Baseline Capture Architecture for Learned Scientific Triggers: A Control Plane for Measuring Selection Before Irreversible Data Loss (EA-SEI-BCA-01 v2.0 — the submitted state)

Sharks, Lee

2026-08-27

Baseline Capture Architecture for Learned Scientific Triggers: A Control Plane for Measuring Selection Before Irreversible Data Loss (EA-SEI-BCA-01 v2.0 — the submitted state)

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

2026-08-27 · Version v1.0 · Methodological specification; technical architecture and selection-metrology paper — submitted manuscript state

AXN:0656.OPERATIVE

https://www.alexanarch.org/s/records/1558/

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


```
{
  "@context": "https://schema.org",
  "@type": "ScholarlyArticle",
 "name": "Baseline Capture Architecture for Learned Scientific Triggers: A Control Plane for Measuring Selecti
SEI-BCA-01 v2.0 — the submitted state)",
  "author": {
    "@type": "Person",
    "name": "Sharks, Lee",
    "identifier": "https://orcid.org/0009-0000-1599-0703",
```

```
    "affiliation": "Crimson Hexagonal Archive / Alexanarch"
  },
  "datePublished": "2026-08-27",
  "identifier": "AXN:0656.OPERATIVE",
  "license": "https://creativecommons.org/licenses/by/4.0/",
  "publisher": {
    "@type": "Organization",
    "name": "Alexanarch"
  },
  "url": "https://www.alexanarch.org/s/records/1558/",
 "description": "The submitted state of EA-SEI-BCA-01, deposited so that the version named in a journal's data
availability statement is the version a reader arrives at. This supersedes deposit #1453 (AXN:05DE, Nobel Glas,
08-11, v1.0), which carried the architecture, the estimation theory, and the accelerator feasibility argument
4 reconstruction-error trigger at a background retention of 1e-3, with four injected anomaly classes including
only archive audits itself as perfect, returning a retention of 1.000 for every class because the denominator p
sensitive pseudo-baseline is not merely imprecise but confidently wrong, estimating a shifted class at 0.76 wit
independent channel measures every class without selection bias, including a retention of exactly zero for the
blind class, with an interval — an estimate no content-sensitive channel in the experiment can produce at any bu
size relation, one observation against an expectation of 0.16, so the baseline states the limits of its own comp
}
```


Abstract

The submitted state of EA-SEI-BCA-01, deposited so that the version named in a journal's data-availability statement is the version a reader arrives at. This supersedes deposit #1453 (AXN:05DE, Nobel Glas, 2026-08-11, v1.0), which carried the architecture, the estimation theory, and the accelerator feasibility argument but no worked demonstration. What is new here is §13, a numerical validation against known ground truth, and the venue reframing that accompanied it. A synthetic stream of four million events is selected by a rank-4 reconstruction-error trigger at a background retention of 1e-3, with four injected anomaly classes including one constructed to lie exactly on the learned manifold, and four acquisition regimes compared at matched storage budgets. The results are the paper's claims made empirical: the trigger-only archive audits itself as perfect, returning a retention of 1.000 for every class because the denominator population no longer exists; a content-sensitive pseudo-baseline is not merely imprecise but confidently wrong, estimating a shifted class at 0.76 with a tight interval against a true 0.016, because the second selector misses what the first misses; the content-independent channel measures every class without selection bias, including a retention of exactly zero for the representation-blind class, with an interval — an estimate no content-sensitive channel in the experiment can produce at any budget; and the rare class exhibits the paper's own sample-size relation, one observation against an expectation of 0.16, so the baseline states the limits of its own competence where the trigger archive cannot. The shadow plane returns miss correlations of 0.55 to 0.75 across visible classes and an undefined correlation on the blind class, where both selectors miss at rate 1.0 and no variation remains to correlate. A preliminary enrichment design keyed to the selector's own scores is reported as a failure rather than omitted: it starved the miss region, and its correction is the architectural rule that the uniform floor is what every other channel stands on. [...abridged for the catalogue; full description in this deposit's record]

Body

---

deposit_number: 1558 hex: 0656 title: “Baseline Capture Architecture for Learned Scientific Triggers: A Control Plane for Measuring Selection Before Irreversible Data Loss (EA-SEI-BCA-01 v2.0 — the submitted state)” creator: Sharks, Lee orcid: 0009-0000-1599-0703 date: 2026-08-27 content_type: Methodological specification; technical architecture and selection-metrology paper — submitted manuscript state license: CC-BY-4.0 substrate: “AI-assisted (substrate) — drafted through the Assembly under MANUS (Lee Sharks) editorial governance. The v1.0 architecture was deposited under the Nobel Glas heteronym, Director of Lagrange Observatory, whose function is the Measurement of Meaning; this state carries the byline under which it was submitted. The numerical validation of §13 was designed, executed and diagnosed in working dialogue with Claude (Anthropic): the experiment is a single seeded script whose figures and table are re-derivable from it, and the mid-course estimator defect it uncovered is reported in the paper rather than silently corrected. Transport D, No-Double-Draw.” axn_schema_version: v2 protocol_version: alexanarch-deposit-protocol/v1 keywords: - baseline capture architecture - learned scientific triggers - selection metrology - content-independent probability sampling - replay bank - shadow deployment - retention estimation - Hajek estimator - inverse-probability weighting - representation-blind class - correlated misses - numerical validation - measurement uncertainty *** # Baseline Capture Architecture for Learned Scientific Triggers: A Control Plane for Measuring Selection Before Irreversible Data Loss (EA-SEI-BCA-01 v2.0 — the submitted state)

Scientific instruments increasingly place learned classifiers, anomaly scores, and other data-dependent mechanisms before durable storage. Where the incoming rate exceeds any feasible archival rate, such selection is unavoidable. The problem is not selection itself but its auditability: a learned selector is characterized only over anticipated classes, while the events of greatest discovery interest are those for which it is weakest. The data needed to measure the selector’s errors may itself be destroyed by the selector.

This paper proposes a **Baseline Capture Architecture (BCA)**: a statistically interpretable control plane for instruments performing irreversible learned selection. It separates four commonly conflated functions — a content-independent probability sample taken before the audited selection (C0); a full-rate tap of the representation presented to the selector (C1); probability-weighted enrichment channels whose inclusion laws are recorded (C2); and shadow selectors whose decisions are logged but do not control acquisition (C3) — feeding a replay bank from which retention surfaces, miss correlation, and selection drift can be estimated retrospectively.

The central principle is deliberately minimal. Under a finite storage budget, no content-sensitive algorithm can constitute a less assumption-laden baseline than a probability sample whose inclusion mechanism is independent of event content and whose inclusion probability is known. More elaborate models may improve discovery efficiency, but they belong to the experimental arm, not the control arm. BCA does not propose an alternative trigger; it proposes the missing **control group for a trigger**.

Existing accelerator systems — CMS Zero Bias data, Level-1 Data Scouting, shadow trigger crates — show these components individually feasible. A controlled numerical study quantifies what binding them yields: a content-sensitive pseudo-baseline estimates the retention

of a shifted class at 0.76 (95% interval [0.71, 0.81]) against a true 0.016, while the content-independent sample measures a deliberately representation-blind class at 0.000 [0, 0.073].

0.4em *Keywords:** learned instruments; selection function; content-independent probability sampling; measurement of retention; inverse-probability estimation; shadow deployment; trigger metrology; measurement uncertainty.

1. The metrological problem

Every high-throughput scientific instrument has a retention function. Some fraction of what occurs at the sensing boundary enters durable scientific memory; the remainder does not. Historically, this function could often be described by explicit thresholds, coincidence rules, geometric acceptance, dead time, detector efficiencies, or other comparatively inspectable conditions. Such systems were never neutral, but their selectivity could at least be stated in variables chosen in advance.

Learned selection changes the epistemic form of the problem. An event may now be retained because a neural representation, reconstruction error, latent-space statistic, distilled anomaly score, or other learned function assigns it a sufficiently exceptional value. In CMS, for example, AXOL1TL is an event-level unsupervised anomaly detector deployed in the Level-1 Global Trigger; it is trained on Zero Bias collision data and selects events according to a learned anomaly score.[1] CICADA provides a distinct instance in which a larger unsupervised teacher is compressed through knowledge distillation into a smaller model suitable for 40 MHz FPGA operation.[2]

The methodological concern is not that such selectors are learned. Nor is it that they sometimes make errors. Any finite-bandwidth acquisition system necessarily makes errors relative to some possible future scientific objective. The concern is narrower: **the data required to measure the selector’s errors may itself be destroyed by the selector.**

Suppose a deployed selection function f_(x) maps an event representation x to a score, and an event is durably retained when

```
f_(x).
```

For a retrospectively specified class Q, its operational retention is

```
R[f](Q;)

= P[X] Q[f_(X)].
```

If events for which f_(X)< disappear before any independent sample is preserved, then Rf may be impossible to estimate once Q becomes scientifically interesting. The instrument has not merely selected the evidence used to answer a scientific question. It has selected the evidence available for auditing its own selection.

The problem is not confined to accelerator physics. Learned functions are entering measurement chains wherever instruments are trained rather than specified: machine-learning models now participate in legally regulated measurement, where their failure modes must be detectable by inspection regimes [14], and in virtual instruments whose outputs require formal uncertainty evaluation before they can stand in for physical meters [15]. In each case the

metrological question is the same one posed here at the acquisition boundary: once a learned component mediates what is measured, the measuring system must retain enough selection-independent information to characterize that mediation. This paper treats the most severe version of the problem—irreversible selection before storage—and its proposal is evaluated, in Section 13, as a measurement method: with explicit estimators, bias, confidence intervals, and sample-size requirements for the retention quantities it is designed to measure.

This circularity is especially consequential for anomaly detection. The advertised scientific objective is sensitivity to classes that were not specified in advance; yet ordinary efficiency studies necessarily use classes that *can* be specified in advance. The unknown class therefore occupies a structurally different position from the benchmark signal. It is the class for which the selector's retention function matters most and for which direct validation is least available.

A control architecture should break that circularity.

2. Baseline capture is necessarily relative, not absolute

The phrase "raw, unbiased data'' is too strong for any real detector. Physical sensors possess thresholds and acceptance regions. Electronics perform shaping, digitization, suppression, clustering, or compression. Hardware architectures may discard information before a software trigger ever sees it.

BCA therefore does not posit an impossible view from nowhere. It introduces the concept of a **fidelity locus**.

The fidelity locus is the earliest representation at which the selection mechanism under audit can practically be bypassed and an independent control sample retained. A BCA claim must always be indexed to that locus.

A hierarchy might be:

The ordering does not imply that a higher-fidelity representation is always economically preferable. It means only that claims about "unbiased'' capture cannot reach upstream of information that has already been irreversibly removed.

This produces an important discipline:

quote **A baseline controls every selection downstream of its declared fidelity locus and none upstream of it.** quote

A 40 MHz stream of Level-1 trigger objects can therefore provide an extraordinarily valuable baseline for auditing an anomaly classifier that consumes those objects, while providing no guarantee about physics signatures already erased during detector readout or trigger-object construction.

CMS Level-1 Data Scouting offers a concrete example of this distinction. The Phase-2 system is designed to capture and process Level-1 trigger information at the full 40 MHz collision rate while bypassing ordinary Level-1 selection.[3,4] That is not "raw reality.'' It is nevertheless an unusually powerful **predecision baseline at a declared representation locus**.

3. The minimal control: content-independent probability sampling

Let X_i denote the event available at the fidelity locus. Before the learned selector's decision is allowed to determine whether X_i survives, define an independent inclusion variable

B_i Bernoulli($_i$).

When $B_i=1$, the event—or the highest-fidelity representation permitted by the control bandwidth—is retained regardless of its anomaly score, reconstructed category, trigger signature, or apparent physical interest.

The strongest and simplest design uses a constant inclusion probability,

$_i=p$,

so that

$P(B_i=1 \mid X_i=x)=p$

for every capturable x.

This is the foundational BCA channel, which we denote **C0**.

The important property is not "randomness' ' in a colloquial sense. It is **content independence with a known inclusion mechanism**. Given the event stream reaching the fidelity locus, the C0 sample does not prefer Standard Model-like events, exotic-looking events, high-complexity events, low-complexity events, sparse events, energetic events, or events occupying a learned latent tail. It samples the stream before those semantic distinctions are permitted to determine retention.

A fixed periodic prescale—every Nth event—may appear equivalent, but periodic accelerator structure, bunch patterns, detector states, or synchronization effects can create accidental correlations. A pseudorandom or suitably hash-derived inclusion mechanism with an auditable seed and known probability is therefore preferable where hardware constraints permit. Randomization should itself be part of the provenance record.

The C0 channel has a second requirement:

$_i>0$

for every event belonging to the population for which baseline claims are made.

This is the familiar positivity principle in a new instrumental setting. An event class assigned zero probability of entering the control archive cannot subsequently be used to estimate what the primary selector would have done to that class. The scientific requirement is therefore not that the control preserve every event, which is impossible at the relevant rates, but that no capturable region be *categorically excluded by content*.

4. Why sketches, random projections, and alternative neural models are not baselines

Several attractive alternatives appear "less biased' ' than a learned trigger while still imposing a content-sensitive selection function.

A streaming histogram that preferentially retains events in sparse bins has a model of interestingness: low empirical occupancy. A quantile sketch preferentially retains tail events. A fixed random projection followed by an outlier threshold selects according to geometry in the projected space. A maximum-entropy flow remains a trained transformation whose behavior depends on its training distribution and objective. A nonparametric distance measure still defines a notion of distance.

These methods may be excellent discovery instruments. Some may possess blind spots almost orthogonal to those of a neural anomaly detector. That is precisely why they belong in a pluralistic acquisition architecture. But none is a statistically neutral control once its output determines inclusion.

The control/experimental distinction can therefore be stated categorically:

Control selection asks: *Was this event selected independently of its content?*

Discovery selection asks: *Does some function of this event make it worth retaining?*

No sophistication of the second question turns it into the first.

This matters because an architecture designed to audit learned priors should not hide a new prior inside its control condition.

[FIGURE 1 — the four metrology channels against a generic learned-trigger chain; see the PDF] The BCA control plane against a generic learned-trigger chain. Solid path: the authoritative selection. Dashed: the four metrology channels of Sections 3–8 — the content-independent probability sample (C0), the predecision representation tap (C1), the statistically legible enrichment plane (C2), and shadow selectors (C3) — feeding the replay bank. fig:bca

5. C1: the full-rate predecision representation tap

Uniform full-fidelity sampling supplies the clean control, but its statistical power against rare phenomena may be limited by the small permissible sampling fraction. A second channel can answer a different question at much higher rate; figure places all four metrology channels of this and the following sections against a generic learned-trigger chain.

Let g(X) be the exact representation consumed by the deployed classifier. The **C1 representation tap** records, at the highest sustainable rate,

$$(g(X_i), f\_(g(X_i)), {}_i, D_i, V_i, C_i),$$

where D_i is the classifier's decision, V_i records the complete model/firmware version, and C_i records relevant detector and run conditions.

C1 need not retain the complete event. Its purpose is to retain **the evidence required to reconstruct the selector's decision surface**.

This distinction may dramatically reduce the bandwidth required for auditability. If an anomaly trigger consumes a compact vector of Level-1 objects, there is no necessity to preserve the full detector payload for every crossing merely to ask, years later, what score the trigger assigned to the input it actually saw.

CMS Level-1 Data Scouting demonstrates the feasibility of the underlying idea at unusual scale: the system is designed to capture Level-1 trigger information at the full 40 MHz colli-

sion rate, and the current Phase-2 baseline bypasses ordinary L1 selection before server-side event building and analysis.[3,4]

C1 does **not** replace C0. A complete record of classifier inputs can establish the selector's behavior on those inputs but cannot recover physical information that the representation g discarded. Nor does it automatically tell us what unknown physical class generated a particular stored vector. C1 is thus a **decision-replay plane**, while C0 is a **population-sampling plane**.

Together they answer different questions:

quote *C0: What did reality sampled at this locus contain?* quote

quote*C1: What did the selector see and decide?*quote

Their conjunction is substantially more powerful than either alone.

6. C2: statistically legible enrichment

A content-independent probability sample is inefficient for phenomena with extremely small prevalence. If a class Q occurs with prevalence q, N events reach the baseline locus, and each is sampled with probability p, then the expected number of captured instances is

```
E[ n[Q] ]=Np q.
```

For small q,

```
P(n[Q]=0) e^-Npq.
```

No baseline architecture defeats this arithmetic. It cannot guarantee that an arbitrarily rare unknown event will survive.

The proper response is not to contaminate the C0 control, but to introduce a separate **C2 enrichment plane**.

Suppose events are partitioned into coarse strata h(X), and events in stratum h are sampled with probability

```
(X)=[h(X)].
```

The sampling probability may deliberately be increased for low-multiplicity events, unusual occupancy patterns, particular detector conditions, extreme values of simple primitives, or regions proposed by a streaming sketch. Provided that the inclusion mechanism is recorded and

```
0<(X) 1,
```

the sample remains statistically interpretable.

For a retrospectively defined class A, the retention of selector f can then be estimated by inverse-probability weighting:

```
R[f](A)

= ᵢ: Bᵢ=1 ᵢ^-1 1(Xᵢ A) 1(f(Xᵢ) ) ᵢ: Bᵢ=1 ᵢ^-1 1(Xᵢ A) .
```

The point is not that this estimator solves every covariate-shift or rare-event problem. It is that **content-sensitive enrichment can remain auditable if its sampling law is explicit**.

Streaming sketches therefore have an important role in BCA—but as proposal mechanisms for C2, not as replacements for C0.

The distinction provides an architecture with both epistemic cleanliness and practical efficiency:

quote *C0 = low-rate, assumption-minimal control* quote

The estimator lineage here is classical rather than novel: whenever inclusion probabilities are known and positive, population *totals* can be estimated without design bias by Horvitz–Thompson weighting [13]. The retention fraction used above is the corresponding inverse-probability-weighted ratio (H'ajek) estimator—a ratio of two estimated totals—which under known positive inclusion probabilities is design-consistent, though not in general exactly unbiased at finite sample size. C2's requirement is precisely that learned enrichment never destroys the knowability of those inclusion probabilities; the statistics is seventy years old, and the only question is whether the acquisition architecture preserves its preconditions.

7. C3: shadow selectors and correlated blind spots

A further plane addresses a different failure mode: even if several selectors individually perform well, their errors may overlap.

Let f[A],f[B],„fₖ be candidate anomaly or classification systems operating on the same event representation. In **C3 shadow mode**, all selectors score the event, but their decisions do not determine whether the event enters the control archive.

For each selector j, define a miss indicator on a subsequently identified class Q,

$$M_j(X)=$$

$$1[\ f_j(X)<_j\ ].$$

A baseline sample then permits measurement not merely of individual miss rates

$$q_j=P(M_j=1),$$

but of joint misses,

$$q_jk=P(M_j=1,\ M_k=1).$$

Under independent failures, the expected overlap is

$$q_j\ q_k.$$

The excess

$$_jk=q_jk-q_j\ q_k$$

measures positive or negative miss dependence at the chosen operating points.

This matters because nominal architectural diversity does not guarantee epistemically useful diversity. Two high-performing selectors whose blind spots coincide may provide less

protection against unforeseen phenomena than two weaker selectors whose misses are complementary.

A shadow deployment is already technologically familiar at CMS. The Global Trigger test crate is a copy of the production Global Trigger that receives the same live inputs while its output is not used to determine detector readout; it has been used to integrate and validate anomaly-detection algorithms without interrupting normal acquisition.[5]

BCA generalizes this engineering convenience into a metrological principle:

quote **Candidate selectors should be compared on common predecision events before any one of them becomes the sole arbiter of which events survive.** quote

This creates an empirical basis for selecting architectures not merely by latency, resource consumption, or benchmark AUC, but by the degree to which they add genuinely independent discovery coverage.

8. The replay bank

The four planes become scientifically durable only if the resulting control corpus can be reinterpreted under future models.

BCA therefore requires a **replay bank**. Each retained event should be bound, where technically possible, to:

- the captured representation at the declared fidelity locus;
- the original event/run/bunch identifiers needed for synchronization;
- the inclusion probability and sampling-channel identity;
- the production selector's score and decision;
- threshold and prescale state;
- complete model, firmware, quantization, and compiler versions;
- calibration and detector-state identifiers;
- shadow-selector outputs;
- a cryptographic or otherwise durable integrity record sufficient to establish that replay is occurring against the original captured representation.

The purpose is temporal separation.

At time t_0, the experiment does not know which future anomaly class will matter. It therefore cannot construct a representative validation panel for that class.

At time t_1, some phenomenon Q may be defined through another experiment, another trigger stream, a later theoretical development, an improved reconstruction, or an archival discovery.

If the relevant C0/C1 material survives, the original selector can then be asked retrospectively:

quote *What would the instrument at t_0 have done to Q?* quote

A replay bank therefore converts some portion of otherwise irrecoverable epistemic uncertainty into a delayed measurement problem.

This is different from simply preserving training data. Training corpora characterize what the model learned from. A replay bank characterizes **what the model would have rejected**.

9. Baseline capture and retention maps

The principal BCA output is not a single "bias score.'' It is a **retention map**.

For a family of retrospectively specified classes or controlled perturbations Qlambda, indexed by physical or representational coordinates lambda,

```
R[f](;)

= P[X Qlambda] [f(X) ].
```

The coordinates might encode multiplicity, sparsity, constituent structure, energy scale, topology, detector occupancy, representation complexity, or any other axis for which a suitable panel can be constructed.

Such a map changes the epistemic meaning of a claim such as "model-independent anomaly detection.'' It does not demand that the detector become equally sensitive to every conceivable phenomenon, an impossible standard. It demands instead that the known retention geometry be published at the operating points that matter.

A learned trigger would then be documented not merely by architecture, training set, loss, latency, resource utilization, and ROC curves on selected benchmarks, but also by:

- fidelity locus;
- control sampling rate;
- baseline exposure;
- retention surfaces;
- directional tests;
- model-version drift;
- shadow-model miss overlap;
- regions for which no meaningful retention bound has yet been established.

That final category is as important as the measured ones. A metrological standard should distinguish **known low retention** from **retention not characterized**.

10. BCA is not another anomaly trigger

A likely objection is that scarce bandwidth should be devoted to the best possible discovery algorithm rather than to deliberately unselective events, most of which will be scientifically ordinary.

The objection mistakes the purpose of the control.

The optimal discovery selector and the optimal measurement of selection bias solve different optimization problems.

A discovery channel asks

```
[f] E[scientific utility of retained events]
```

subject to a bandwidth constraint.

The control channel asks something closer to

`I(B;X)`

subject to a required baseline sample rate, where I(B;X) denotes dependence between event content and inclusion.

For ideal C0 sampling,

`I(B;X)=0`

relative to the declared fidelity-locus population.

Any attempt to make the control “smarter’ ’ by preferentially saving unusual-looking events increases its dependence on a prior definition of unusualness. That may increase discovery yield. It simultaneously weakens its status as a control.

The correct architecture therefore does not choose between intelligence and neutrality. It **budgets separately for them**.

11. Existing accelerator infrastructure as proof of feasibility

BCA does not require the invention of every component from first principles. Current CMS infrastructure already demonstrates three of the central operations independently.

First, **Zero Bias data** supply collision data independent of a targeted physics signature and are sufficiently established that AXOL1TL itself is trained on a Zero Bias dataset.[1]

Second, **Level-1 Data Scouting** is explicitly designed to capture Level-1 trigger information at the full 40 MHz collision rate while bypassing ordinary Level-1 selection. The 2026 Phase-2 demonstrator uses FPGA readout and preprocessing before server-side event building and online analysis.[3,4]

Third, the **Global Trigger test crate** receives the same collision inputs as the main GT but does not control CMS readout, making live shadow evaluation of candidate algorithms possible.[5]

These systems were developed for their own experimental purposes; none should be retroactively redescribed as an implementation of BCA. Their significance is architectural. Together they demonstrate that the three operations BCA requires most urgently—selection-independent sampling, predecision/full-rate representation access, and non-authoritative shadow scoring—are not speculative hardware fantasies.

The missing step is to connect them through a common statistical and provenance protocol whose explicit object is **selection metrology**.

11a. Relation to existing selection-independent practice

BCA’s components have partial precedents across all four large LHC experiments, and the proposal should be read against that practice rather than beside it.

Trigger-level and scouting streams. ATLAS Trigger-Level Analysis records reduced trigger-object information at rates far above full-readout bandwidth and has carried published dijet searches from the first Run 2 result to the full Run 2 dataset [6,7]. CMS data scouting, prototyped in 2011 and developed continuously since, together with data parking, is now documented as a comprehensive program-level strategy [8]. LHCb's Turbo stream and real-time analysis model go furthest: trigger-level candidates are the analysis data, with the Tesla application persisting online-reconstructed objects in analysis-ready form [9,10]. These systems demonstrate, at production scale, that reduced predecision representations can be retained at rates orders of magnitude above conventional readout — which is C1's feasibility claim made by other means.

What existing practice does not yet supply. Each of these streams is signal-directed: scouting and TLA record what *passed* a (lowered) selection, and Turbo persists candidates a trigger line chose. None is, or claims to be, a content-independent probability sample of the predecision population, and none binds its channels to a common statistical protocol whose explicit estimand is the selector's own miss structure. Zero-bias and prescaled minimum-bias samples do supply content-independence, but at rates set by bandwidth accounting rather than by the variance requirements of selection metrology, and without coupling to shadow scoring or replay. The machine-learning real-time-analysis literature [11] likewise concentrates on what learned triggers can *find*, not on the control channels needed to measure what they foreclose. BCA's contribution is not any single component but the protocol that connects them with selection metrology as the declared object.

Preservation context. The replay bank inherits its warrant from the data-preservation tradition: DPHEP has argued for two decades that scientific value accrues to data whose future questions cannot be specified at acquisition time [12]. BCA extends that argument one stage upstream — from preserving what was recorded to preserving the evidentiary basis for estimating what was not.

12. Minimal implementation

A minimally viable BCA need not begin as a large new acquisition system.

For a learned trigger operating on representation g(X), the minimum implementation would contain:

C0. A randomized bypass stream. A known probability p of events reaching the relevant fidelity locus are retained independently of classifier score.

C1. A decision ledger. At maximum feasible rate, preserve the exact classifier input or sufficient exact representation, score, threshold, decision, and version state.

C3. At least one shadow selector. A materially different model or nonlearned selector scores the same inputs without controlling readout.

Replay provenance. Preserve sufficient versioning to reproduce each historical decision.

Retention reporting. At every major model release, replay standard withheld panels and publish retention at operationally relevant rate points, including directional tests where classes can exchange training and anomaly roles.

C2 enrichment can be added as resources permit.

This architecture is intentionally modular. An experiment that cannot preserve raw detector events at useful C0 rates might preserve trigger primitives. An experiment unable to maintain a full-rate C1 archive might retain rolling windows, compressed sufficient inputs, or statistically sampled representation records. The quality of the baseline should be described, not idealized.

A BCA implementation therefore requires a **Baseline Capture Statement** analogous to an instrument calibration statement:

quote fidelity locus; upstream irreversible transformations; C0 inclusion law; effective sample exposure; retained representation; known data-dependent exclusions; C1 coverage; C2 enrichment laws; C3 shadow models; model and firmware provenance; replay horizon. quote

Such a statement would make explicit exactly what the control can and cannot audit.

13. Numerical validation: retention measurement under four acquisition regimes

The architecture's estimation claims can be demonstrated on a controlled stream in which the true retention of every class is known. A synthetic stream of N=410⁶ events xR⁸ was generated from an anisotropic Gaussian background. The deployed selector f is the reconstruction error of a rank-4 linear autoencoder trained on a background-only zero-bias sample, with threshold set to a background retention of 10^-3; a shadow selector f' of the same family was trained on a disjoint sample with one input feature removed. Four anomaly classes were injected: Q_1, off-manifold events the selector is built to notice (true Rf=0.816); Q_2, distribution-shifted events (R=0.016); Q_3, *representation-blind* events constructed to lie exactly on the learned manifold, so that R=0 although the class is physically distinct; and Q_4, a rare variant of Q_1 at prevalence q_4=210^-5, included to exhibit the sample-size relation of Section 6. All code is seeded and deposited (see Data and code availability).

Four acquisition regimes were compared (figure , table). Regime A retains only what the trigger selects. Regime B adds a content-sensitive tap—events flagged by f'—and treats it as a baseline. Regime C is the C0 channel: uniform content-independent sampling at p=210^-3. Regime D halves the uniform floor to p_0=10^-3 and adds a C2 enrichment band around the decision threshold (0.8 f< at =210^-2; f at 210^-3), with inverse-probability estimation and design-based variance. Interval construction is stated for reproducibility: within-class proportions from uniform samples (regimes B and C) carry 95% Wilson score intervals; regime D uses the H'ajek ratio estimator with the design-based variance Var(R)=$_i$ w_i(w_i-1)(y_i- R)²/($_i$ w_i)² for Bernoulli sampling, w_i=$_i$^-1, and normal-approximation intervals truncated to [0,1].

The results reproduce, quantitatively, each claim of the architecture. Under regime A the archive audits itself as perfect: every retained event was, by construction, retained, so an archive-only audit returns R=1 for every class—an error of +0.98 for Q_2 and +1.0 for the blind class—and no confidence statement is even defined, because the denominator population no longer exists. Under regime B the pseudo-baseline is not merely imprecise but confidently wrong: it estimates Rf=0.76 with a tight interval [0.71,0.81] against a true value of 0.016, because f' misses nearly the same Q_2 events f misses, and it returns no estimate at all for Q_3, which f' also cannot see. The shadow plane quantifies this directly: the miss indicators of f and f' have -correlations of 0.55–0.75 across background and visible classes, and both selec-

tors miss the blind class at rate 1.0, where the correlation is undefined because there is no variation left to correlate—the correlated-blindness datum of Section 7 observed in numbers.

Under regime C, content-independent uniform sampling permits direct within-class estimation of every retention value, free of the selection bias that regimes A and B introduce: R(Q_1)=0.841 with 95% interval [0.71,0.92] covering the true 0.816; R(Q_2)=0.000 with an honest upper bound; and, decisively, the blind class is *measured*—R(Q_3)=0 on n=49 baseline events with interval [0,0.073]—an estimate no content-sensitive channel in this experiment could produce at any budget. The rare class exhibits the design equation: with Npq_4=0.16 expected baseline events, one was observed, and the relation P(n[Q]=0) e^-Npq tells the designer in advance that this budget cannot resolve Q_4—the baseline, unlike the trigger archive, states the limits of its own competence. Under regime D, enrichment with recorded inclusion probabilities preserves validity (Q_1: 0.874 [0.77,0.98]; Q_2: 0.021 [0,0.064]) at two-thirds of regime C's storage, and the blind class is still reached—but only through the uniform floor p_0, since any enrichment stratum defined through the selector's own representation inherits that representation's blindness. A preliminary design in which the enrichment stratum was keyed to the selector's top scores concentrated the budget on events the selector already retained and starved the miss region; with weights strongly correlated with outcomes, naive effective-sample-size intervals also overstate precision. Both defects are corrected by band enrichment around the threshold and design-based variance, and both are reported here deliberately: they are the form miscalibration takes *inside* the control plane when C2 is designed for discovery convenience rather than audit coverage, and they reinforce the architectural rule that C0 is the floor on which every other channel stands.

[figure: bca_validation.pdf] Estimated retention Rf for three anomaly classes under four acquisition regimes at comparable storage budgets (A: trigger archive only; B: content-sensitive tap as pseudo-baseline; C: C0 uniform probability sample; D: C0 floor with C2 near-threshold enrichment and inverse-probability estimation). Dotted lines mark true values. Regimes A and B (red) produce confident error; regimes C and D (blue) produce calibrated estimates, including a measured retention of zero for the representation-blind class Q_3, which regime B cannot see at all. fig:validation

Regime & Stored & Q_1 (true 0.816) & Q_2 (true 0.016) & Q_3 (true 0.000) & Q_4 (true 0.862)

A trigger-only & 20,659 & 1.000 (no CI) & 1.000 (no CI) & 1.000 (no CI) & 1.000 (no CI)

B content-sensitive & 18,181 & 0.991 [0.989,0.992] & 0.759 [0.705,0.807] & no estimate & 0.983 [0.907,0.997]

C C0 uniform & 7,925 & 0.841 [0.706,0.921] & 0.000 [0,0.077] & 0.000 [0,0.073] & n=1; Npq=0.16

D C0+C2, IPW & 5,371 & 0.874 [0.771,0.978] & 0.021 [0,0.064] & 0.000 (n=25, all missed) & n=0 (off-band)

Retention estimates with 95% intervals under the four regimes of figure); class prevalences 510^-3, Q_4: 210^-5). Regime B's Q_2 row is the correlated-miss failure mode: a precise interval around a wrong value. Regime C's Q_3 row is the architecture's central capability: the selector's blind spot measured, with uncertainty, from the control channel. tab:validation

14. Limitations

BCA does not solve the unknown-unknown problem. A probability sample can contain a phenomenon without anyone recognizing it; a phenomenon sufficiently rare may not enter the sample at all; and a physical signature erased upstream of the fidelity locus cannot be reconstructed through downstream sampling.

Nor does BCA eliminate representation dependence. Its purpose is almost the opposite: to make the consequences of representation dependence measurable.

The architecture also consumes real resources. Every bit reserved for control capture is a bit unavailable for some other acquisition purpose. The optimal baseline fraction is therefore an experimental-design question, not a universal constant. Its justification should be evaluated in the same way as other calibration, monitoring, commissioning, and systematic-uncertainty budgets: by the information it preserves about the reliability of the scientific instrument.

Finally, the existence of a control stream does not automatically produce a useful audit. Retrospective classes still require labeling, simulation, injections, alternative reconstruction, or independent discovery channels. BCA creates the evidentiary substrate upon which those analyses can operate; it does not predetermine them.

15. Discussion: from trigger performance to trigger metrology

Contemporary learned-trigger work has made impressive progress on a difficult engineering problem: how to execute sophisticated inference within latency and resource constraints that only recently would have excluded such models entirely. CMS now operates a signal-agnostic event-level anomaly trigger at Level-1; CICADA performs distilled anomaly inference at 40 MHz; and Level-1 Data Scouting is being developed precisely to open physics access beyond ordinary Level-1 accept constraints.[1–4]

The next methodological step is different. It is not another improvement in classification performance. It is the construction of an independent observational channel from which classification performance can be measured in directions the classifier did not define.

This becomes particularly important when the selector’s scientific justification is its ability to retain the unforeseen. Benchmark performance on known alternatives is necessary evidence for such a system, but it cannot by itself characterize sensitivity to the space of alternatives that were not represented by those benchmarks. Where the classifier acts before irreversible data loss, this uncertainty acquires an unusual form: the events needed to expose the blind spot may be preferentially absent from the surviving record.

Baseline capture addresses that problem by separating **discovery selection from selection metrology**.

The proposal is modest in its strongest claim. It does not assert that learned triggers are uniquely unreliable. It does not assert that an unknown physical phenomenon has already been discarded. It does not require a hypothetical storage system capable of retaining an entire high-rate detector stream.

It requires that a small portion of the acquisition architecture remain outside the selector’s definition of interestingness.

16. Conclusion

A scientific trigger is usually evaluated by asking what interesting events it retains. A learned anomaly trigger invites a second question:

quote **What evidence remains from which we could discover what it systematically failed to regard as interesting?** quote

For systems in which the answer is "only the events the trigger retained,'' the audit is structurally circular.

Baseline Capture Architecture breaks that circle through a layered control plane. A content-independent probability sample supplies the statistically clean baseline. A full-rate representation tap preserves the inputs and decisions of the audited selector. Probability-weighted enrichment improves coverage while retaining inferential traceability. Shadow models expose correlated blind spots before alternative selectors become authoritative. A versioned replay bank allows future anomaly classes to be evaluated against historical instruments.

The architecture does not promise unfiltered access to reality. No detector can. Its claim is more precise:

quote **Every irreversible selector should be accompanied, at a declared fidelity locus, by an observation channel whose inclusion law does not depend on the selector's ontology.** quote

At high data rates, the primary trigger decides what the experiment can afford to remember.

The baseline channel preserves the experiment's ability to determine **what that decision cost**.

Funding

No external funding was received for this work.

Conflict of interest

The author declares no conflicts of interest.

Data and code availability

The numerical validation of Section 13 uses synthetic data generated by a single seeded script (Python/NumPy; seed 20260827); the script, its tabulated output, and the figure are provided with the submission and are sufficient to reproduce every number reported. No experimental data were otherwise generated or analysed: the feasibility argument uses published accelerator trigger and acquisition systems, cited to their public documentation. The canonical archival version of this text, its content-derived identifier, and the associated provenance record are available through Crimson Hexagonal Archive deposit #1453 at https://www.alexanarch.org/s/records/1453/.

Declaration of AI-assisted technology

AI-assisted tools (Anthropic Claude; the present revision and the numerical validation of Section 13 with Claude Fable 5) were used in the preparation and critical revision of this manuscript: in drafting and redrafting prose from the author's specifications, in locating and retrieving candidate literature for the related-practice section, and in typesetting. The author formulated the architecture and its claims, directed the research, verified every cited source and technical claim against its primary publication, determined the final argument and wording, and takes full responsibility for the content of the submitted work. No AI tool is credited as an author.

Acknowledgments

The author thanks the CMS, ATLAS, and LHCb collaborations for the public documentation of their trigger, scouting, and real-time-analysis systems, without which the feasibility argument of Section 11 could not be made concrete.

References

- CMS Collaboration, **Anomaly detection with AXOL1TL at the CMS Level-1 Trigger in 2024 and 2025**, CMS Detector Performance Summary CMS-DP-2025-061, CDS record 2942560 (2025).
- CMS Collaboration, **Model-Independent Real-Time Anomaly Detection at the CMS Level-1 Calorimeter Trigger with CICADA**, CDS record 2917884 (2024).
- R. Ardino, on behalf of the CMS Collaboration, **Development and demonstration of the CMS Phase-2 Level-1 trigger Data Scouting baseline system for HL-LHC**, **JINST 21** (2026) C01024.
- D. S. Rabady, on behalf of the CMS Collaboration, **A 40 MHz Level-1 trigger scouting system for the CMS Phase-2 upgrade**, **Nucl. Instrum. Meth. A 1047** (2023) 167805.
- M. Quinnan, on behalf of the CMS Collaboration, **Anomaly Detection in the CMS L1 Trigger**, **EPJ Web Conf. 337** (2025) 01032.
- ATLAS Collaboration, Search for low-mass dijet resonances using trigger-level jets with the ATLAS detector in pp collisions at s=13 TeV, **Phys. Rev. Lett. 121** (2018) 081801 [arXiv:1804.03496].
- ATLAS Collaboration, Search for electroweak-scale dijet resonances using trigger-level analysis with the ATLAS detector in 132 fb^-1 of pp collisions at s=13 TeV, **Phys. Rev. D 112** (2025) 092015 [arXiv:2509.01219].
- CMS Collaboration, **Enriching the physics program of the CMS experiment via data scouting and data parking**, **Phys. Rept. 1115** (2025) 678 [arXiv:2403.16134].
- LHCb Collaboration, R. Aaij et al., **Tesla: an application for real-time data analysis in High Energy Physics**, **Comput. Phys. Commun. 208** (2016) 35 [arXiv:1604.05596].
- LHCb Collaboration, R. Aaij et al., **A comprehensive real-time analysis model at the LHCb experiment**, **JINST 14** (2019) P04006 [arXiv:1903.01360].
- **Review of Machine Learning for Real-Time Analysis at the Large Hadron Collider experiments ALICE, ATLAS, CMS and LHCb**, arXiv:2506.14578.

- DPHEP Collaboration, **Data preservation in high energy physics**, **Eur. Phys. J. C 83** (2023) 795.
- D. G. Horvitz and D. J. Thompson, **A generalization of sampling without replacement from a finite universe**, **J. Am. Stat. Assoc. 47** (1952) 663.
- A. G. da Silva Santos, L. F. Rust Carmo and C. B. do Prado, **Machine learning in legal metrology—detecting breathalyzers' failures**, **Meas. Sci. Technol. 35** (2024) 045015.
- N. Bayazit, M. Marschall, M. Straka and S. Schmelter, **A novel framework for the uncertainty evaluation of virtual flow meters**, **Meas. Sci. Technol. 36** (2025) 076012.

document

Suggested Citation

Sharks, Lee. “Baseline Capture Architecture for Learned Scientific Triggers: A Control Plane for Measuring Selection Before Irreversible Data Loss (EA-SEI-BCA-01 v2.0 — the submitted state)” *Alexanarch*, 2026-08-27. https://www.alexanarch.org/s/records/1558/

Deposit Information

This paper is deposit #1558 in Alexanarch, the self-governing successor to the Crimson Hexagonal Archive on Zenodo (community `crimsonhexagonal`, terminated 2026-06-19). The AXN identifier `AXN:0656.OPERATIVE` is sovereign — minted and maintained by Alexanarch.

- Canonical URL: https://www.alexanarch.org/s/records/1558/
- 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/1558/

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