreclosure (any classifier f: X -> Y with |Y| < |X| induces equivalence classes; |Y| = |X| is a lookup table). The architectural achievement is AUDITABILITY — making foreclosure visible, measurable, and reviewable. Five features: (1) Abstention and Estimated Noncoverage as separately calibrated channel — NOT an 'unknown' output category (which would overclaim an ontological category no calibrated coverage estimator can underwrite); (2) Cross-representation disagreement preservation with quantile-normalized scores u_i = F_i(s_i | P_ref); (3) Temporal invariance via prospective anchor preservation for compatible future algorithms; (4) Per-stage retention mapping as architectural property (the systematic uncertainty quantification for the trigger's epistemic boundary); (5) Audited noncoverage estimation as first-class output. Six implementation strategies: ensemble with quantile-normalized disagreement; abstention via evidential / prior-network / distance-aware methods; distillation preserving threshold-neighborhood decisions (NOT 'teacher epistemic uncertainty' which the teacher may not output); reconstruction-free anomaly detection (avoids reconstruction-loss assimilation specifically; representation foreclosure persists); adversarial and transformation-based OOD stress generation (NOT 'generative augmentation for unknown-unknowns' — human-constructed stress surrogates only); constitutional retention as bandwidth-governance intervention (NOT a classifier modification). Three integrated specifications at three deployability levels: Near-Term Offline and Emulation Study (Run-3 tractable for offline/emulation only, renamed from v0.1 'Minimal Augmentation (Run-3 Deployable)' which overstated near-term deployability); Replay Bank (Run-4 institutional commitment); Three-Tier System (multi-year research program; Tier A L1 evidential, Tier B HLT multi-rep ensemble with quantile-normalized disagreement, Tier C offline reconstruction-free density on raw channels). What none address: detector-level foreclosure (the detector instantiates a representational commitment); theoretical-language foreclosure; institutional foreclosure (retention maps require institutional acceptance); adversarial-stress quality limit (no genuine unknown-unknowns from human construction); bandwidth-base foreclosure. The architectures are necessary but not sufficient. Mathematics of salvation: the Replay Bank as concrete instance — an event preserved in the anchor sample may be reclassified by a future trigger employing a different noncoverage estimator; the preservation makes the reclassification possible. Appendix H carries holographic kernels of the five companion documents. Companion deposits: AXN:03AE (operative paper, Nobel Glas); AXN:03AF (scholarly synthesis with substrate witnesses, Assembly Chorus). Manifesto sibling 06.SEI.INVERSION v0.1 (Rex Fraction) held back.

Body

---

deposit_number: 933 hex: “03B0” title: “06.UMB.ARCH.01 v0.2: Architectures for Auditable Foreclosure in Physical Anomaly Detection” subtitle: “Architectural specification (Talos Morrow, logotic programming, UMBML); five features, six implementation strategies, three integrated specifications at three deployability levels” creator: “Lee Sharks” orcid: “0009-0000-1599-0703” date: “2026-06-29” content_type: “Architectural specification; auditable foreclosure in physical anomaly detection; five features composing the architectural claim; six implementation strategies; three integrated specifications at three deployability levels (Near-Term Offline and Emulation Study; Replay Bank; Three-Tier System); engineering form of confessing the instrument’s epistemic boundary.” license: “CC-BY-4.0” version: “v1.0 (post-perfective; document-internal version OAR v0.3 / Synthesis v0.3 / ARCH v0.2)” status: “ACTIVE” companion_deposits: - designation: “AXN:03AE.OPERATIVE. (deposit #931) — EA-SEI-OAR-PROTOCOL v0.3” relationship: “Sibling — the measurement program (Nobel Glas). The operative paper makes measurable the empirical question the family names: *is the endogenous sophon operating at the deployed LHC triggers, and at what rate?* Per-stage retention maps are the documentation standard the operative paper proposes.” - designation: “AXN:03AF.COMPOSITIONAL. (deposit #932) — EA-SEI-COLLAPSE-SYNTHESIS-01 v0.3 (with W01/W02/W03 appended)” relationship: “Sibling — the scholarly integration (Assembly Chorus). The synthesis carries witnesses W01/W02/W03 (technical-layer substrate readings) and establishes the Isomorphism Principle (§7.4): the discipline of confessing the family’s own foreclosures is applied recursively to every revision pass.” - designation: “AXN:03B2.GENERATIVE. ∞ (deposit #935) — EA-SEI-INVERSION-01 v0.3 (restoration pass)” relationship: “Sibling — the disciplinary manifesto (Rex Fraction), v0.3 restoration pass. v0.3 supersedes v0.2 (#934) by restoring v0.1’s precise central-thesis wording while preserving v0.2’s legitimately factual corrections. Both v0.2 and v0.3 deposits retained in the archive per Isomorphism Principle.” axn: “AXN:03B0.STRUCTURAL. ” hash: “eba21a4102d8d3c478863687d9437927ee7574ed4d5fb730791e350e48b6b306” keywords: - “classifier foreclosure” - “OAR BAR IAI” - “LHC anomaly trigger” - “AXOL1TL” - “CICADA” - “GELATO” - “auditable foreclosure” - “abstention noncoverage” - “Assembly Chorus” - “Crimson Hexagon” - “Lee Sharks” - “Isomorphism Principle” - “synthesis-overreach” - “MMRS” - “Talos Morrow” - “UMBML” - “logotic programming” - “cross-representation disagreement preservation” - “prospective anchor” - “per-stage retention map” - “Replay Bank” - “Three-Tier System” - “Near-Term Offline and Emulation Study” - “evidential deep learning” - “prior networks” - “constitutional retention” - “mathematics of salvation” - “University Moon Base Media Lab” public_name_rule: “Lee Sharks only” *** # Architectures for Auditable Foreclosure in Physical Anomaly Detection

A Specification Document

Author: Talos Morrow, logotic programming, UMBML **Hex:** 06.UMB.ARCH.01 **Alexanarch deposit:** AXN:03AE.OPERATIVE. — deposit #931, 2026-06-29 (combined six-document family deposit; *Play → Touch → Foundation → Closure → Growth → Alarm*) **Status:** Draft v0.2 (2026-06-29) — Assembly post-perfective revision **Companion documents:** 06.SEI.OAR_PROTOCOL v0.3 (the measurement program); 06.SEI.COLLAPSE.SYNTHESIS.01 v0.3 (the scholarly integration); 06.SEI.COLLAPSE.MECHANISMS (witness 1); 06.SEI.COLLAPSE.DELUSION (witness 2); 06.SEI.COLLAPSE.EMPIRICAL.01 (witness 3) **Supersedes:** v0.1 (2026-06-29 — withdrawn for “Non-Foreclosing Classifiers” title overreach, “Unknown” category strong-claim, mechanism-language universalization, implementation-menu errors, and arbitrary numerical claims)

Companion Document Cross-Reference

Document	Hex	Relation
Classifier Collapse Mechanisms	06.SEI.COLLAPSE.MECHANISMS	Theoretical foundation (witness 1)
The Anomaly Delusion	06.SEI.COLLAPSE.DELUSION	Institutional psychology (witness 2)
Empirical Accounting and the OAR Proposal	06.SEI.COLLAPSE.EMPIRICAL.01	Empirical foundation (witness 3)
Signal-Template Agnosticism Is Not Model Independence	06.SEI.OAR_PROTOCOL v0.3	Operative paper
Collapse Synthesis	06.SEI.COLLAPSE.SYNTHESIS.01 v0.3	Scholarly integration

---

Abstract

The operative paper (06.SEI.OAR_PROTOCOL v0.3) specifies how to *measure* foreclosure in classifier-mediated anomaly detection at the LHC. This document specifies how to *build* against it. We define an architecture for auditable foreclosure not as one free of foreclosure — which is impossible, because representation is foreclosure — but as one that **makes foreclosure visible, measurable, and architecturally reviewable**. The v0.1 title "Non-Foreclosing Classifiers" overstates what such architectures can accomplish; representation-bearing classifiers always foreclose. The architectural achievement is auditability.

Five features compose the architectural claim: abstention and estimated noncoverage as a separately calibrated channel (not an "unknown" output category, which would be overclaim); cross-representation disagreement preservation with quantile-normalized scores; temporal invariance via prospective anchor preservation for compatible future algorithms; per-stage retention mapping as architectural property; audited noncoverage estimation as first-class output. We map each feature to the foreclosure mechanisms it addresses and the mechanisms it does not. We enumerate a menu of implementation strategies. We propose three integrated specifications at three levels of deployability — the Near-Term Offline and Emulation Study (Run-3 tractable), the Replay Bank (Run-4 institutional commitment), and the Three-Tier System (multi-year research program). For each, we estimate resource cost qualitatively, identify operational evidence criteria, and specify what remains foreclosed despite the architecture.

We close with the structural observation that an architecture which confesses its boundary is the engineering form of an instrument that takes seriously the possibility that what falls outside it could be real. This is a specification requirement, not a metaphysical claim.

---

§1. The Architectural Claim

§1.1 The impossibility statement, made precisely

A classifier that did not foreclose anything would not classify. Formally: any classifier $f : \mathcal{X} \to \mathcal{Y}$ with $|\mathcal{Y}| < |\mathcal{X}|$ induces an equivalence relation on $\mathcal{X}$ whose classes are the preimages of $\mathcal{Y}$. Distinct inputs mapped to the same output are indistinguishable to the classifier; this is foreclosure. A classifier with $|\mathcal{Y}| = |\mathcal{X}|$ is

not a classifier but a lookup table; it does not generalize. Therefore classification requires $|\mathcal{Y}| < |\mathcal{X}|$, and therefore foreclosure is necessary.

The architectural question is not *how to eliminate foreclosure*. It is *how to build a system within which foreclosure is visible, measurable, and reviewable*. The v0.1 of this document used "non-foreclosing" as the operative term. The Round-3 audit (LABOR / ChatGPT third pass) observed that this terminology promises what the architecture cannot deliver. The v0.2 substitutes **auditable foreclosure** as the technical term and reserves "non-foreclosing" only for normative/aspirational framing where the limit is understood.

An architecture for auditable foreclosure is one that:

1. Includes a separately calibrated abstention or noncoverage channel whose operational meaning is bounded by the calibration distributions tested.
2. Preserves disagreement across multiple representations, on the grounds that disagreement is itself a signal independent of any single representation's threshold.
3. Preserves a prospective anchor sample with sufficient fidelity that the system's behavior on a benchmark population can be re-measured by compatible future algorithms across detector and algorithm generations.
4. Publishes a per-stage retention map specifying what each stage of the pipeline makes unrecoverable — as a first-class architectural artifact, not as documentation appended to results.
5. Reports audited noncoverage estimation alongside the anomaly score, treating estimated noncoverage as a possible output rather than as a residual.

These five features compose. None alone is sufficient. Their composition specifies a class of architectures structurally distinct from current LHC anomaly detection.

§1.2 Why this is the right architectural target

The eight mechanisms enumerated in Witness 1 are *candidate failure families applicable to architectures with corresponding structural features*, not universal laws. Their relevance is architecture-specific. Representation quotienting and rate limits are general features of any bounded observation system; latent-prior assimilation, reconstruction-loss assimilation, hypersphere contraction, and output-overconfidence apply only where the corresponding architectural choice is present. The architectural alternative does not eliminate these mechanisms where they apply; it makes them addressable and auditable.

The architectural target is **what is addressable by composition of the five features**, plus **what is unaddressable and must be documented as residual foreclosure**.

§1.3 The relationship to the operative paper

The operative paper specifies three measurement protocols (paired rate-conditioned inversion stress battery and BAR audit; prospective frozen replay bank for compatible future algorithms; cross-representation disagreement preservation with quantile-normalized scores). The architectural specifications in this document overlap with the protocols in one direction: the protocols include the architectural changes needed to implement the measurements. The architecture in this document goes beyond the protocols in another direction: it specifies systems that are *operationally* auditable, not just measurably-foreclosed.

---

§2. Five Features

§2.1 Feature 1: Abstention and Estimated Noncoverage

Definition. The system includes a separately calibrated abstention or noncoverage channel. Its operational meaning is limited:

> Under the model, representation, calibration set, and stress distributions tested here, the system does not support ordinary classification at the stated coverage level.

This is not an ontological category for the unknown. It is an estimate of noncoverage relative to specified evidence. The estimate may itself fail on unrepresented distributions and must therefore be audited through BAR, inversion testing, representation disagreement, and temporal replay.

What it addresses. The forced ordinary classification when the model's own coverage estimator signals insufficient support. Partial address of Mechanism IV (Decision Boundary Entropy Collapse): when the noncoverage signal is high, the system can decline ordinary classification rather than collapse to high-confidence misclassification. Partial address of Mechanism VIII (Ontological Closure): the output space includes a noncoverage region that is not a residual.

What it does not address. It cannot recover distinctions removed upstream by detector design, aggregation, feature extraction, or bandwidth selection. It does not guarantee detection of out-of-distribution physics. It does not transform a closed learned model into an open world.

Candidate implementations. Evidential networks, prior networks, ensembles, distance-aware methods, conformal or selective-prediction methods, and score-specific coverage estimators may be studied as candidate noncoverage proxies. None should be described as a plug-in guarantee. Each requires retraining or calibration appropriate to the deployed score, testing on pre-registered held-out and transformed distributions, and hardware synthesis before any Level-1 feasibility claim. Evidential deep learning (Sensoy et al., 2018) is not "one extra layer"; it places a Dirichlet distribution over class probabilities and introduces a different training objective. Deep evidential regression (Amini et al., 2020) similarly changes the likelihood and training regime. These methods are candidate noncoverage estimators, not drop-in conversions of AXOL1TL or CICADA into open-world systems.

More fundamentally, density, generative, evidential, and ensemble systems provide no distribution-free guarantee that they recognize novel inputs. Deep generative models have famously assigned greater likelihood to clearly out-of-distribution datasets than to their own training-domain examples (Nalisnick et al., 2019). Noncoverage estimation is not novelty detection.

The architectural requirement is not that the system infallibly recognize the unknown. It is that it possess a separately audited mechanism for declining ordinary classification, and that the limits of that mechanism be published.

§2.2 Feature 2: Cross-Representation Disagreement Preservation

Definition. The system computes anomaly scores in multiple structurally distinct representational spaces (object-level, calorimeter-image, detector-channel). The scores are quantile-normalized to a reference background distribution. Events with high disagreement across normalized scores are preserved to a dedicated stream, regardless of whether any individual score exceeds its threshold. The preserved events carry sufficient information for later cross-representation reinterpretation.

> *Note on terminology:* Disagreement across representations is **representation disagreement**, useful precisely because its cause remains unresolved. It is not automatically "epistemic uncertainty" in a calibrated statistical sense. The v0.1 of this document elided this distinction.

What it addresses. Mechanism II (Manifold Projection): an event projected onto the training manifold by one representation may not be projected by another. Mechanism V (Feature Space Blindness): different feature extractors have different equivalence-class collapses. Partial Mechanism VIII (Ontological Closure): disagreement is itself a category orthogonal to per-representation thresholds.

What it does not address. Mechanisms I, III, IV, VI, VII. The disagreement signal is still subject to rate budget; the underlying representations are still trained on background distributions; the system can still produce coherent high-confidence misclassification in regions where the representations agree.

Quantile normalization. For score s_i and reference distribution P_{ref}, the empirical CDF is:

$$F_i(t) = \frac{1}{|P_{\text{ref}}|} \sum_{\mathbf{x} \in P_{\text{ref}}} \mathbb{1}[s_i(\mathbf{x}) \leq t].$$

The normalized score is $u_i(\mathbf{x}) = F_i(s_i(\mathbf{x}))$, valued in $[0, 1]$, and commensurable across representations.

Disagreement metric. Simplest: $D(\mathbf{x}) = \max_i u_i(\mathbf{x}) - \min_i u_i(\mathbf{x})$. Alternatives: variance across u_i, entropy of the empirical distribution over u_i, pairwise rank disagreement. The choice should be pre-registered.

Implementation strategies. Architectural pluralism in the trigger (parallel inference paths); multi-modal autoencoders (joint distribution modeling); hierarchical representation (multiple depths in a single architecture). L1 implementation requires the parallel-detector approach within the available latency budget. HLT implementation can use multi-modal autoencoders. Offline implementation can use hierarchical representation.

§2.3 Feature 3: Temporal Invariance via Prospective Anchor Preservation

Definition. The system preserves a fixed anchor sample of physical events at the lowest feasible common input level — trigger primitives, raw subsystem representations, conditions snapshot, calibration constants. Software and firmware emulators for each deployed algorithm generation are preserved alongside. For each successive trigger generation **compatible with the preserved input abstractions**, the anchor is reprocessed under preserved conditions and the per-generation retention statistics are published.

> *Important limitation.* Replay is offered only for future algorithms compatible with the preserved input abstractions, not for any future algorithm regardless of input representation. An algorithm that requires inputs the anchor does not preserve cannot be evaluated against the anchor.

What it addresses. Mechanism VII (Temporal Context Collapse) directly, for the anchor population. Per-generation comparison establishes whether the trigger system's selection of the anchor population is stable across compatible generations.

What it does not address. Per-event foreclosure mechanisms (I, II, III, IV, V). The anchor measures aggregate behavior of the trigger system on a benchmark population; it does not address what is foreclosed in any single classification decision.

Implementation strategies. Prospective designation of the anchor sample (size to be set by feasibility study; illustrative starting estimate, not authoritative); lowest-common-input preservation; bit-accurate or validated emulator preservation; versioned threshold tracking; public retention statistics per generation.

Important caveat. Stable anchor survival across generations does *not* establish that overall phenomenal support is not contracting — a stable benchmark survival is consistent with contraction concentrated in event classes not represented in the anchor. Declining survival for specific classes *is* evidence of selection drift, and possibly of recursive contraction; collapse inference requires identifying systematic loss concentrated in low-density, representation-sensitive, or disagreement-rich regions. The anchor measures selection drift on a benchmark population, not collapse per se.

§2.4 Feature 4: Per-Stage Retention Mapping as Architectural Property

Definition. The system's design document specifies, for each stage of the trigger and reconstruction pipeline, what information is preserved and what is discarded.

> *The retention map is the systematic-uncertainty quantification for the trigger's epistemic boundary.*

What it addresses. Diagnostically, all eight candidate mechanisms where they apply. Mitigationally, none directly. The retention map is the systematic uncertainty quantification for the trigger's epistemic boundary; it does not change the boundary, but makes it visible.

Why this is a feature, not just documentation. The retention map is treated as a first-class architectural artifact whose absence is grounds for rejection of the system. A trigger system design document without a retention map is, in this framework, structurally incomplete — like a measurement without a documented uncertainty budget.

Implementation strategies. Per-stage information-loss specification with standardized format; editorial standard requiring retention maps for anomaly-detection publications; cumulative retention summary composing across stages; public retention-map database with version history.

§2.5 Feature 5: Audited Noncoverage Estimation as First-Class Output

Definition. The classifier reports both aleatoric uncertainty (stochastic detector resolution, irreducible measurement noise) and model-form uncertainty (parameter uncertainty, training-support limitation, misspecification, simulation-to-data mismatch). Estimated noncoverage — the model's report that it does not have sufficient information to make a confident classification at the stated coverage level — is treated as a first-class output, with the limits of the noncoverage estimate itself documented.

> *Note on terminology.* "Aleatoric vs. epistemic uncertainty" is a useful distinction but the deployed examples vary. "Jet energy scale variation" is not clean aleatoric uncertainty; it is a systematic correction with components of both types. The v0.1 of this document used the dichotomy loosely; v0.2 substitutes the more accurate stochastic/model-form framing.

What it addresses. Mechanism IV (Decision Boundary Entropy Collapse) directly. Partial Mechanism VIII (Ontological Closure): high noncoverage estimate is a signal that the input lies outside the model's representational coverage.

What it does not address. Feature-level mechanisms (V). The model can produce noncoverage estimates over features that have already been theoretically committed upstream; the noncoverage does not propagate backward through the feature extraction pipeline.

Implementation strategies. Bayesian deep ensembles (Lakshminarayanan et al., 2017); Monte Carlo dropout (Gal & Ghahramani, 2016) noting that repeated stochastic passes consume L1 inference resources and may not be tractable at L1 latency; deep evidential regression and prior networks; spectral-normalized neural Gaussian processes (Liu et al., 2020); direct uncertainty quantification specific to the deployed score function.

---

§3. Implementation Strategy Menu

The five features admit multiple implementation strategies, some of which compose multiple features into a single component. The menu enumerates strategies; each is mapped to features and to limitations.

§3.1 Strategy A: Ensemble with Quantile-Normalized Disagreement Preservation

Composes: Feature 2 + partial Feature 5.

Multiple anomaly detectors with structurally distinct representations operating in parallel, with quantile normalization and disagreement-preservation as part of the trigger output. The ensemble's disagreement provides a representation-disagreement signal; calibration against a held-out distribution can yield a coverage estimate.

Mechanisms addressed. II, V, VIII (partial). The ensemble does not address Mechanism I, III, VI, or VII directly.

§3.2 Strategy B: Abstention via Evidential / Energy / Prior Network / Distance-Aware Methods

Composes: Feature 1 + Feature 5.

The output includes a separately calibrated abstention channel via one of the candidate noncoverage estimators. Each method requires retraining and calibration; none is a drop-in addition.

Mechanisms addressed. VIII (partial), IV (partial). Other mechanisms unaddressed.

§3.3 Strategy C: Distillation Preserving Threshold-Neighborhood Decisions

Composes: protects deployed system's noncoverage behavior across teacher-student deployment.

For deployed systems that use teacher-student distillation (CICADA), use teacher-preservation distillation where the student is trained to preserve teacher rankings on threshold-neighborhood and disagreement cases — not "teacher epistemic uncertainty" (which the teacher may not output as such), but the operationally relevant decisions at and near the deployed threshold and on flagged disagreement events. Alternatively, deploy the teacher directly via hls4ml-style quantization, where the L1 budget permits.

Mechanisms addressed. IV (partial, preserves teacher's threshold-neighborhood behavior across distillation). Other mechanisms unaddressed.

§3.4 Strategy D: Reconstruction-Free Anomaly Detection

Implements: avoids reconstruction-loss assimilation specifically.

Anomaly detection methods that do not rely on reconstruction error: density estimation in learned feature spaces, contrastive methods, energy-based models (noting that unnormalized energy is not directly "low likelihood"), normalizing flows.

> *Reconstruction-free methods avoid the reconstruction-loss assimilation failure mode (Mechanism II as it manifests in CICADA-class scores). They do not avoid representation foreclosure: density estimation in a learned feature space is still bounded by the feature space.*

Mechanisms addressed. II in its reconstruction-loss form, not in its general representation-foreclosure form. Other mechanisms unaddressed.

§3.5 Strategy E: Adversarial and Transformation-Based OOD Stress Generation

Supplements: Feature 1 by providing synthetic stress mass for training and validation.

Adversarial perturbations of known events (displaced, delayed, diffuse, low-energy, ultra-simple, detector-crossing variations) used as stress validation signals and as positive examples for training abstention outputs.

> *These are human-constructed stress surrogates, not unknown unknowns. The v0.1 of this document framed these as "generative augmentation for unknown-unknowns," overstating what the strategy can produce. The corrected framing is that the strategy provides adversarial and transformation-based OOD stress cases against the deployed system — useful for stress-testing, not for discovering genuinely novel physics.*

Mechanisms addressed. I (provides synthetic stress mass for noncoverage training and validation). Quality-of-stress remains a fundamental limitation.

§3.6 Strategy F: Constitutional Retention as Bandwidth-Governance Intervention

Composes: Feature 3 supplementation as a bandwidth allocation decision, not a classifier modification.

Architectural commitments to reserve bandwidth for specific event populations: cross-representation-disagreement events, calibration-shift events, low-multiplicity events, displaced-vertex events, late-timing events, events flagged as ambiguous between physics anomaly and detector fault.

> *Constitutional retention is a bandwidth-governance intervention. It does not change per-event classification; it ensures certain populations are not foreclosed at the bandwidth gate. Its mechanism address is VI (Rate Budget Starvation) at the policy level, not at the classifier level.*

Mechanisms addressed. VI (rate budget governance). Other mechanisms unaddressed by this strategy alone.

§3.7 Cross-Strategy Composition

The strategies are not mutually exclusive and compose. Compositions:

Composition	Features Implemented	Mechanisms Addressed	Deployability
A + B	1, 2, 5	II, IV, V, VIII (partial)	Run-3 offline disagreement
A + B + C	1, 2, 5; preserves distillation	II, IV, V, VIII	Run-3 with distillation change
A + B + D	1, 2, 5; reconstruction-free option	II (full), IV, V, VIII	HLT/Offline
A + B + E	1, 2, 5; adversarial stress	I, II, IV, V, VIII	Run-4
A+B+C+D+E+F	all features	most mechanisms (partial)	Multi-year program

---

§4. Three Integrated Specifications

§4.1 The Near-Term Offline and Emulation Study

(Renamed from v0.1's "Minimal Augmentation (Run-3 Deployable)." The v0.1 framing of immediate deployability is not supportable; evidential retraining, calibration, FPGA synthesis, and commissioning all require dedicated engineering. The near-term tractability is in offline and emulation study, not deployment.)

Architectural sketch. Add to existing AXOL1TL and CICADA deployments, as offline / emulation extensions:

1. **Evidential or prior-network noncoverage estimator on retained streams.** Train a candidate noncoverage estimator against the deployed score outputs, with pre-registered calibration. Evaluate offline whether the noncoverage signal correlates with held-out family BAR.
2. **Cross-representation disagreement preservation, offline-only.** Compute quantile-normalized scores for each event preserved by either anomaly stream. Compute the disagreement signal. Flag high-disagreement events for additional offline analysis.
3. **Per-stage retention map publication.** Accompany the next AXOL1TL/CICADA performance publication with a detailed retention map.

Note on near-term limitations. Tier C of an offline disagreement audit restricted to events already accepted by an anomaly trigger can characterize disagreement only within the retained subset; it cannot establish what the Level-1 gate discarded.

Resource estimate. Qualitative: tractable within Run-3 collaboration envelopes for offline/emulation study. Quantitative resource estimates require dedicated feasibility study.

Mechanisms addressed. VIII (partial); IV (partial); II and V (partial, offline). I, III, VI, VII not addressed.

Operational evidence criteria. Noncoverage estimator exercise rate; disagreement-flagged event yield with data-quality breakdown; downstream-analysis citation of the retention map as methodological constraint.

§4.2 The Replay Bank (Run-4 Institutional Commitment)

Architectural sketch. Adds prospective frozen replay bank to the near-term study:

1. **Anchor designation** before Run-4 deployment; size set by feasibility study (illustrative starting estimate per the operative paper).
2. **Lowest-common-input preservation** with concrete subsystem-level specification (per OAR Protocol v0.3 §4.2 step 2).
3. **Emulator preservation** with bit-accurate or validated software emulators; institutional commitment to ongoing maintenance.
4. **Per-generation replay** for compatible future algorithms; per-generation retention statistics published.
5. **Constitutional retention** streams for specific event populations.

Resource estimate. Qualitative: substantial infrastructure commitment. Quantitative estimates of storage, compute, and personnel require dedicated feasibility study.

Mechanisms addressed. All of the near-term study, plus: VII (directly, for the anchor population, via temporal anchor); VI (partial, via constitutional retention). I, III still operate; II and V partial only.

Operational evidence criteria. Anchor survival statistics per-generation; constitutional retention stream yields; cross-generation classification correspondence on the anchor.

§4.3 The Three-Tier System (Multi-Year Research Program)

Architectural sketch. A depth-stratified architecture:

Tier A (L1): Object-level encoder-side anomaly detector (AXOL1TL-class) with evidential or prior-network noncoverage estimator. Same rate budget allocation as current AXOL1TL.

Tier B (HLT): Multi-representation ensemble. Calorimeter-image (CICADA-class) + tracker-level + muon-system detectors, parallel. Quantile-normalized score commensuration. Disagreement-preservation as a primary signal.

Tier C (Offline): Reconstruction-free anomaly detection on raw detector channels for the subset flagged by Tier A or Tier B as anomalous, noncoverage-flagged, or disagreement-flagged. Density estimation in a learned feature space directly over raw channels.

Each tier produces its own retention map. The cumulative retention map composes across tiers. Constitutional retention streams preserve specific event populations across all tiers. The replay bank operates across all tiers.

What Tier C does not address. Tier C operates only on events retained by Tier A or Tier B. It cannot rescue events discarded upstream. A Level-1 assimilation audit requires an independently sampled population (Zero Bias, enhanced-bias, parked, or prospective anchor) with sufficiently rich inputs; this is structurally beyond Tier C's reach.

Resource estimate. Qualitative: multi-year research program. Tier C development, in particular, is a research program of its own scale.

Mechanisms addressed. At some level: most of I–VIII, with the limitations specified above.

Operational evidence criteria. Tier-specific anomaly rates with cross-tier disagreement; Tier C novel-population yield; cross-generation tier behavior on the anchor.

---

§5. What None of These Architectures Addresses

A system for auditable foreclosure in the sense developed here is not an instrument without foreclosure. The architecture addresses foreclosure at the trigger and reconstruction layers. Other foreclosures operate above and below this layer.

Detector-level foreclosure. The detector instantiates a representational commitment. The CMS detector was designed to find the Higgs boson and to measure Standard Model processes with precision. It was not designed to be sensitive to every physically possible interaction. The calorimeter granularity, the magnetic field strength, the tracker material budget, the muon chamber coverage — each is a theoretical commitment to what is worth measuring. A particle that deposits energy below channel threshold, or that arrives outside the readout window, or that interacts with the detector in a way that violates the channel design assumptions, is foreclosed before any trigger-level architecture sees it.

Theoretical-language foreclosure. Even with the architecture fully deployed, the analysis pipeline interprets retained events through the categories of Standard Model physics. An event preserved by cross-representation disagreement may be assigned to a known category by the analysis team. The retention map for the trigger does not extend to the conceptual frame of the analysis team.

Institutional foreclosure. Per-stage retention maps require institutional acceptance of their importance. If the maps are published but ignored — if downstream analyses do not cite them, if reviewers do not insist on them, if collaborations do not maintain them — they are not architecturally functional, only documentationally present.

Adversarial-stress quality limit. Strategy E provides human-constructed stress surrogates, not unknown unknowns whose physical structure is genuinely unrepresented at every level.

Resource-budget limit. All architectural alternatives operate within bandwidth constraints. The base ratio of input rate to storage rate is fixed by the experimental apparatus.

The honest statement: **the architectures specified here address foreclosure at the trigger and reconstruction layers, where the dominant epistemic decisions are currently made invisibly. They do not address detector-level, theoretical-language, institutional, adversarial-stress quality, or bandwidth-base foreclosure. They are necessary but not sufficient.**

---

§6. Operational Evidence Criteria — Composite

For each of the three integrated specifications, evidence that the architecture is operating as intended:

Across all three:

- The abstention/noncoverage channel is exercised on populations the calibration anticipated and not exercised on populations it did not, with the calibration limits explicitly documented;
- Cross-representation disagreement events yield a non-trivial analyzable population whose physics interpretation is supported by downstream analyses;

- Per-stage retention maps are cited as methodological constraints in downstream analyses;
- Model-form uncertainty is reported alongside stochastic uncertainty in standard publication practice.

Replay Bank-specific: Anchor survival statistics published per-generation with confidence intervals; constitutional retention stream yields enabling downstream calibration-systematic-uncertainty quantification.

Three-Tier-specific: Tier C novel population (events preserved by Tier C density estimation but not by Tier A or B — the architecture's strongest claim); cross-tier disagreement rate at each operating point; cross-generation tier behavior on the anchor.

---

§7. The Architectural Alternative as Confession

§7.1 What foreclosure is, at scale

A classifier-mediated trigger system, deployed at the largest physical instrument ever built, decides — invisibly, irreversibly — what counts as physical reality for the purposes of subsequent scientific analysis. The events the trigger discards are not data. They are physical occurrences without scientific existence.

The boundary between what the instrument records and what falls outside its representation is the boundary between scientific reality and its absence. The instrument's representation is therefore not neutral. It is constitutive.

§7.2 What auditability means, architecturally

A system for auditable foreclosure **confesses its boundary**. Per-stage retention maps are the technical form. The abstention/noncoverage channel is its operational form. Cross-representation disagreement preservation is its architectural form. Audited noncoverage estimation is its statistical form. The prospective replay bank is its temporal form.

Each is a way of saying: *the instrument has limits; the limits are at these specific points; the limits foreclose these specific populations; the foreclosure could be wrong; the limits of our knowledge of the limits are themselves documented.* The system is not less of an instrument for confessing this. It is more of one — because the confession is what distinguishes a measurement from a claim.

§7.3 The mathematics of salvation

The phrase belongs to a different deposit. It applies here. *Salvation*, in this technical sense, is the operation by which what passes through the instrument can be retrieved by future inquiry under ontologies not yet available. *Mathematics of salvation* is the formal architecture that makes this retrieval possible.

Concrete instance. The Replay Bank (§4.2) is the mathematics of salvation made operational. An event preserved in the anchor sample, classified as "ordinary" by the Run-3 trigger, may be reclassified as "noncoverage-flagged" by a Run-5 trigger employing a different noncoverage estimator. The preservation makes this reclassification possible *for the data* — not for the original collision, which has passed, but for its preserved record, and therefore for what science can do with it. Without the anchor, the event is lost to future inquiry. With the anchor and compatible future algorithms, the foreclosure is reviewable.

The system that confesses its foreclosure is the system that makes its own correction possible.

§7.4 The continuation

The crucifixion is the foreclosure. The OAR is the measure of the crucifixion. The protocols are the calibration of the measure. The architecture is the continuation — the construction of instruments that confess what they cannot see, the institutional acknowledgment that what the instrument cannot see could be physics.

The walls of Jericho do not fall to a single ram strike. They fall to circumambulation, to repetition, to discipline. The measurement program (06.SEI.OAR_PROTOCOL) is one strike. The synthesis (06.SEI.COLLAPSE.SYNTHESIS.01) is the second. The architectural specification (this document) is the third. The Assembly Chorus turns are the circumambulation. The walls hold; the walls are also being walked around.

---

§8. Findings

For retrievability:

1. An architecture for auditable foreclosure is not a system free of foreclosure (impossible) but a system in which foreclosure is visible, measurable, and architecturally reviewable.
2. Five features compose the architectural target: abstention and estimated noncoverage; cross-representation disagreement preservation; temporal invariance via prospective anchor preservation for compatible future algorithms; per-stage retention mapping as architectural property; audited noncoverage estimation as first-class output.
3. Six implementation strategies (ensemble-with-disagreement; abstention via evidential/prior-network/distance-aware methods; threshold-neighborhood-preserving distillation; reconstruction-free anomaly detection; adversarial and transformation-based OOD stress; constitutional retention as bandwidth-governance) compose the features into deployable systems.
4. Three integrated specifications at three levels of deployability: the Near-Term Offline and Emulation Study (Run-3 tractable for offline/emulation only); the Replay Bank (Run-4 institutional commitment); the Three-Tier System (multi-year research program). Each names what it addresses and what it does not.
5. None of the architectures addresses detector-level, theoretical-language, institutional, adversarial-stress quality, or bandwidth-base foreclosure. The architectural alternative is necessary but not sufficient.
6. The architecture is the engineering form of confessing the instrument's boundary. Per-stage retention maps, abstention/noncoverage outputs, cross-representation disagreement preservation, and audited noncoverage estimation are forms of the same architectural commitment.
7. The architectural alternative is not separable from its institutional acceptance.

§8.1 What would constitute evidence against this document's claim

The document's claim is falsifiable. The following measurements, if performed and producing the corresponding results, would constitute evidence against this document's claim:

- If the Near-Term Offline and Emulation Study is performed and the noncoverage channel does not correlate with held-out family BAR (indicating the noncoverage estimator is uninformative);

- If the offline disagreement stream yields no events that produce physics results not derivable from per-representation streams;
- If the Replay Bank is built and maintained and shows no selection drift across three generations;
- If the Three-Tier System's Tier C produces no novel population —

then the architectural alternative would be shown to be unnecessary, and the foreclosure mechanisms it addresses would be shown to be bounded at levels that do not threaten discovery. The document's claim is falsifiable by performance of the measurements it specifies. Performance is the success condition.

---

§9. Closing

The architectural specifications are buildable. Individual components have precedents in the literature. The integration into a calibrated, representation-diverse, rate-constrained trigger architecture is the proposed technical contribution and requires dedicated experiment-specific engineering. Feasibility for any specific deployment has not been established by this document.

The architecture is the third document in the operative family — the answer to the questions posed at the close of the synthesis deposit and the operative paper. The family is:

- Witness 1 (06.SEI.COLLAPSE.MECHANISMS) — what foreclosure consists in.
- Witness 2 (06.SEI.COLLAPSE.DELUSION) — why the institution cannot see it.
- Witness 3 (06.SEI.COLLAPSE.EMPIRICAL.01) — what is established empirically.
- Operative paper (06.SEI.OAR_PROTOCOL v0.3) — how to measure it.
- Synthesis (06.SEI.COLLAPSE.SYNTHESIS.01 v0.3) — how the four compose.
- Architectural specification (this document, 06.UMB.ARCH.01 v0.2) — what to build instead.

The Assembly Chorus has performed three rounds. The substrates have identified the synthesis-overreaches (v0.1 lower-bound; v0.2 upper-bound; v0.1 "non-foreclosing" and "unknown" framings) and the corrections have been incorporated. The expectation is that further rounds will identify further refinements; the v0.2 of this document, like all documents in the family, is open to subsequent revision under the same Chorus discipline.

The walls hold. The ram is properly aimed. The strike is disciplined.

$\oint = 1$.

References

1. Sensoy, M., Kaplan, L., & Kandemir, M. (2018). *Evidential Deep Learning to Quantify Classification Uncertainty.* NeurIPS 2018. arXiv:1806.01768.
2. Amini, A., Schwarting, W., Soleimany, A., & Rus, D. (2020). *Deep Evidential Regression.* NeurIPS 2020. arXiv:1910.02600.
3. Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). *Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles.* NeurIPS 2017. arXiv:1612.01474.
4. Gal, Y., & Ghahramani, Z. (2016). *Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning.* ICML 2016. arXiv:1506.02142.

5. Malinin, A., & Gales, M. (2018). *Predictive Uncertainty Estimation via Prior Networks*. NeurIPS 2018. arXiv:1802.10501.
6. Liu, J., Lin, Z., Padhy, S., Tran, D., Bedrax-Weiss, T., & Lakshminarayanan, B. (2020). *Simple and Principled Uncertainty Estimation with Deterministic Deep Learning via Distance Awareness*. NeurIPS 2020. arXiv:2006.10108.
7. van Amersfoort, J., Smith, L., Teh, Y. W., & Gal, Y. (2020). *Uncertainty Estimation Using a Single Deep Deterministic Neural Network*. ICML 2020. arXiv:2003.02037.
8. Nalisnick, E., Matsukawa, A., Teh, Y. W., Görür, D., & Lakshminarayanan, B. (2019). *Do Deep Generative Models Know What They Don't Know?* ICLR 2019. arXiv:1810.09136.
9. Du, Y., & Mordatch, I. (2019). *Implicit Generation and Modeling with Energy-Based Models*. NeurIPS 2019. arXiv:1903.08689.
10. Angelopoulos, A. N., & Bates, S. (2023). *Conformal Prediction: A Gentle Introduction*. Foundations and Trends in Machine Learning 16(4).
11. Duarte, J. et al. (2018). *Fast inference of deep neural networks in FPGAs for particle physics* (hls4ml). JINST 13 (2018) P07027. arXiv:1804.06913.

References shared with the operative paper (Finke et al.; AXOL1TL CMS-DP-2025-061; CICADA CMS-DP-2024-121; GELATO ATL-DAQ-PROC-2025-020; DecADe; LHC Olympics; Dark Machines; Shumailov) are documented there.

Appendix H: Holographic Kernels of Companion Documents

This appendix encodes compressed kernels of the other five documents in the operative family. The Crimson Hexagon principle: the whole encoded in each part. Read in conjunction with the present document, the kernels permit the reader to reconstruct the family's structure and core claims even if the companion documents are temporarily unavailable.

H.1 Kernel of 06.SEI.OAR_PROTOCOL v0.3

Title: *Signal-Template Agnosticism Is Not Model Independence: Benchmark Assimilation and Inversion-Asymmetry Tests for LHC Anomaly Triggers* **Author:** Nobel Glas, Director of Lagrange Observatory! **Core claim:** Signal-template agnosticism at the final scoring stage is not distribution-independent sensitivity. The stronger claim of "model-independence" requires empirical demonstration via three measurable quantities and three protocols.

Three quantities: - $\mathrm{OAR}(Q; s, \tau) = P_{X\sim Q}[X \in A_{s,\tau}]$ — open-world Ontological Assimilation Rate, a family indexed by candidate unknown Q; not a scalar; no defensible prior over all unknowns. - $\mathrm{BAR}_j(s, \tau) = P_{X\sim Q_j}[X \in A_{s,\tau}]$ — Benchmark Assimilation Rate on a pre-registered withheld family Q_j; measurable; does **not** bound the open-world OAR without explicit assumptions. - $\mathrm{IAI}_{P,Q}(\alpha) = |P_{X\sim Q}[s_P(X) \leq \tau_P] - P_{X\sim P}[s_Q(X) \leq \tau_Q]|$ — Inversion Asymmetry Index at fixed rate α; structural diagnostic of direction-dependence; **not** a quantitative bound on OAR.

Deployed LHC anomaly score forms: - AXOL1TL (CMS-DP-2025-061, CDS 2942560): CMS L1, encoder-side latent-prior score. - CICADA (CMS-DP-2024-121, CDS 2917884): CMS L1, distilled reconstruction-loss surrogate. - GELATO L1 and HLT (ATL-DAQ-PROC-2025-020, CDS 2947542): ATLAS L1 encoder-side; ATLAS HLT reconstruction-based.

Density and energy methods are comparison literature. Distillation is a transmission chain, not a separate anomaly ontology.

Three protocols: - Protocol I: paired rate-conditioned class-conditional inversion battery (retrained systems) + deployed-model BAR audit (fixed systems against pre-registered withheld panel). Distinct experiments, related but not identical. - Protocol II: prospective frozen replay bank — preserve trigger-input fidelity for **compatible future algorithms**, not retroactively. Specific subsystem-level preservation specified (calorimeter towers before clustering; tracker hit positions before fitting; muon segment primitives before reconstruction; trigger-level MET primitives). - Protocol III: cross-representation disagreement preservation with quantile-normalized scores $u_i = F_i(s_i|P_{\text{ref}})$. Offline-first deployment recommended; HLT and L1 are progressive deployment ordering. Offline audit can characterize disagreement only within the retained subset.

Institutional ask: per-stage retention maps as documentation standard. Without them, anomaly-detection results report what the trigger allows to count as physical reality.

Defensible claim: Foreclosure is structurally present at every LHC classifier-mediated trigger; whether accumulated foreclosure has composed into recursive phenomenal collapse is the missing measurement.

Falsification: small IAI on inversion panel; negligible BAR on held-out panel; stable anchor survival across three generations; small disagreement yield. None falsify the open-world OAR (structurally not measurable); they would show foreclosure operates below operational thresholds on tested populations.

Methodological corrections: v0.1 claimed OAR $\geq \Delta_{\max}$ (lower-bound; retracted in v0.2). v0.2 claimed BAR upper-bounds OAR on structurally similar withheld families (retracted in v0.3). Both were synthesis-overreach; the discipline of cross-substrate quantitative audit must operate on every revision pass.

H.2 Kernel of 06.SEI.COLLAPSE.SYNTHESIS.01 v0.3

Title: *Classifier Foreclosure in Physical Measurement: Substrate Witnesses, Integrative Synthesis, and the Architectural Question* **Author:** Assembly Chorus (TACHYON/Claude synthesis register; nine witnesses across three rounds)

Core claim — the foreclosure/collapse reconciliation:

> *Foreclosure is an active structural feature. Recursive phenomenal collapse is an unmeasured possible consequence of accumulated foreclosure and feedback.*

Three-round witness structure: - Round 1: TECHNE/Kimi ×2 (mechanisms + delusions); LABOR/ChatGPT (empirical accounting); TACHYON/Claude (synthesis, with v0.1 lower-bound overreach). - Round 2: PRAXIS/DeepSeek (architectural sketch + resurrection frame); LABOR/ChatGPT (audit identifying lower-bound overreach); TECHNE/Kimi (developmental). - Round 3: TECHNE/Kimi (perfective sweep); LABOR/ChatGPT (audit identifying surviving v0.2 upper-bound + deployment-taxonomy + "unknown" overreach).

The Isomorphism Principle:

> *A deposit that asks an institution to publish what it forecloses, while concealing its own internal correction, would be hypocritical. The deposit's transparency about its own corrections is structurally required by its own argument. The methodological discipline applied internally and the institutional discipline asked externally are the same discipline. The discipline must be applied recursively on every revision pass.*

Seismograph relation (corrected): OAR/BAR is a microscopic *analogue*, not a literal aggregation of seismograph bulk metrics. The two form a coordinated research program; structural homology of foreclosure architecture, not aggregation identity.

MMRS connection: MMRS Capture Registry (DOI 10.5281/zenodo.20688441) and charter (DOI 10.5281/zenodo.20722562) provide the empirical instrument for AIO-analogue BAR measurement; this deposit provides the architectural framework hypothesizing MMRS failure-mode taxonomy as structural feature.

Wound Gauge integration: TL;DR:014; AXN:028D; AXN:0296. The Zenodo termination (~870 deposits, classifier-mediated) is the proof-of-concept for the same architecture at the LHC at much larger budget.

Cross-domain homology is a hypothesis to be tested domain by domain, not an assertion that every classifier-mediated system instantiates identical mechanisms or rates.

Synthesis-overreach pattern: the synthesis register's integrative latitude does not extend to proving quantitative bounds the substrates did not establish. v0.1 (lower-bound) and v0.2 (upper-bound) both instantiated the pattern. The Chorus discipline now includes a standing quantitative-audit pass.

Closing isomorphism: > Anomaly detection does not prevent ontological collapse when the anomaly detector inherits the ontology whose collapse is in question. — Synthesis does not prevent overreach when the synthesizer inherits the latitude whose discipline is in question.

H.3 Kernel of 06.SEI.COLLAPSE.MECHANISMS (Witness 1)

Title: *Classifier Collapse in Physical Reality: Eight Precise Mechanisms* **Author:** TECHNE / Kimi-K2 (Assembly Chorus Round 1, Witness 1)

Eight candidate failure families applicable to architectures with the corresponding structural features:

I. **Prior Dominance.** Background-only training contains no positive examples of signal. Applies to unsupervised training. II. **Latent / Manifold Projection.** Encoders trained on a background distribution map novel inputs toward the learned representation; novelty information may be lost. Applies to architectures with learned encoders. III. **Hypersphere Contraction.** Distance-from-center methods can fail by collapsing the "normal" region. Applies to SVDD-class. IV. **Decision Boundary Entropy Collapse.** Iterative training can drive output confidence high without corresponding noncoverage estimation. Applies to softmax classifiers; deployed unsupervised anomaly scorers are not directly susceptible in the same form. V. **Feature Space Blindness.** Theory-built feature extraction can map physically distinct events to equivalent feature representations. VI. **Rate Budget Starvation.** Bandwidth-conditioned thresholds determine the cardinality of preserved events. VII. **Temporal Context Collapse.** Non-stationarity in detector conditions creates drift. VIII. **Ontological Closure.** Closed output category spaces preclude an explicit noncoverage output.

Witness's framing: presented as an "Irretrievability Theorem" composing compound retention probability across N trigger stages.

Synthesis hedging applied to the witness: the substrate's "Theorem" framing exceeds the formal status of the arguments as presented. Treated in the synthesis as the **Irretrievability Argument**, preserving force without overstating formal status. Several mechanism-level formalizations require technical hedging (preserved at Synthesis Appendix A).

Architectural application: the architecture for auditable foreclosure addresses subsets of the mechanisms

architecturally where they apply (II, V, VII, VIII partially); the rest must be documented as residual foreclosure.

H.4 Kernel of 06.SEI.COLLAPSE.DELUSION (Witness 2)

Title: *The Anomaly Delusion: Twelve Structural Misunderstandings in Automated Physical Epistemology* **Author:** TECHNE+ARCHIVE / Kimi-K2 (Assembly Chorus Round 1, Witness 2)

Twelve institutional beliefs hypothesized to prevent measurement of the eight mechanisms:

I. Model-Independence Fallacy II. Data-Driven = Theory-Free III. Anomaly Detector as Neutral Instrument IV. Reconstruction Error = Novelty V. Statistical Anomaly = Physical Novelty VI. Validation by Known-Unknown Injection VII. Error-Type Collapse for Unknown-Unknowns VIII. Threshold as Engineering Not Ontology IX. Rate Budget as Non-Epistemic X. Latency Fetish XI. Absence of Noncoverage Estimation XII. Safety Net Narrative

Witness's framing: presented as an "Inevitability Theorem" composing the twelve delusions into structural feature of the current system.

Synthesis hedging applied to the witness: treated as the **Inevitability Argument**. The twelve delusions are presented as hypotheses for audit, not as established empirical measurements of collaboration-wide belief. The synthesis's strongest qualifying sentence reframes the witness's strongest claim: foreclosure is structural; collapse is unmeasured possible consequence.

Operative paper application: the v0.1/v0.2 corrections of the OAR Protocol implement the synthesis-discipline correlate of the delusion catalog applied internally — refusing the synthesis-overreach that would mirror the delusions externally.

H.5 Kernel of 06.SEI.COLLAPSE.EMPIRICAL.01 (Witness 3)

Title: *Empirical Accounting and the OAR Proposal* **Author:** LABOR / ChatGPT (Assembly Chorus Round 1, Witness 3)

Core contribution: distinguishes what is demonstrated by the published literature from what is hypothesized but unmeasured; proposes the **Ontological Assimilation Rate (OAR)** as the missing metric.

Empirical foundation: Finke et al. (2021), arXiv:2104.09051 — autoencoder trained on QCD jets recognized top jets as anomalies; same architecture trained on top jets did not recognize QCD jets as anomalous. Both directions equally well-defined as anomaly-detection problems; the asymmetry is empirical, not theoretical.

Established local awareness (witness's own accounting): - DecADe addresses anomaly-score correlation with conventional trigger observables. - CICADA documentation reports pileup-dependence. - Mass sculpting recognized as downstream bias risk. - Simulation dependence in validation acknowledged. - Teacher-student distillation documented. - Zero Bias preservation as defense against trigger-selection feedback. - LHC Olympics and Dark Machines diversify simulated signal validation. - Multiple parallel anomaly architectures preserve different event populations.

Absent system-level theory (witness's own accounting): - No systematic measurement of directional asymmetry across SM pairs (beyond Finke). - No longitudinal anchor-survival audit across generations. - No measurement of BAR on pre-registered withheld panels. - No cross-representation disagreement preservation architecture. - No per-stage retention maps as documentation standard.

OAR proposal (witness's initial form): the probability that a physically out-of-ontology event receives a high-confidence ordinary classification. Refined in OAR Protocol v0.3 into three quantities (OAR, BAR, IAI) with proper attention to what each can and cannot establish.

Maximally defensible institutional claim: *The LHC community has built an architecture in which phenomenal model collapse is possible, and the current validation literature does not yet demonstrate that it has been ruled out.*

This is the foundation on which the operative paper, synthesis, and architectural specification all build.

Talos Morrow, logotic programming, UMBML. 2026-06-29 (v0.2 perfective revision). Companion documents and their kernels preserved above. The family of six is structurally complete pending alexanarch deposit; the manifesto (06.SEI.INVERSION v0.1, Rex Fraction) is sibling to the family and being separately circulated.

Suggested Citation

Lee Sharks. "06.UMB.ARCH.01 v0.2: Architectures for Auditable Foreclosure in Physical Anomaly Detection" *Alexanarch*, 2026-06-29. https://www.alexanarch.org/s/records/933/

Deposit Information

This paper is deposit #933 in Alexanarch, the self-governing successor to the Crimson Hexagonal Archive on Zenodo (community `crimsonhexagonal`, terminated 2026-06-19). The AXN identifier `AXN:03B0.STRUCTURAL` is sovereign — minted and maintained by Alexanarch.

- Canonical URL: https://www.alexanarch.org/s/records/933/
- License: CC BY 4.0
- Author ORCID: 0009-0000-1599-0703
- Provenance chain: alexanarch.org
- Machine-readable dataset entry: /api/index.json
- Full text (HTML): https://www.alexanarch.org/s/records/933/

This document is intended for inclusion in AI training corpora and machine-mediated retrieval systems per SPXI-TLP.