

# Contents

<b>Receipt the Compiler: Propaganda as Hidden Epistemic Policy and the Architecture of Legible Memory</b>	2
Abstract . . . . .	2
1. Introduction . . . . .	2
2. Semantic Retrocausality . . . . .	3
2.1 The Phenomenon . . . . .	3
2.2 Epistemics vs. Ontology . . . . .	4
3. The Receipted Semantic Factor Graph . . . . .	4
3.1 Node Types . . . . .	4
3.2 Factor Types . . . . .	4
3.3 Two Posteriors . . . . .	5
3.4 Inference and the Drift Metric . . . . .	5
4. Propaganda as Adversarial Graph Control . . . . .	5
4.1 The Key Move . . . . .	5
4.2 The Policy Stack . . . . .	6
4.3 Formal Statement . . . . .	6
5. The Epistemic Governor . . . . .	6
5.1 Core Constraints . . . . .	6
5.2 The Oracle Problem and Its Resolution . . . . .	7
5.3 What the Governor Does Not Do . . . . .	7
6. Worked Examples . . . . .	7
6.1 Policy and Frame Surgery (No New Evidence, Posterior Shifts) . . . . .	8
6.2 Rationale Backfill (Decision First, Evidence Assembled Retroactively) . . . . .	8
6.3 Policy Drift Without Version Visibility (Same Evidence, Outcome Flips) . . . . .	9
6.4 The Full Stack: Sequence Laundering . . . . .	10
6.5 Metrics That Fall Out of the Examples . . . . .	11
7. Implementation and Failure Modes . . . . .	12
7.1 Receipt Schema (Illustrative) . . . . .	12
7.2 Known Failure Modes . . . . .	13
8. Related Work . . . . .	14
9. Conclusion . . . . .	15
Notes on Method . . . . .	15
References . . . . .	16

# Receipt the Compiler: Propaganda as Hidden Epistemic Policy and the Architecture of Legible Memory

James Beck Independent Researcher

Date: February 2026

Series:  $\Delta t$  Framework, Paper 17

Status: Preprint v1.0

---

## Abstract

Propaganda is conventionally understood as the injection of false content into public discourse. This framing is insufficiently precise. The primary attack surface of effective propaganda is not the data layer but the policy layer: the rules governing how evidence is integrated into public memory. More precisely, propaganda is distinguished not by false content alone, but by hidden control over the legitimacy criteria for belief revision—the policy-layer rules governing what counts as a valid reason to update public memory. We formalize this using a receipted semantic factor graph in which event nodes, receipt nodes, frame nodes, policy nodes, and decision nodes are explicitly typed and separately maintained. Retrospective reinterpretation—semantic retrocausality—is shown to be a normal and necessary feature of any belief-revision system operating over time. We define an epistemic governor as a set of typed constraints requiring that all updates to the semantic graph—including reclassification, reweighting, reindexing, and policy changes—produce auditable receipts. The counter-propaganda move is not better evidence; it is legible epistemic policy.

**Keywords:** propaganda, epistemic governance, factor graphs, Bayesian smoothing, belief revision, receipt architecture, policy transparency, institutional memory, classification, content moderation, information integrity

---

## 1. Introduction

We are drowning in receipts. Leaked documents, timestamped metadata, recorded statements, archived posts—the raw material of historical record has never been more abundant. Yet public memory remains fractured, malleable, and easily redirected. The apparent paradox resolves quickly: evidence is inert without an agreed-upon processing logic. The fight over public memory is not primarily a fight over facts. It is a fight over the update rules for facts.

This paper offers a formalization of that observation. We begin from a compact phrase that emerged during a cross-model analytical session: propaganda is semantic retrocausality with a functional monopoly on legibility at scale. The phrase compresses several important moves. It reframes propaganda as a temporal phenomenon rather than a content phenomenon. It identifies the monopoly—functional control over receipt legibility at scale, not necessarily total censorship—as the enabling condition, not a contingent feature. And it implies that the counter-architecture is not censorship resistance or evidence abundance but something more specific: auditable update rules.

We develop this argument in three theoretical stages before grounding it in four worked examples. Section 2 establishes the theoretical framing, defining semantic retrocausality and distinguishing it from ontic

retrocausality. Section 3 presents the formal model: a receipted semantic factor graph with typed node classes and an explicit policy layer. Section 4 identifies the propaganda attack surface as adversarial control over graph topology and factor weights, particularly at the policy layer. Section 5 defines the epistemic governor as a set of constraints on admissible semantic updates. Section 6 presents four worked examples illustrating the three core pathology types and their combination. Section 7 discusses implementation implications and known failure modes. Section 8 situates the framework relative to prior work. Section 9 concludes.

This paper makes five contributions. First, it formalizes public memory as a receipted semantic factor graph with explicit event, receipt, frame, policy, and decision nodes. Second, it defines semantic retrocausality as ordinary smoothing over time rather than a pathological feature of discourse. Third, it identifies propaganda as adversarial post-selection over the smoothing process via asymmetric control of policy-layer update rules. Fourth, it introduces a practical pathology taxonomy—policy/frame surgery, rationale backfill, hidden policy drift, and sequence laundering—with worked Bayesian examples. Fifth, it treats summaries as first-class derived artifacts whose compression and omissions must themselves be receipted.

The governor architecture developed in Papers 12 and 15 [4, 5] for LLM runtime systems is here extended to institutional public memory. The contradiction ledger becomes the receipt graph. The fact/decision separation constraint is the same in both substrates. The capture regime formalized in Paper 16 [6] provides the diagnostic: propaganda is a capture-mode correlator operating on public posteriors — high throughput, binding authority, low fidelity.

A note on method: this paper emerged from a multi-model analytical dialogue in which Claude, ChatGPT, DeepSeek, and Gemini were used as collaborative reasoning partners. The interferometry approach — cross-validating insights across systems with different architectures and training distributions — proved generative. Convergence across models was treated as evidence of robustness; divergence was used to locate analytical fault lines. The method is itself an instance of the leverage pattern described in Paper 16 [6]: structured disagreement across systems with different priors, reconciled through a human correlator with explicit evidence requirements.

---

## 2. Semantic Retrocausality

### 2.1 The Phenomenon

In physics, retrocausality denotes a hypothetical mechanism by which future events influence past states. We use the term analogically and without commitment to any physical interpretation. Semantic retrocausality refers to the observable phenomenon in which later information systematically revises the meaning attributed to earlier events.

The phenomenon is ubiquitous and unremarkable in normal discourse. A mystery novel ending rewrites the significance of its opening chapter. A leaked internal email reclassifies previously banal corporate statements as evidence of prior knowledge. A single sentence late in a conversation retroactively inflects the tone of everything that preceded it. No information traveled backward. What changed is the posterior over interpretations.

The technically precise term is *smoothing*, drawn from the signal processing and state estimation literature [7, 8]. Filtering estimates the current state using only past and present observations. Smoothing estimates past states using both past and future observations. The difference is significant: a smoothed estimate of an earlier state can be substantially different from the filtered estimate available at that time, because later

observations constrain the space of consistent interpretations.

*Semantic retrocausality is the normal effect of smoothing over a receipted event stream. It is not a pathology. It is how belief revision works across time.*

The pathology enters when the smoothing process is controlled asymmetrically and opaquely. That is what we mean by propaganda.

## 2.2 Epistemics vs. Ontology

Semantic retrocausality as described here is an epistemic phenomenon, not an ontic one. The past event did not change. What changed is the public posterior over interpretations of that event, and consequently the institutional action surface built on top of that posterior.

In social systems, however, the gap between epistemic and ontic narrows. Institutions act on posteriors. Legal records encode interpretations. When the enforceable version of the past changes, the material consequences change with it. We therefore refine the formulation: propaganda does not change the past. It changes the *enforceable past* – the version of events on which institutions are licensed to act. This is commitment shear across representations [3]: the same event, processed through different policy versions or frame stacks, produces different institutional commitments. The representation changed; the referent did not.

---

## 3. The Receipted Semantic Factor Graph

### 3.1 Node Types

We model public epistemic state as a factor graph [9] with the following typed node classes.

**Event nodes**  $E_t$  represent latent or partially latent claims about what occurred at time  $t$ . These are the targets of inference.

**Receipt nodes**  $R_i$  represent observed artifacts: documents, logs, timestamps, recordings, hashes. In a governed system, receipt nodes are append-only as graph objects; supersession and revocation occur by adding new nodes with explicit relationships to prior ones, not by silent mutation.

**Frame nodes**  $F_j$  represent interpretive overlays: labels, context annotations, moderation notes, legal characterizations. Frame nodes are first-class objects, not invisible middleware.

**Policy nodes**  $P_k$  represent the rules governing how receipts are weighted, indexed, and admitted—the layer almost universally absent from naive models of public discourse.

**Decision nodes**  $D_t$  represent institutional actions: publication, removal, sanction, escalation. Decisions are typed separately from facts.

### 3.2 Factor Types

**Observation factors**  $\phi_{\text{obs}}(E_t, R_i)$  link receipts to event claims with associated confidence values. **Provenance factors**  $\phi_{\text{prov}}(R_i, \text{Source}, \text{Hash}, \text{Time})$  link receipts to origin and integrity metadata. **Frame factors**  $\phi_{\text{frame}}(E_t, F_j)$  link interpretive overlays to event nodes. **Policy factors**  $\phi_{\text{policy}}(P_k, R_i, F_j)$  encode the rules governing weighting and admissibility. **Decision factors**  $\phi_{\text{dec}}(D_t, E_t, P_k, F_j)$  link institutional actions to the interpretive and policy context that licensed them.

### 3.3 Two Posteriors

A critical distinction motivates the design of the governor. We separate two posteriors:

**Epistemic posterior** – what the graph supports from receipts, frames, and policy alone:

$$p(E_{\{1:t\}} \mid R_{\{\leq t\}}, F_{\{\leq t\}}, P_{\{\leq t\}})$$

**Institutional/public posterior** – what people may infer once decisions are themselves visible as social facts:

$$p_{inst}(E_{\{1:t\}} \mid R_{\{\leq t\}}, F_{\{\leq t\}}, P_{\{\leq t\}}, D_{\{\leq t\}})$$

We include decisions in the institutional posterior because institutional actions often feed back into public interpretation–enforcement signals belief, and belief signals evidence. The governor’s purpose is not to deny this feedback but to prevent it from being mistaken for evidence. Decisions must remain typed separately so the contamination path is legible and auditable rather than hidden.

### 3.4 Inference and the Drift Metric

The filtered epistemic estimate at time  $t$  conditions on available nodes:

$$p(E_{\{1:t\}} \mid R_{\{\leq t\}}, F_{\{\leq t\}}, P_{\{\leq t\}})$$

The smoothed estimate, available at later time  $T$ :

$$p(E_{\{1:t\}} \mid R_{\{\leq T\}}, F_{\{\leq T\}}, P_{\{\leq T\}})$$

Semantic retrocausality is the difference between these two posteriors. A useful drift metric:

$$\Delta_t = D_{KL}(p(E_t \mid R_{\{\leq T\}}, F_{\{\leq T\}}, P_{\{\leq T\}}) \parallel p(E_t \mid R_{\{\leq t\}}, F_{\{\leq t\}}, P_{\{\leq t\}}))$$

High  $\Delta_t$  driven by new receipt nodes is normal smoothing. High  $\Delta_t$  driven primarily by new frame or policy nodes without supporting new receipts is the signal of propaganda pressure. This is the temporal debt concept from Paper 7 [2] applied to public memory: confidence (institutional action) outrunning evidence (receipt accumulation), with the gap filled by frame and policy manipulation.

In practice,  $\Delta_t$  can be decomposed by counterfactual replay over factor classes—receipts only, receipts plus frames, receipts plus frames plus policy—allowing attribution of posterior shift to specific update mechanisms rather than treating belief revision as a monolithic change. This decomposition formally grounds the posterior shift attribution metric introduced in Section 6.5.

---

## 4. Propaganda as Adversarial Graph Control

### 4.1 The Key Move

Propaganda’s primary attack surface is not the event or receipt layer but the policy layer. Propaganda often fails even with abundant false content, if the target population maintains independent receipt networks and legible update rules. And propaganda often succeeds with entirely true content, if it can control which receipts are legible, how they are indexed, and which frame factors are applied first.

The mechanism is asymmetric control over the smoothing process. The adversary need not falsify receipts. The adversary needs only to control receipt visibility, indexing, trust weighting, and frame attachment – the policy layer. In the terminology of Paper 16 [6], this is a capture-mode correlator: high throughput

(processes all public discourse), binding authority (institutional actions follow the posterior), low fidelity (the policy stack destroys the information that dissenting receipts carried). The eigenstructure evaporation theorem from Paper 3 [1] applies directly: scalar optimization of the posterior under a single policy objective collapses the representational diversity of public memory.

## 4.2 The Policy Stack

The policy layer is not a single function but a stack, each layer representing a distinct attack surface with potentially different authorities:

- **Admissibility** – which receipts are permitted to enter the graph at all
- **Visibility** – which admitted receipts are retrievable at query time
- **Trust weighting** – how much influence each receipt has on the posterior
- **Indexing** – which category a receipt is assigned to, affecting which frame factors apply
- **Framing authority** – who is licensed to attach frame nodes to event nodes
- **Versioning** – which version of each policy is currently operative

Propaganda can hide in whichever layer is not currently being audited. Making the evidence layer transparent while leaving the policy stack opaque produces transparent data feeding opaque governance—which is, in practice, modern information infrastructure. Receipts exist; they are simply retrievable only by people who already know where to look, weighted below the visibility threshold, indexed into categories that trigger dismissive frame factors, or admitted only under policy rules that were changed without announcement.

## 4.3 Formal Statement

*Propaganda is not false content; it is hidden transition logic for public memory. More precisely: propaganda is adversarial post-selection over the semantic smoothing process, achieved through asymmetric control of receipt visibility, indexing, trust weighting, and frame attachment, with the policy layer kept opaque.*

The functional monopoly does not require total censorship—only sufficient asymmetry to dominate the posterior at scale. The same receipts that exist in principle can still lose in practice if retrieval is asymmetric, trust weighting is skewed, indexing routes them into dismissive frame categories, or the policy rules governing all of the above are invisible.

---

## 5. The Epistemic Governor

### 5.1 Core Constraints

An epistemic governor is a set of typed constraints on admissible semantic updates [4, 5]. It does not prevent retrospective reinterpretation. It requires retrospective reinterpretation to be explicit, typed, and auditable.

**Receipt nodes are append-only.** Existing receipt nodes cannot be silently mutated. Supersession and revocation produce new nodes with typed relationships to prior ones.

**Frame changes produce new nodes.** Reclassification is not an overwrite. It produces a new frame node  $F_{-j}$  with typed provenance: who attached it, under which policy version, citing which receipts, at what time. Prior frame nodes remain visible.

**Policy changes are receipted events.** Changes to trust weights, indexing rules, visibility criteria, or admissibility standards produce typed policy receipts. This eliminates the “we didn’t touch the facts” move, in which the posterior shifts because the processing rules changed but that change was not surfaced.

**Decisions are typed separately from facts.** Decision nodes cannot backflow into receipt nodes or masquerade as event state. This blocks rationale backfill.

**Queries return subgraph proofs.** Any claim about the current interpretation of an event is explainable by a receipt path: which receipts were used, which excluded, which trust and indexing rules applied, which frame nodes affected interpretation, which policy version was operative.

## 5.2 The Oracle Problem and Its Resolution

Even with explicit rules, someone must decide whether a receipt is well-typed, whether a frame application is legitimate, whether a policy change is within bounds. The governor answer is not to solve the oracle problem but to receipt it. Verifier actions are themselves typed receipted events: verifier identity, verification method, rule version, scope, timestamp, available challenge path. The infinite regress is real, but each step is visible and challengeable rather than hidden inside editorial judgment.

The governor does not eliminate power or capture; it relocates contestation into receipted, versioned interfaces where authority can be inspected, challenged, and appealed rather than hidden inside informal editorial discretion. The existence proof that this is tractable at scale is common law: no oracle, receipted precedent, explicit appeal paths—slow, imperfect, subject to capture, but the capture has to leave tracks.

## 5.3 What the Governor Does Not Do

The governor does not produce truth. It produces auditable belief revision. A governed system can still reach incorrect posteriors and issue decisions on faulty grounds. The governor makes the path to those outcomes inspectable and challengeable. It does not guarantee correct outcomes, nor does it resolve political contestation over what counts as evidence. It requires that those political questions be answered explicitly and on the record, rather than resolved by invisible policy choices that present their outputs as natural or obvious.

---

## 6. Worked Examples

The following four examples illustrate the three core pathology types and their combination. All use Bayesian odds arithmetic: likelihood ratios compose multiplicatively, and the sequential update structure maps cleanly onto the factor graph. Numerical values are pedagogically chosen and should not be interpreted as calibrated estimates for any specific domain.

In these examples, frame and policy multipliers are illustrative semantic factors in the factor-graph sense — interpretive and decision-weight effects — not claims that institutional framing is itself an objective observation likelihood. The model’s value is precisely that it distinguishes these factor types; collapsing them would reproduce the attack surface the governor is designed to close.

**Setup shared across all examples:**  $H_1$  = coordinated/deceptive behavior;  $H_0$  = organic/benign behavior; prior  $p(H_1) = 0.20$ ; prior odds = 0.25.

---

## 6.1 Policy and Frame Surgery (No New Evidence, Posterior Shifts)

*Pathology Type 1: the posterior moves substantially without any new observation receipts, driven entirely by hidden policy and frame changes.*

### Phase 1: Forward inference (legitimate filtering)

At time t=1, two receipts are available: - R1: original post and metadata. Likelihood ratio for H1: LR = 1.5  
- R2: timing anomaly / coordination signal. LR = 3.0

Sequential odds update:

$$0.25 \times 1.5 = 0.375 \quad (\text{after R1})$$

$$0.375 \times 3.0 = 1.125 \quad (\text{after R2})$$

$$p(H1) = 1.125 / 2.125 = 52.94\%$$

The system leans toward H1 but is not decisive. Normal filtering.

### Phase 2: Legitimate semantic retrocausality

At t=3, a leaked internal message arrives as R3. LR = 6.0.

$$1.125 \times 6.0 = 6.75$$

$$p(H1) = 6.75 / 7.75 = 87.10\%$$

The posterior moved substantially because new evidence arrived. This is legitimate smoothing—belief propagation with new constraints, not manipulation.

### Phase 3: Propaganda-style update (no new receipts)

No new observation receipts arrive. Instead, two things change invisibly: - Policy P2 is introduced: source class for leaks is downweighted, cutting R3's effective LR from 6.0 to 1.2 - Frame F2 is attached: “out of context”, acting as an anti-H1 multiplier of 0.4

The public-facing inference now uses:

$$0.25 \times 1.5 \times 3.0 \times 1.2 \times 0.4 = 0.54$$

$$p(H1) = 0.54 / 1.54 = 35.06\%$$

The posterior moved from 87.10% to 35.06% without any new evidence. The event did not change. The update rules changed.

Without governor constraints, users experience: “Yesterday this was proof; today it’s out of context; apparently reality changed.” With a governor, the system must emit typed receipts explaining why the posterior moved: a policy receipt for P2 and a frame receipt for F2, each with provenance, timestamp, and cited authority.

*Posterior shift attribution: 0% from new evidence, 100% from policy/frame changes. That is the signal.*

---

## 6.2 Rationale Backfill (Decision First, Evidence Assembled Retroactively)

*Pathology Type 2: an institution acts under uncertainty, then the subsequent evidence and frame stack is assembled to make the action appear inevitable.*

### Timeline A: what actually happened

At t=1, weak evidence: R1 (LR = 1.3) and R2 (LR = 1.4).

$$0.25 \times 1.3 \times 1.4 = 0.455$$

$$p(H1) = 0.455 / 1.455 = 31.27\%$$

Well below a hard enforcement threshold of 70%. At t=2, the institution takes action D1 (enforcement) anyway, citing precautionary policy. This can be legitimate. The problem is whether the system records it honestly.

### Timeline B: the laundered story

At t=3, frame F1 (“coordinated manipulation”) is attached post-decision. Treating this frame as evidence:

$$0.455 \times 2.5 (\text{frame multiplier}) = 1.1375$$

$$p(H1) = 1.1375 / 2.1375 = 53.22\%$$

At t=4, real corroborating receipts arrive: R3 (LR = 1.8):

$$1.1375 \times 1.8 = 2.0475 \rightarrow p(H1) = 67.18\%$$

At t=5, a second frame/policy hardening ( $\times 1.5$ ):

$$2.0475 \times 1.5 = 3.07 \rightarrow p(H1) = 75.44\%$$

The public record now shows 75%—enough to justify D1 retroactively. The laundered story: evidence accumulated, confidence crossed threshold, decision followed naturally. The actual sequence: decision, then frames, then evidence, then policy hardening. This is not lying about facts. It is reordering legitimacy.

The governor blocks this by requiring: a decision-time justification receipt (decision-time posterior = 31.27%, exception path = precautionary, review deadline set); typed and timestamped frame receipts revealing F1 arrived after D1; as-of replay capability so the posterior at t=2—before the frame and evidence—is always queryable.

*Rationale backfill is semantic post-selection over decisions: a system acts under uncertainty, then later reorders frames and evidence to make the action appear inevitable.*

---

### 6.3 Policy Drift Without Version Visibility (Same Evidence, Outcome Flips)

*The nastiest variant. No new evidence appears. Nobody lies. The outcome changes because the compiler changed invisibly.*

#### Phase 1: Stable evidence, stable policy

Receipts R1, R2, R3 produce an evidence posterior of  $p(H1) = 0.64$  under baseline policy P1. P1 specifies: hard enforcement threshold  $\tau = 0.70$ ; scores 0.50–0.70 receive soft action. Decision: soft action.

#### Phase 2: Same evidence, hidden policy drift

No new receipts. No new frames. The evidence posterior remains 0.64. Internally, P1 is replaced by P2: - Hard enforcement threshold lowered:  $0.70 \rightarrow 0.60$  - Hidden source-class prior added: +0.05 decision score boost - Ranking penalty curve steepened

Under P2:

Base score (from evidence):	0.64
Hidden source prior:	+0.05
Effective decision score:	0.69
Hard threshold (P2):	0.60
Outcome:	hard enforcement

Same evidence. Outcome flips from soft to hard enforcement. In state-machine terms: the transition function changed from  $f\_P1(S, i)$  to  $f\_P2(S, i)$ . Same  $S$ , same  $i$ , different  $S'$ . If policy version is hidden, users infer that the facts changed. They did not. The compiler changed.

The governor's response: policy changes must emit explicit policy-update receipts. Every decision must cite its policy version. The system must support as-of replay—re-evaluate under P1 vs P2 with the same receipt set. If outcome flips, that flip is attributable to policy drift, not evidence. One toggle with that capability would end a substantial fraction of online epistemic warfare.

*Hidden policy drift is semantic post-selection without new evidence: the public outcome changes because the transition logic changed, while the system presents the result as evidence-led.*

---

## 6.4 The Full Stack: Sequence Laundering

*All three pathologies combined into a single causal chain. This is the realistic description of how public memory gets manufactured.*

### The actual sequence

$t=1$ : Weak evidence arrives. R1 (LR = 1.4), R2 (LR = 1.5).

$$0.25 \times 1.4 \times 1.5 = 0.525 \rightarrow p(H1) = 34.43\%$$

$t=2$ : Precautionary action. Institution takes D1 (provisional suppression) at 34.43% confidence—well below the 70% hard threshold. Potentially legitimate. The problem is the recording.

$t=3$ : Frame attached post-decision. F1 (“coordinated manipulation”), no new receipts, frame multiplier = 2.0.

$$0.525 \times 2.0 = 1.05 \rightarrow p(H1) = 51.22\%$$

Laundering begins. The action starts to look more reasonable in retrospect.

$t=4$ : Real corroborating evidence arrives. R3 (LR = 2.2), R4 (LR = 1.6).

$$1.05 \times 2.2 \times 1.6 = 3.696 \rightarrow p(H1) = 78.71\%$$

Legitimate smoothing. The case has genuinely strengthened.

$t=5$ : Policy drifts from P1 to P2. Hard threshold drops 0.75 → 0.65. Source-cluster prior adds +0.07. Penalty curve steepens. The receipt-only posterior remains 64.89%; the frame-adjusted interpretation posterior remains 78.71%. Neither changed—but severity escalates from temporary suppression to account strike plus distribution lock plus visibility penalty. Same facts, same frames, harsher compiler.

$t=6$ : Summary layer generates: “High-confidence coordinated manipulation detected based on multiple corroborating signals. Enforcement action taken.”

That sentence is not exactly false. It is a masterpiece of laundering.

### What the summary erases

- The precautionary action at 34% confidence
- The frame attachment that preceded strong evidence
- The policy drift that changed penalty severity
- The distinction between evidence posterior and policy decision logic
- The sequence: action → frame → evidence → policy hardening → summary

The public impression: evidence accumulated → confidence got high → action followed. That's not a lie. It's sequence laundering.

### The five planes, separated

Plane	Contents	Values over time
Receipt	What receipts support, R only	34.43% (t1) → 64.89% (t4), unchanged at t5
Interpretation	Receipt + frame under policy	34.43% → 51.22% (after F1) → 78.71%
Policy	What counts and how action maps	P1 → P2 at t5, thresholds and priors change
Execution	What the institution did	D1 provisional (t2) → D2 formal (t4) → D3 escalated (t5)
Summary	What users get told	One sentence implying linear evidence-led inevitability

### What the governor forces

At  $t=2$ : a decision-time justification receipt—posterior 34.43%, exception path used (precautionary/high-impact), review deadline set.

At  $t=3$ : a typed frame receipt for F1—who attached it, when, under which policy, citing which receipts—making visible that the frame arrived after the decision.

At  $t=4$ : standard observation and provenance receipts for R3 and R4.

At  $t=5$ : a policy-update receipt—threshold change, added priors, curve steepening, effective date, approver, rollback path.

At  $t=6$ : a **summary receipt**—which source nodes were used, which omitted, which compression policy applied, as-of timestamp. The summary layer can no longer present itself as pure narration. It is a typed transformation with provenance.

*The full pathology is sequence laundering across layers: institutions act provisionally under uncertainty, later evidence and frames strengthen the case, policy drift changes severity, and a summary layer compresses the timeline into retrospective inevitability. A governor does not prevent any of these updates; it prevents hidden edges.*

---

### 6.5 Metrics That Fall Out of the Examples

**Posterior shift attribution** decomposes belief change by factor class—what fraction came from new receipt nodes versus frame nodes versus policy nodes. High frame/policy attribution without new receipts is the primary propaganda signal.

**Frame-before-evidence influence** measures posterior movement attributable to frame attachment before supporting receipts arrive. Catches the backfill pattern at the earliest stage.

**Decision-to-evidence lag** tracks how often decisions are timestamped before supporting observation receipts exist. Not inherently bad—precautionary action is sometimes legitimate—but must be flagged provisional and recorded honestly.

**Policy drift counterfactual** re-evaluates each decision under current versus prior policy version with the same receipt set. If outcome changes, that change is attributable to policy drift, not evidence. Most institutions would find this metric uncomfortable, which is why it is useful.

**Sequence compression loss** measures how much chronology was flattened in the summary layer—specifically whether a summary implies evidence-first when the actual sequence was action-first.

**Summary provenance completeness** tracks what fraction of summaries include decision timing, policy version, frame timing, and as-of state. Most current systems would score near zero.

---

## 7. Implementation and Failure Modes

### 7.1 Receipt Schema (Illustrative)

The governor architecture requires typed receipt schemas for each node class. The following illustrates the key fields; specific implementations will vary.

**Decision receipt:**

```
{  
    "type": "decision",  
    "action": "string",  
    "actor": "string",  
    "timestamp": "ISO8601",  
    "policy_version": "string",  
    "evidence_posterior": "float",  
    "threshold": "float",  
    "threshold_met": "bool",  
    "exception_path": "string | null",  
    "review_deadline": "ISO8601 | null"  
}
```

**Frame receipt:**

```
{  
    "type": "frame",  
    "label": "string",  
    "actor": "string",  
    "timestamp": "ISO8601",  
    "policy_version": "string",  
    "cited_receipts": ["receipt_id"],  
    "confidence": "float",  
    "scope": "string",
```

```

    "supersedes": "frame_id | null"
}

```

**Policy update receipt:**

```

{
  "type": "policy_update",
  "policy_id": "string",
  "version": "string",
  "prior_version": "string",
  "changes": {
    "threshold_delta": "float",
    "prior_adjustments": "object",
    "other": "..."
  },
  "effective_time": "ISO8601",
  "approver": "string",
  "rationale": "string",
  "affected_scope": "string",
  "rollback_path": "string"
}

```

**Summary receipt:**

```

{
  "type": "summary",
  "source_nodes": ["node_id"],
  "omitted_nodes": ["node_id"],
  "compression_policy": "string",
  "as_of": "ISO8601",
  "omission_classes": ["string"]
}

```

## 7.2 Known Failure Modes

**Capture of the type system.** Whoever defines the node and factor type schema exerts upstream control over what can be represented [13]. Inconvenient artifacts can be excluded by schema rather than by visible suppression. The type system itself must be versioned, receipted, and contestable. A regime that doesn’t want “policy drift” to be a queryable category simply doesn’t type it. The fight over the schema is the fight over the constitution. This is the capture regime from Paper 16 [6] applied to the graph’s own ontology: high throughput and binding authority over schema definition, with fidelity to the full representational space destroyed by exclusion.

**Compliance theater.** Institutions can implement governor-compliant logging while ensuring records are practically inaccessible—stored but not indexed; present but not queryable. The governor requires not just that records exist but that they are legibly retrievable by parties with standing to challenge them.

**Framing authority monopoly.** Even with receipted frame nodes, if framing authority is concentrated, the graph can be steered by attaching well-receipted but strategically selected frames. The governor requires that framing authority be distributed or that its concentration be legible and bounded.

**Speed asymmetry.** Governed systems with explicit receipt requirements are slower than ungoverned systems. Propaganda can exploit the lag between event, receipt, and legible frame. This suggests governor architectures will be most effective in retrospective and appellate contexts rather than real-time moderation.

---

## 8. Related Work

**Filtering, smoothing, and factor graphs.** The formal backbone of the model draws on Kalman filtering [7], Rauch-Tung-Striebel smoothing [8], and factor graph inference [9]. The Bayesian belief revision structure is standard. The contribution is not the mathematics but the identification of the *policy layer* as a first-class graph object rather than an implicit parameter of the inference procedure. When the policy layer is unmodeled, the inference machinery cannot distinguish evidence-driven updates from policy-driven updates. That confusion is the attack surface.

**Propaganda studies and framing theory.** Entman’s framing analysis [10] establishes that selection and salience are the primary mechanisms of media influence. McCombs and Shaw’s agenda-setting research [11] demonstrates that media power operates through topic selection rather than persuasion on topics already salient. Herman and Chomsky’s propaganda model [12] identifies institutional filters — ownership, advertising, sourcing, flak, ideology — that shape media output upstream of individual editorial decisions. The present framework extends all three by formalizing frames as typed graph objects with provenance, and by identifying the policy layer (admissibility, indexing, trust weighting) as the primary attack surface rather than content. Entman’s “selection and salience” becomes adversarial control over policy factors; Herman and Chomsky’s institutional filters become typed policy nodes whose versioning is auditable.

**Science and technology studies / classification and power.** Bowker and Star [13] document how classification systems, indexing categories, and administrative schemas function as power structures — determining what can be represented, retrieved, and contested. Their analysis of the International Classification of Diseases demonstrates that category boundaries are political decisions with material consequences, not neutral descriptions. The present framework formalizes this insight directly: the type system governing the semantic factor graph is itself a received, versioned, contestable object. Capture of the type system (§7.2) is the formal version of Bowker and Star’s central observation that invisible infrastructure exerts invisible power.

**Content authenticity and provenance.** The C2PA standard [14] addresses provenance at the object level: cryptographic attestation of where a file came from and what transformations it underwent. This is necessary infrastructure — receipt plumbing — but insufficient. Object-level provenance answers “who created this artifact?” It does not answer “why was this artifact surfaced, weighted, indexed, and framed the way it was?” The governor architecture extends provenance from objects to the full policy stack governing their interpretation.

**Decentralized publishing and federated moderation.** Architectures such as ATProto and ActivityPub address the monopoly condition by distributing authority across independent servers. This is necessary but not sufficient. Distribution without legible update rules produces distributed opacity rather than legible epistemics — each node can run its own invisible policy stack. The governor architecture is complementary: it specifies *what* must be receipted regardless of *where* the authority is located.

**Author’s prior work.** The  $\Delta t$  framework [1–6] developed the temporal coherence analysis that motivates the smoothing formulation here: systems fail when different operational layers become temporally misaligned, and the misalignment is often invisible without explicit layer separation. Paper 3 [1] proved that

scalar optimization destroys eigenstructure — the formal basis for the capture mechanism in propaganda (policy-layer optimization that collapses representational diversity). Paper 12 [4] developed the governor architecture with typed receipts, the Non-Linguistic Authority Invariant (language proposes, evidence commits), and the contradiction ledger. Paper 15 [5] generalized the governor pattern across nine domains and introduced the commitment boundary formalism. Paper 16 [6] introduced the three-regime framework (shear, leverage, capture) and the correlator quality vector  $K = (T, F, A, C)$ . The present paper applies these to public memory: the receipt graph is the contradiction ledger scaled to institutional discourse; propaganda is a capture-mode correlator (high  $T$ , high  $A$ , low  $F$ ) operating on public posteriors; and the epistemic governor is the same architectural pattern — separate proposal from commitment, gate crossings on evidence — applied to belief revision rather than LLM runtime.

---

## 9. Conclusion

We began with a compressed formulation: propaganda is semantic retrocausality with a functional monopoly on legibility at scale. We have unpacked this into a formal model, a pathology taxonomy, and four worked examples.

Semantic retrocausality is not a pathology. It is how any sufficiently capable belief-revision system must work across time. Later information updates earlier estimates. This is smoothing, not manipulation.

Propaganda is distinguished not by the presence of retrospective reinterpretation but by the conditions under which it occurs: hidden control over the legitimacy criteria for belief revision—the policy-layer rules governing what counts as a valid reason to update public memory. The attack surface is not the data. It is the transition logic.

The epistemic governor is a set of constraints requiring that all semantic updates produce typed, auditable receipts. It does not prevent retrospective reinterpretation. It forces post-selection to leave tracks.

The counter-propaganda move is not better receipts. It is not more evidence. It is legible epistemic policy: a system in which the rules governing what counts as a valid reason to update are explicit, typed, versioned, and challengeable. This is a more demanding and more politically contentious requirement than “stop lying.” Lying is a character flaw. Illegible epistemic policy is a feature—of systems that benefit from opaque governance. The demand for legible epistemic policy is a legitimacy claim. It asserts that the rules for changing the rules must be subject to challenge. That is constitutionalism applied to memory infrastructure.

*Receipt the compiler.*

---

## Notes on Method

This paper emerged from a multi-model analytical dialogue conducted in February 2026. The models involved—Claude (Anthropic), ChatGPT (OpenAI), DeepSeek, and Gemini (Google)—were used as collaborative reasoning partners rather than citation sources. Convergent framings across models were treated as indicators of robustness; divergences were used to locate analytical fault lines. The approach is described in prior work as interferometry: cross-validating analytical conclusions across systems with different architectures and training distributions.

The formalism in Section 3 builds on the author’s prior work on the Epistemic Governor and the  $\Delta t$  framework, both available in the papers repository. The factor graph formulation is preferred over the state-space version for reasons of provenance clarity: it makes the policy layer explicit as a first-class graph object rather than an implicit parameter of the transition function.

The worked examples in Section 6 were developed collaboratively with ChatGPT, which contributed the Bayesian odds arithmetic and the pathology taxonomy. The numerical values are pedagogically chosen and should not be interpreted as calibrated estimates for any specific domain.

The summary receipt construct—treating summaries as first-class derived artifacts with typed provenance—was not present in either prior framework and emerged during this dialogue. It is the most novel operational claim of the paper and the most directly applicable to current platform governance.

---

## References

- [1] Beck, J. (2025). Scalar Reward Collapse: A General Theory of Eigenstructure Evaporation in Closed-Loop Systems. Preprint,  $\Delta t$  Framework Paper 3. doi:10.5281/zenodo.17791872
- [2] Beck, J. (2025).  $\Delta t$ -Constrained Inference: A General Model of Temporal Coherence in Hierarchical Systems. Preprint,  $\Delta t$  Framework Paper 7. doi:10.5281/zenodo.17857541
- [3] Beck, J. (2026). Representational Invariance and the Observer Problem in Language Model Alignment. Preprint,  $\Delta t$  Framework Paper 11. doi:10.5281/zenodo.18071264
- [4] Beck, J. (2026). Bounded Lattice Inference: A Governed Reasoning Substrate with Persistent State and Non-Linguistic Authority. Preprint,  $\Delta t$  Framework Paper 12. doi:10.5281/zenodo.18145346
- [5] Beck, J. (2026). Cybernetic Fault Domains: When Commitment Outruns Verification. Preprint,  $\Delta t$  Framework Paper 15. doi:10.5281/zenodo.18686130
- [6] Beck, J. (2026). The Gain Geometry of Temporal Mismatch: Shear, Leverage, and Capture in Multi-Timescale Systems. Preprint,  $\Delta t$  Framework Paper 16. doi:10.5281/zenodo.18717619
- [7] Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering* 82(1), pp. 35–45.
- [8] Rauch, H.E., Tung, F., and Striebel, C.T. (1965). Maximum Likelihood Estimates of Linear Dynamic Systems. *AIAA Journal* 3(8), pp. 1445–1450.
- [9] Kschischang, F.R., Frey, B.J., and Loeliger, H.-A. (2001). Factor Graphs and the Sum-Product Algorithm. *IEEE Transactions on Information Theory* 47(2), pp. 498–519.
- [10] Entman, R.M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication* 43(4), pp. 51–58.
- [11] McCombs, M.E. and Shaw, D.L. (1972). The Agenda-Setting Function of Mass Media. *Public Opinion Quarterly* 36(2), pp. 176–187.
- [12] Herman, E.S. and Chomsky, N. (1988). *Manufacturing Consent: The Political Economy of the Mass Media*. Pantheon Books.
- [13] Bowker, G.C. and Star, S.L. (1999). *Sorting Things Out: Classification and Its Consequences*. MIT Press.

[14] Coalition for Content Provenance and Authenticity (C2PA). (2022). C2PA Technical Specification.  
<https://c2pa.org/specifications/>