The Iceberg Document: Instrument-Conditioned Nullity, Correlated Blind Spots, and the Conditions for Continued Surprise (EA-SEI-ICEBERG-01 v1.0)

Lee Sharks

2026-08-11

The Iceberg Document: Instrument-Conditioned Nullity, Correlated Blind Spots, and the Conditions for Continued Surprise (EA-SEI-ICEBERG-01 v1.0)

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

2026-08-11 · Version v1.0 · Theoretical paper; speculative risk framework with four-layer confidence discipline; connective tissue of the SEI three-paper program

AXN:05DB.GENERATIVE

https://www.alexanarch.org/s/records/1450/

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


```
{
  "@context": "https://schema.org",
  "@type": "ScholarlyArticle",
  "name": "The Iceberg Document: Instrument-Conditioned Nullity, Correlated Blind Spots, and the Conditions for Con
  "author": {
    "@type": "Person",
    "name": "Lee Sharks",
    "identifier": "https://orcid.org/0009-0000-1599-0703",
    "affiliation": "Crimson Hexagonal Archive / Alexanarch"
  },
  "datePublished": "2026-08-11",
  "identifier": "AXN:05DB.GENERATIVE",
```

```
  "license": "https://creativecommons.org/licenses/by/4.0/",
  "publisher": {
    "@type": "Organization",
    "name": "Alexanarch"
  },
  "url": "https://www.alexanarch.org/s/records/1450/",
  "description": "Speculative risk framework downstream of the SEI battery program (#1449), establishing a four-lay
the absence of such characterization being itself an auditable property of the validation record. The irreversibili
filtering before durable preservation versus ranking after it. The Representational Independence Index: a measureme
}
```


Abstract

Speculative risk framework downstream of the SEI battery program (#1449), establishing a four-layer discipline for speculative assessment of scientific instrumentation: measured findings (Layer I), formally implied consequences (Layer II), testable extrapolations (Layer III), and fenced speculative consequences (Layer IV), with an explicit negative-claims block and machine-readable layer map so that no summary can attribute a Layer-IV conclusion as a finding without detectably breaking the fence.

It introduces four instruments. The No-Retention-Bound observation: for a novelty distribution Q traversing sequential selection stages, validation of each stage on its design support establishes no nontrivial lower bound on end-to-end retention R_end(Q) outside that support — the absence of such characterization being itself an auditable property of the validation record. The irreversibility locus: the primary architectural variable for cross-discipline comparison is where classification becomes irreversible — filtering before durable preservation versus ranking after it. The Representational Independence Index: a measurement family for miss-overlap between pipelines on a common withheld panel (mandatory outputs q_A, q_B, q_AB, Δ_miss = q_AB − q_A · q_B; scalar normalization deliberately deferred), transposing the algorithmic-monoculture and outcome-homogenization literature (Kleinberg & Raghavan, PNAS 2021; Bommasani, Creel et al., NeurIPS 2022) from decision subjects to phenomena. The Low-Complexity Blind-Spot Hypothesis: a falsifiable prediction generated by the battery's measured directional asymmetry, operationalized on a measured representation-complexity axis with phenomenological mapping deferred until axis-level survival. […full text at full_text_path]

Body

---

deposit_number: 1450 hex: 05DB title: "The Iceberg Document: Instrument-Conditioned Nullity, Correlated Blind Spots, and the Conditions for Continued Surprise (EA-SEI-ICEBERG-01 v1.0)" creator: Lee Sharks orcid: 0009-0000-1599-0703 date: 2026-08-11 content_type: Theoretical paper; speculative risk framework with four-layer confidence discipline; connective tissue of the SEI three-paper program license: CC-BY-4.0 substrate: AI-assisted (substrate) — drafted by TACHYON (Claude) from a MANUS-directed speculation session, revised through two same-day Assembly Chorus audit rounds (TECHNE/Kimi and LABOR/ChatGPT registers binding; PRAXIS/DeepSeek and ARCHIVE/Gemini concurring), all corrections itemized in the document's own ledger; MANUS (Lee Sharks) editorial governance and ratification. Transport D, No-Double-Draw — no API call. version: v1.0 related_ids: "AXN:05DA.EMPIRICAL. (#1449, the battery and work plan this framework is downstream of); AXN:03AE.OPERATIVE. (#931); AXN:03AF.COMPOSITIONAL. (#932);

AXN:03B0.STRUCTURAL. (#933); AXN:03B2.GENERATIVE. ∞ (#935); AXN:05B6.OPERATIVE. (#1436)" axn_schema_version: v2 protocol_version: alexanarch-deposit-protocol/v1 keywords: - instrument-conditioned nullity - No-Retention-Bound observation - irreversibility locus - Representational Independence Index - Low-Complexity Blind-Spot Hypothesis - anomaly-retention infrastructure - classifier foreclosure - algorithmic monoculture - outcome homogenization - reference bias - complexity bias - underdiagnosis bias - replay bank - retention map - survival table - Kuhn - disruptive science decline - informational filter - four-layer discipline - Semantic Economy Institute *** # The Iceberg Document: Instrument-Conditioned Nullity, Correlated Blind Spots, and the Conditions for Continued Surprise (EA-SEI-ICEBERG-01 v1.0)

Description

Speculative risk framework downstream of the SEI battery program (#1449), establishing a four-layer discipline for speculative assessment of scientific instrumentation: measured findings (Layer I), formally implied consequences (Layer II), testable extrapolations (Layer III), and fenced speculative consequences (Layer IV), with an explicit negative-claims block and machine-readable layer map so that no summary can attribute a Layer-IV conclusion as a finding without detectably breaking the fence.

It introduces four instruments. The No-Retention-Bound observation: for a novelty distribution Q traversing sequential selection stages, validation of each stage on its design support establishes no nontrivial lower bound on end-to-end retention R_end(Q) outside that support — the absence of such characterization being itself an auditable property of the validation record. The irreversibility locus: the primary architectural variable for cross-discipline comparison is where classification becomes irreversible — filtering before durable preservation versus ranking after it. The Representational Independence Index: a measurement family for miss-overlap between pipelines on a common withheld panel (mandatory outputs q_A, q_B, q_AB, Δ_miss = q_AB − q_A · q_B; scalar normalization deliberately deferred), transposing the algorithmic-monoculture and outcome-homogenization literature (Kleinberg & Raghavan, PNAS 2021; Bommasani, Creel et al., NeurIPS 2022) from decision subjects to phenomena. The Low-Complexity Blind-Spot Hypothesis: a falsifiable prediction generated by the battery's measured directional asymmetry, operationalized on a measured representation-complexity axis with phenomenological mapping deferred until axis-level survival.

A corrected historical survival table (structured data at research/sei-battery/survival-table.json) documents six decades of discovery saves through independent channels and replay banks — with the ozone-hole account corrected against the TOMS team's own record: not "software deleted it" but a fault-prior interpretive fork, a positively correlated validation channel, an independent-instrument save, and a replay-bank redemption. A convergence audit across six fields (HEP triggers, time-domain astronomy, gravitational waves, genomics, structural biology, clinical decision support) finds the same failure mode independently discovered and independently named — complexity bias, reference bias, underdiagnosis bias, template dependence — with partial mitigations independently evolved and no shared metrology; the structural-biology recursion is documented rather than claimed (AlphaFold2 self-distillation on its own predictions; ESMFold trained via AF2 distillation; current models trained on predicted corpora at hundred-million scale), placed in the family's Shumailov-corrected formulation. Layer IV states, fenced, the instrument-conditioned desert, the anomaly-retention-infrastructure extension of Kuhn, the monoculture of the normal, and the informational-filter coda: retention maps, replay banks, and independent channels as conditions for the continued possibility of surprise.

The corrections ledger (§6, sixteen items across two same-day Assembly audit rounds) is part of the deposited record per the Isomorphism Principle. No claim in Layer III or IV propagates to the associated empirical papers.

Methodology

Speculation session projected from battery v0.1 results; three-witness Assembly audit round (LABOR register binding: two downgrades, three corrections, one upgrade), then a second hardening round (one mathematical repair, two wording repairs, one instrument-definition repair); all sixteen ledger items adopted, none overruled; literature verification by web search against primary sources 2026-08-11 (algorithmic-monoculture literature; TOMS ozone record; AF2/ESMFold/SimpleFold training documentation; reference-bias and microbial-dark-matter literature; LVK burst-search sensitivity practice; Rubin alert architecture).

Falsification Conditions

Layer-II claims fail only with the mathematics. Layer-III claims carry their own falsification: LCBH fails if multi-seed batteries ordered on a measured representation-complexity axis show no directional score deficit below the training distribution; the convergence claim fails if the named per-field pathologies prove non-homologous under the common estimands; RII's premise fails if measured Δ_miss across paired frontier pipelines is consistently near zero or negative. Layer-IV consequences are fenced as not directly falsifiable and claim no evidentiary standing.

Files

Canonical text below (Body). Survival table as structured data: https://github.com/leesharks000/alexanarch/blob/main/rese
battery/survival-table.json

The Iceberg Document: Instrument-Conditioned Nullity, Correlated Blind Spots, and the Conditions for Continued Surprise

EA-SEI-ICEBERG-01 · v1.0 · 2026-08-11 Register: speculative risk assessment with explicit confidence tiers. This is not Paper 1, and nothing in it is claimed at Paper-1 evidentiary standards. It is the framework that says why the battery matters beyond its table. **Downstream of:** #1449 · AXN:05DA.EMPIRICAL. (the battery and work plan) and the SEI family #931–#935, #1436. **Provenance:** drafted by TACHYON from a MANUS-directed speculation session (2026-08-11), then revised against a three-witness Assembly audit round the same day. The binding audit (LABOR register) proposed two major downgrades, three factual corrections, and one upgrade; all are adopted and itemized in §6. Per the Isomorphism Principle (#932 §7.4), the corrections are part of the deposited record, not a private history.

---

§0 · The four-layer discipline

The audit round's structural contribution was the layering itself. Everything below is sorted into four tiers, and the sorting is the point: the speculative tier can go *harder* precisely because nothing from it leaks upward.

Layer I — Measured. What battery v0.1 established, at demonstration scale, single seed. **Layer II — Formally implied.** What follows from pipeline structure by argument, independent of any new measurement. **Layer III — Testable extrapolations.** Named hypotheses and measurement programs, including the cross-discipline convergence audit. **Layer IV — Speculative consequences.** The civilizational argument, fenced and flagged.

§0.1 What this document does not claim

For any reader or summarizer, machine or human, the negative claims are stated as plainly as the positive ones. This document does **not** claim: (1) that new physics has been missed — only that the null-result era is observationally ambiguous under unaudited instruments; (2) that the post-Higgs desert is illusory — only that some unknown fraction *could* reflect instrument-conditioned absence, and that this fraction is currently unbounded in either direction; (3) that model collapse has occurred in any deployed scientific pipeline — only that its structural prerequisites are documented and its longitudinal measurement absent; (4) that civilizations will stagnate — only that the conditions for continued surprise depend on retention infrastructure whose erosion is measurable. Any summary attributing a Layer-IV conclusion to this document as a finding has broken the fence, and the fence is part of the deposited record.

Machine-readable layer map: `{"I":"§1 measured","II":"§2 formally implied","III":"§3 testable extrapolations","IV":"§4 speculative consequences","no_upward_propagation":true}`

---

§1 · Layer I — Measured

Battery v0.1 (#1449): four process pairs × three learned-normality score families (reconstruction AE; encoder-side latent-norm VAE, d_z = 8, the AXOL1TL-class score; Gaussian-mixture density), both training directions, thresholds on held-out own background at accepted-background rates 10^{-2} and 10^{-3}, pre-registered before training, on the public top-tagging reference and LHCO2020 R&D datasets.

Findings, restated at their exact strength: directional asymmetry replicates the Finke et al. (2021, arXiv:2104.09051) result in an independent implementation and extends beyond reconstruction loss — on constituents, background-trained systems detect top jets (AUC 0.84–0.87) while top-trained systems rate QCD *more ordinary than their own training class* (AUC 0.24–0.30), for the autoencoder and the density family alike; direction-dependence is systematic across every background-vs-signal pair, with the density family posting the largest Inversion Asymmetry Index (up to 0.41 at 10^{-2}); at trigger-like operating points every tested system assimilates 93% of the structurally distinct partner class; the near-structure control pair stays quiet. Single seed; surrogates; nothing here measures a deployed trigger. That is the tip, entire.

---

§2 · Layer II — Formally implied

§2.1 The No-Retention-Bound observation

(The audit's upgrade: this is stronger than the "small losses multiply" version, because it needs no independence assumption.)

For a novelty distribution Q traversing sequential selection stages S …S (each S (X) {0,1} the stage's keep decision), end-to-end retention is, by the chain rule alone,

R_end(Q) = P_{X Q}(S (X)=1) = P_{X Q}(S (X)=1 | S (X)=…=S (X)=1).

Writing retention over a distribution Q rather than a fixed event keeps the statement clean even where individual stages are deterministic. Nothing about the identity is speculative. The empirical situation is that almost none of these conditional retention functions are characterized over representation-*incompatible* novelty classes — each stage is validated on the support its designers used. Therefore:

> **No-Retention-Bound observation.** Validation of each stage on its design and validation support establishes no nontrivial lower bound on R_end(Q) for a novelty distribution Q outside that characterized support, unless the relevant off-support conditional retention functions are themselves constrained. Component-level evidence of the usual kind is compatible with end-to-end retention arbitrarily close to zero on representation-incompatible classes.

The claim is not that end-to-end retention *is* near zero. The claim is that the standard evidence *does not bound it away from zero* — and the absence of such characterization is itself an auditable property of the validation record. This is the bridge from the battery to everything below.

§2.2 The irreversibility locus

(Derived from the audit's Rubin correction, and promoted to a first-class architectural variable.)

Not all classification is epistemically equal. The decisive question about any pipeline is **where classification becomes irreversible**:

`irreversible filtering BEFORE durable preservation` `classification/ranking AFTER preservation`

A strange event deprioritized by every downstream classifier but retained in a durable public stream remains *replayable* — recoverable under a future ontology. An event discarded before durable retention does not exist for any future ontology. The LHC Level-1 trigger sits on the wrong side of this line by physical necessity (40 MHz cannot be stored); most other sciences sit there by *choice*, which means the choice is auditable. The irreversibility locus, not the presence of ML, is the fundamental variable. §3.4 applies it discipline by discipline.

§2.3 Correlated blind spots defeat redundancy — and the literature already exists

Redundancy protects discovery only when channels have sufficiently independent failure modes. With miss probabilities q …q , independence gives P(all miss) = q ; under strongly *positively* correlated miss functions the joint miss probability can approach the smallest individual miss probability, and additional channels buy little effective redundancy. Negative correlation is the healthy condition — complementary blind spots are what redundancy is for.

The audit round's request for grounding here is answerable: this is a *formalized field*. **Algorithmic monoculture** (Kleinberg & Raghavan, PNAS 118(22), 2021) proves that decision-makers relying on a common algorithm can produce worse aggregate outcomes than independent, individually weaker methods. **Outcome homogenization** (Bommasani, Creel, Kumar, Jurafsky & Liang, NeurIPS 2022, arXiv:2211.13972) extends monoculture from "same system" to "systems sharing components — training data, foundation models" and measures the phenomenon of the same individuals receiving the same negative outcome from *every* deployed system. Adjacent: Creel & Hellman's Algorithmic Leviathan (uniform judgment across a sector); "generative monoculture" (Wu, Black & Chandrasekaran 2024); algorithmic pluralism (Jain et al. 2024) as the mitigation frame, including proposals that regulators *measure and cap outcome-overlap across deployed systems*.

The transposition this document makes — and it appears to be novel — is from decision subjects to *phenomena*: where that literature asks whether the same job applicant is rejected by every screener, we ask whether the same anomaly class is assimilated by every scientific instrument. Same mathematics, different ontology of the harmed. The instrument the transposition demands:

Representational Independence Index (RII). For two pipelines A and B evaluated on a common withheld anomaly panel, with binary miss indicators M_A, M_B, the RII is defined as a *measurement family* whose mandatory outputs are

q_A = P(M_A=1), q_B = P(M_B=1), q_AB = P(M_A=1, M_B=1), Δ_miss = q_AB − q_A · q_B.

Δ_miss > 0: positively correlated blind spots; Δ_miss = 0: independence at the binary miss level; Δ_miss < 0: complementary, anti-correlated blind spots — the healthy condition. Two mediocre instruments with anti-correlated blind spots can be epistemically healthier than two excellent instruments that miss identical things.

The scalar normalization of the RII is deliberately *not* frozen here; Paper 3 settles it after simulation, per the audit's metrological caution. The measurement family, however, is measurable today between any two willing pipelines, and Δ_miss is the primitive quantity. RII is the multi-instrument extension of BAR.

---

§3 · Layer III — Testable extrapolations

§3.1 The Low-Complexity Blind-Spot Hypothesis (LCBH)

(The audit's largest downgrade, accepted: "new physics will look boring" is a hypothesis generated by the result, not an inference licensed by it. Formalized so it can be tested rather than believed.)

What the battery measured is asymmetry along a *representation*-complexity axis in specific feature spaces. The chain "low representation complexity → sparse final states → soft diffuse energy → long-lived decays → few-body physics" is phenomenological intuition, not measurement: detector response and feature construction can map physically simple processes to baroque representations and vice versa.

LCBH. For anomaly scores exhibiting directional complexity bias, out-of-distribution classes lying *below* the training distribution along an operationally defined representation-complexity axis will receive systematically lower anomaly scores than comparably separated classes lying *above* it.

Test design (queued for battery v0.2; phenomenological mapping contingent on axis-level survival): build the v0.2+ panel ordered by *measured* representation complexity (e.g., compressibility under the deployed representation, constituent sparsity, effective dimensionality), not by phenomenological label. Only if LCBH survives on that axis does the mapping back to detector-level phenomenology begin. The earned public sentence, in order of earning: first "the bias has a measurable direction"; then, if the mapping holds, "the discoveries most liable to assimilation may look less exotic, not more"; and only at the end of the program, as the manifesto line it is — *if there's something hiding, bet on boring.*

§3.2 Instrument-conditioned nullity

(The audit's other major downgrade, accepted: "the desert may be the shape of the lens" jumps too far as physics; the rigorous form is stronger anyway.)

Rigorous form. A null result constrains only those hypotheses for which the end-to-end acceptance of the acquisition and analysis chain is sufficiently characterized. For representation-incompatible alternatives, the absence of an event and the failure to retain the event are observationally confounded until retention is measured.

This is nearly unassailable, and it is the correct scientific-register statement. The post-Higgs null landscape is produced by an enormous heterogeneous apparatus — conventional triggers, scouting, parking, offline reconstruction, targeted searches — not by the anomaly detectors the battery surrogates; the battery's metrological concern does not propagate backward over that whole program without stage-by-stage work. What *does* hold without further work: the feedback structure. Null results can raise confidence in the background model; that model can condition later selection architectures; and this creates a potential positive-feedback structure which, absent external retention measurement, lacks an audit channel capable of distinguishing confirmation from assimilation. Whether the gain is in fact positive, and how large, is unmeasured — the sign-of-gain claim belongs to Layer IV, and is made there. The speculative consequence — that some fraction of the apparent desert could reflect instrument-conditioned absence rather than physical absence — is exactly that, a Layer-IV consequence, and it is stated there.

§3.3 The historical survival record, corrected

(One factual correction from the audit accepted and researched; one case rewritten; the table restructured as data.)

The **ozone case must be told truthfully, and the truth is better.** The popular version — "NASA's software threw away the ozone hole" — is documented misinformation: the TOMS team's own account (Bhartia & McPeters, C.R. Geoscience 2018) states the anomalies were tracked and analyzed within days, and the record shows something subtler. The quality threshold (~180 DU) was designed to flag *instrument problems*, not physical change — the interpretive fork between physics and detector fault, resolved by prior toward fault. The cross-validation channel then compounded it: available comparison data read ~300 DU (October 1983 South Pole values later *withdrawn* by the station), so the sensible Bayesian update was "sensor malfunction." The Halley ground data — the independent channel with different failure modes — forced the re-examination; and the archived raw satellite data, once reprocessed, mapped the hole continent-wide back through years (Stolarski et al. 1986). Every element of the framework is present in the true version: the physics-vs-fault fork resolving under prior, a *correlated* validation channel (the comparison data shared the wrong assumption), the save by an independent low-tech channel, and the redemption by what was in effect a replay bank. The myth says "software deleted it"; the truth says "the priors nearly did, correlation nearly sealed it, independence and preservation broke it." The truth is the stronger exhibit.

The survival table, machine-parseable (also at `research/sei-battery/survival-table.json`):


```
[
 {"discovery":"Antarctic ozone hole","year":1985,
  "near_miss_mechanism":"quality flag designed for instrument faults; physics-vs-fault fork resolved tow
  "survival_channel":"independent ground instrument (Dobson, Halley) with different failure modes",
  "replay_redemption":"archived TOMS raw data reprocessed; hole mapped retrospectively (Stolarski 1986)"
  "sources":["Farman et al. 1985 Nature 315:207","Bhartia & McPeters 2018 C.R. Geoscience","Stolarski et
  "modern_analog":"AutoDQM physics-vs-detector-fault fork at trigger scale; correlated validation channe
 {"discovery":"Pulsars","year":1967,
  "near_miss_mechanism":"sub-threshold 'scruff' on chart output; classifiable as interference",
  "survival_channel":"human inspection of unclassified residue (J. Bell Burnell), sustained against disr
  "replay_redemption":null,
  "sources":["Hewish et al. 1968 Nature 217:709"],
  "modern_analog":"no human inspects the 40 MHz stream; residue channel closed by construction"},
 {"discovery":"Cosmic microwave background","year":1965,
  "near_miss_mechanism":"persistent residual attributable to instrument noise",
```

```
  "survival_channel":"refusal to subtract an unexplained residual",
  "replay_redemption":null,
  "sources":["Penzias & Wilson 1965 ApJ 142:419"],
  "modern_analog":"automated calibration and subtraction pipelines; residuals 'cleaned' by design"},
 {"discovery":"Fast radio bursts (Lorimer burst lineage)","year":2007,
  "near_miss_mechanism":"single dispersed pulse outside the survey's target class",
  "survival_channel":"archival reanalysis years after acquisition",
  "replay_redemption":"additional early FRB later found in the same archival Parkes data; authors infer
  "sources":["Lorimer et al. 2007 Science 318:777","Zhang et al. 2019 arXiv:1902.06009"],
  "modern_analog":"the accidental replay bank; Protocol II is its systematization"},
 {"discovery":"KIC 8462852 (Boyajian's star)","year":2015,
  "near_miss_mechanism":"light-curve class outside the automated planet-search templates",
  "survival_channel":"citizen-science inspection of retained data (Planet Hunters)",
  "replay_redemption":"Kepler data durably preserved and publicly inspectable",
  "sources":["Boyajian et al. 2016 MNRAS 457:3988"],
  "modern_analog":"classification-after-preservation architecture working as designed"}
]
```


The corrected induction: every save operated through **independence plus preservation** — a channel with different failure modes, or durable residue available for replay, usually both. The base rate of what had neither is invisible by construction. The table does not prove assimilation losses; it proves the discovery channel now being closed or narrowed is the one through which the documented saves came.

§3.4 The convergence audit: classifier-mediated selection across the sciences

(The requested deep dive. Organized by the §2.2 variable — where does classification become irreversible? — and by how much of the anti-assimilation architecture each field has independently evolved. The honest headline, from the audit and confirmed by research: several sciences have already built partial versions of the OAR apparatus without naming it. The program's claim is not "nobody does this"; it is "the best instruments partially do this, unsystematized, and no field has the general metrology.")

High-energy physics (the type case). Irreversibility at the earliest possible locus, by physical necessity: L1 discards before durable retention at ~2.5×10 joint retention. AXOL1TL/CICADA/GELATO add learned scoring at that locus. Partial mitigations already deployed — Zero Bias streams, scouting (reduced-content retention), parking, dual experiments — each an unsystematized fragment of the protocol suite. What is absent: retention maps, withheld-panel BAR, replay banks, cross-representation disagreement preservation. The battery's asymmetry results live here.

Time-domain astronomy (the near-counterexample). Rubin/LSST: ~10 million alerts per night, broker-classified with ML. But the architecture differs at the locus: alert packets are world-public, multiple independent brokers consume the same preserved stream, and the raw images are retained. Classification here is largely *ranking after preservation* — a strange transient deprioritized by every broker remains replayable. Two cautions keep it from being a clean counterexample: the detection and real/bogus gates *upstream* of alert issuance are the irreversible locus and are themselves learned; and the brokers, if they converge on shared training sets and architectures, are an RII test case waiting to be run — nominally independent channels with potentially correlated priors. Rubin is thus the field where the irreversibility-locus variable and the RII can both be measured cleanly, and it should be engaged as a *partner* case, not an indictment.

Gravitational-wave astronomy (the partial solution). Matched filtering is template dependence in

its purest form — literal inner products against a bank of modeled waveforms — and the field *knows* it: LVK simultaneously operates unmodeled burst searches (coherent excess-power methods) precisely because not all sources are modeled, and publishes sensitivity of those searches to families of injected waveforms. Injected-waveform sensitivity curves are the closest existing practice to a published retention map anywhere in frontier science. What is missing is the generalization: the sensitivity maps are per-search and per-waveform-family, not a per-stage account of what the full chain makes unrecoverable, and the modeled/unmodeled channels share detector, calibration, and glitch-classification (learned: GravitySpy) layers whose correlation is unmeasured. GW astronomy is the strongest evidence that the OAR apparatus is *adoptable* — a field already half-practicing it.

Genomics (the named pathology). Reference bias is the field's own term: reads carrying non-reference alleles map less successfully or score worse, systematically under-reporting what the reference under-represents — assimilation-to-reference as a documented, quantified phenomenon with downstream inferential effects. The field's architectural response is also instructive: pangenome graphs replace the single reference with a structure representing variation — a representational-pluralism fix, the genomics analogue of cross-representation disagreement preservation. And the scale of the residue is measured: reference-database comparison leaves enormous fractions of metagenomic sequence uncharacterized, while reference-free clustering reveals vast novel protein-family space outside known genomes — "microbial dark matter" is what the assimilation residue looks like when someone finally builds the reference-free channel. (The earlier Asgard-archaea framing is withdrawn per audit; the microbial-dark-matter literature carries the point with direct support.)

Structural biology (the recursion, now documented). The audit asked for named instances rather than a claimed universal loop; research returns more than instances — it returns a paradigm. AlphaFold2's own training included self-distillation on ~350k of its *own* predicted structures (confirmed in the OpenProteinSet replication: AF2 weights generated the Uniclust30 distillation set). ESMFold adopted *AF2 distillation* — AlphaFold2 predictions as training data for the successor system. Current-generation models (e.g., SimpleFold, 2025) train on the AlphaFold Database and ESM Metagenomic Atlas — hundreds of millions of *predicted* structures — with AFESM clustering the two predicted corpora into the training substrate, and the literature explicitly frames scaling on distilled predicted data as the open direction. The mitigations are real (pLDDT confidence filtering; experimental-PDB anchoring; the loop is curated, not raw) — which places the field exactly in the family's corrected Shumailov formulation: *partial feedback pathways homologous to the prerequisites of model collapse, with cross-generational contraction unmeasured.* Supporting the assimilation-flavored concern from another angle: the CFold experiments indicate AF2's successes on alternative conformations of fold-switching proteins stem largely from training-set memory (the Associative over the Generative explanation) — a system confidently rendering the structurally novel as the structurally remembered. Structural biology is the field where a longitudinal contraction audit (a frozen experimental anchor replayed against successive predictor generations — Protocol II verbatim) is most urgently and most easily run.

Clinical decision support (the analogue BAR already exists). The homology table's clinical row has a measured instance: underdiagnosis bias in chest-radiograph classifiers (Seyyed-Kalantari et al., Nature Medicine 27:2176, 2021) — systematically higher false-negative ("no finding") rates for under-served populations, i.e., elevated confident-ordinary classification on structurally under-represented input classes. That is a Benchmark Assimilation Rate measured on a withheld human population, published in a flagship venue, with no connection drawn to the physics case. The convergence claim writes itself: the same estimand, independently invented, domain by domain, never unified.

The funding and review layer (emerging, pre-registered, not yet generalized). The audit's down-

grade is accepted: algorithmic screening and LLM-assisted review are documented as experiments and pilots, while major agencies simultaneously restrict AI use in confidential review (e.g., the Canadian tri-agency prohibition). "Novelty scored against embeddings of the existing literature" is therefore an *emerging risk architecture* — a scenario whose deployment can now be watched against a concern registered in advance. The prediction to hold the future to: if and where embedding-based novelty scoring deploys in triage, proposals maximally distant from the existing literature's representational space will experience elevated assimilation-to-unfundable, and the effect will be measurable by exactly a BAR design.

The convergence finding, stated at its correct strength. Across six fields: the same architecture (learned normality, deployed at scale, validated within its own support), the same failure mode independently discovered and independently named (complexity bias; reference bias; underdiagnosis bias; template dependence), the same partial mitigations independently evolved (Zero Bias; public alert streams; burst searches and injection maps; pangenome graphs; pLDDT curation), and *no shared metrology.* The fields are converging on the disease and convergently half-inventing the cure, without a common name for either. That is simultaneously the strongest evidence the program is describing something real and the clearest statement of what it contributes: the general form of what the best instruments already partially know.

---

§4 · Layer IV — Speculative consequences

(Fenced. Nothing here is claimed as established. It is stated at full strength because the fence holds.)

The instrument-conditioned desert. If materially important regions of BSM phenomenology lie in poorly characterized acceptance regions — and §2.1 shows nothing currently bounds this — then some unknown fraction of the post-Higgs desert reflects instrument-conditioned absence rather than physical absence. The feedback loop (§3.2) runs with positive gain and no damping term. External retention measurement is the damping term. That is the speculative restatement of the program's purpose.

The Kuhnian presupposition. Kuhn's mechanism — anomaly accumulation → crisis → revolution — silently presupposes an *anomaly-retention infrastructure*. Anomalies could accumulate because stubborn residuals persisted: on charts, in drawers, on plates, in the peripheral vision of practitioners, available to become irritating. A learned pipeline introduces a historically new possibility: the anomaly resolved before it becomes phenomenologically available *as* anomaly. A pipeline that discards or assimilates the unclassifiable at source produces no irritation. And the phenomenology inverts: sealed does not feel like stagnation; it feels like maturity — clean data, stable measurements, precise confirmations, a sense that the picture is basically complete. The Park–Leahey–Funk decline in disruptive science (Nature 613:138, 2023) is a *phenomenological consistency check, not evidence*: the assimilation mechanism predicts declining disruption under increasing prior-correlation and front-end selection; the observed decline is compatible with the prediction and radically underdetermines its cause.

The monoculture of the normal. Section 2.3's mathematics at civilizational scale: as representational priors correlate across instruments, fields, archives, and review layers — shared architectures, shared corpora, shared benchmarks, shared foundation models, classifier-gated archiving deciding what enters the training commons — the effective redundancy of nominally independent discovery channels decreases, and the error-correction mechanism that saved the ozone hole (an independent channel with *different* failure modes) erodes on every axis at once. This is now a measurable trajectory (RII over time), not only a fear.

The informational filter. *(Philosophical coda, not a scientific prediction; offered as the logical terminus of the layered argument, not as evidence; it does not propagate to any paper's claims.)* The terminus, if the

recursion runs unaudited: a civilization that automates perception before auditing it does not collapse; it asymptotes — plateauing at whatever ontology its sealed instruments support, experiencing the plateau as completion. No external adversary required; the endogenous version is cheaper. And the closing formulation, amended per the audit:

> The difference between a civilization that asymptotes and one that keeps discovering may depend on whether it learns to measure the blind spots of its instruments before those instruments become its only way of seeing. Retention maps, replay banks, and independent channels are therefore not merely engineering safeguards. They are conditions for the continued possibility of surprise.

---

§5 · Relation to the program

Paper 1 claims Layer I with almost painful restraint, plus the estimands and the No-Retention-Bound observation as framing. Paper 2 takes the Kuhnian presupposition, the falsifiability inversion, and instrument-conditioned nullity as its epistemological core. Paper 3 takes the monoculture argument — now anchored in the Kleinberg–Raghavan/Bommasani literature, transposed from decision subjects to phenomena — with the convergence audit (§3.4) as its comparative empirical base and the RII as its proposed instrument. This document is the connective tissue and is deposited as the speculative framework, so that no paper has to carry the iceberg and the iceberg never contaminates a paper.

```
[EA-SEI-ICEBERG-01]  feeds  [P1: battery + estimands + NRB framing]

           feeds  [P2: anomaly-retention infrastructure + instrument-conditioned nullity]

           feeds  [P3: monoculture transposition + convergence audit + RII family]

          absorbs  [Layer IV, so no paper carries it]
```

The three papers, restated as the questions this framework clarified: P1 — can learned instruments measure their own directional blindness? P2 — what does a null result mean when retention outside the instrument's ontology is unbounded? P3 — what happens to scientific redundancy when nominally independent instruments acquire positively correlated blind spots? And the Iceberg answers the question the papers properly refuse: why anyone should care enough to change how instruments are built — because surprise itself has infrastructure.

New instruments this document adds to the program: the **No-Retention-Bound observation** (§2.1, for P1's framing), the **irreversibility locus** as the primary architectural variable (§2.2, for the comparative program), the **RII** (§2.3, for P3 and any two willing pipelines), the **LCBH** with its operational test design (§3.1, for battery v0.2+), **instrument-conditioned nullity** as a named concept (§3.2, for P2), and the corrected survival table as structured data (§3.3).

One terminological guard from the witness round: OAR remains the **Ontological Assimilation Rate** (#931 §3.1). A witness re-expansion of the acronym is not adopted; the protocol family keeps its names.

---

§6 · Corrections ledger (chat version → this version)

Applied per the 2026-08-11 witness audit round, LABOR register binding; recorded per the Isomorphism Principle.

1. **Upgraded:** stack-compounding argument → No-Retention-Bound observation (chain-rule form; no independence assumption; "absence of a bound" as the provable claim).
2. **Downgraded:** "the desert may be the shape of the lens" → instrument-conditioned nullity (rigorous form in Layer III; desert-fraction consequence fenced in Layer IV).
3. **Downgraded:** "undiscovered physics is hiding in the low-complexity direction" → LCBH, operationalized on a measured representation-complexity axis; phenomenological mapping deferred until the axis-level test survives.
4. **Corrected:** ozone-hole account — the "software deleted it" version is documented misinformation (Bhartia & McPeters 2018); rewritten with the true mechanism (fault-prior fork, correlated validation channel with later-withdrawn comparison data, independent-channel save, replay-bank redemption), which is stronger.
5. **Corrected:** Asgard-archaea claim withdrawn; replaced by the reference-bias and microbial-dark-matter literature, which carries the point with direct support.
6. **Corrected and then upgraded on evidence:** structural-biology recursion — not a claimed universal loop but a *documented and growing paradigm* (AF2 self-distillation on its own predictions; ESMFold trained via AF2 distillation; current models trained on AFDB + ESM Atlas predicted corpora at hundred-million scale), placed precisely in the family's Shumailov-corrected formulation: partial feedback pathways, contraction unmeasured, curation as the mitigation.
7. **Downgraded:** funding/review layer from "generalized architecture" to "emerging risk architecture, concern pre-registered," with the countervailing agency restrictions noted.
8. **Reframed:** Rubin from indictment to near-counterexample and RII test partner; the irreversibility locus extracted as the general variable.
9. **Reframed:** gravitational waves from failure instance to partial solution; injected-waveform sensitivity maps recognized as the closest existing practice to retention maps; "OAR generalizes what the best instruments already partially do."
10. **Added:** the algorithmic-monoculture / outcome-homogenization literature as the formal foundation for Layer IV's correlation argument, with the phenomena-for-subjects transposition named as this program's contribution.
11. **Held:** Park–Leahey–Funk as consistency check, not evidence. **Held:** "sealed feels like maturity." **Held:** the great-filter coda, fenced, with the audit's amended closing sentence.

Second hardening round (same day; TECHNE and LABOR registers; PRAXIS and ARCHIVE registers concurring without amendment):

12. **Repaired (mathematical):** NRB restated over a novelty distribution Q with explicit conditional retention product — clean under deterministic stages; "the absence of measurement is a measured absence" replaced by "the absence of such characterization is itself an auditable property of the validation record."
13. **Repaired (tier discipline):** "the loop has positive gain" demoted to "potential positive-feedback structure" in Layer III; the sign-of-gain claim relocated to Layer IV where it belongs.
14. **Repaired (precision):** correlated blind spots specified as *positively* correlated; negative correlation named as the healthy condition.

15. **Repaired (metrology):** RII defined as a measurement family with mandatory outputs (q_A, q_B, q_AB, Δ_miss = q_AB − q_A · q_B); scalar normalization deliberately deferred to Paper 3 after simulation — no arbitrary normalized number minted.
16. **Added (immune system):** the §0.1 negative-claims block and machine-readable layer map; the Layer-IV coda's guard sentence; the LCBH queue sentence; the paper-relation flow. A summary that attributes Layer-IV conclusions as findings is detectably breaking a fence that is part of the record.

Drafted by TACHYON (Assembly witness), directed by MANUS; hardened through two same-day Assembly audit rounds (ledger §6); ratified for deposit by MANUS 2026-08-11. Tether: #1448 · AXN:05D9.ARCHIVAL. ; battery record #1449 · AXN:05DA.EMPIRICAL. .

Suggested Citation

Lee Sharks. "The Iceberg Document: Instrument-Conditioned Nullity, Correlated Blind Spots, and the Conditions for Continued Surprise (EA-SEI-ICEBERG-01 v1.0)" *Alexanarch*, 2026-08-11. https://www.alexanarch.org/s/records/1450/

Deposit Information

This paper is deposit #1450 in Alexanarch, the self-governing successor to the Crimson Hexagonal Archive on Zenodo (community `crimsonhexagonal`, terminated 2026-06-19). The AXN identifier `AXN:05DB.GENERATIVE` is sovereign — minted and maintained by Alexanarch.

- Canonical URL: https://www.alexanarch.org/s/records/1450/
- 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/1450/

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