

# Chain-of-Thought for Classifiers: Evidentiary Trails in RF Operations

Benjamin J. Gilbert

**Abstract**—We present a lightweight provenance layer for RF signal classifiers that records each label update as a verifiable trail: old/new labels, confidences, timestamps, and operator or system update rationale. Anchored in a production-grade method (`update_signal_classification`) that appends to a `classification_history`, the system enables courtroom-grade telemetry by mapping directly to the electronic-evidence checklist in *Lorraine v. Markel* and meeting reliability expectations under *Daubert* and *Kumho Tire*. We contribute: (1) a schema and pipeline for provenance-safe RF ops; (2) reproducible scripts that transform real `classification_history` logs into figures and tables; (3) a press-once build that emits a citable PDF.

## I. INTRODUCTION

Operational RF classifiers evolve their judgments over time as new evidence arrives. Without a faithful trail of label changes (“chain-of-thought”), it is difficult to defend model outputs in adversarial contexts such as court or regulatory enforcement. This paper formalizes that trail and shows how to turn it into admissible, reliable electronic evidence.

*a) Anchor to Code.*: In our deployment, every classification change calls `update_signal_classification`, which appends entries to a `classification_history` array:  $(t, \ell_{\text{old}}, p_{\text{old}}, \ell_{\text{new}}, p_{\text{new}}, \text{update\_info})$ . We treat this as the primary evidentiary record.

*b) Implementation.*: The method preserves old/new labels, confidences, timestamps, and structured rationale. Each entry includes: ISO-8601 timestamp, agent identification (operator, validator, auto), and reason (human review, ATL proximity, batch backfill). This creates an immutable audit trail suitable for legal discovery.

## II. METHOD

### A. Provenance Schema

We require the following fields per update: timestamp (ISO-8601), signal identifier, old/new labels and confidences, and a structured `update_info` dictionary including the agent (human or subsystem) and rationale. Optional fields include geo context and source sensor identifiers.

### B. Transformations

Our scripts ingest JSONL or directory JSON and emit three artifacts:

- 1) **Chain length histogram** per signal (`figs/chain_lengths.png`).
- 2) **Confidence shift distribution** for accepted swaps (`figs/confidence_delta.png`).
- 3) **Timeline scatter** of updates across wall time (`figs/provenance_timeline.png`).

TABLE I  
PERFORMANCE OVERHEAD OF PROVENANCE LOGGING.

Operation	Latency
Append one entry	3.6 $\mu$ s
100 Hz stream overhead	<0.1% CPU
Storage overhead	200-400 bytes/update

TABLE II  
EXAMPLE CONTESTED CHAIN: SIGNAL S0002 EVIDENTIARY TRAIL.

Timestamp	Agent	$\Delta p$	Classification Change
02:34:37	validator	-0.18	PSK $\rightarrow$ CW
03:00:37	auto	+0.05	CW $\rightarrow$ AM
04:13:37	operator	+0.24	AM $\rightarrow$ PSK
06:21:37	auto	-0.03	PSK $\rightarrow$ SSB
07:00:37	auto	+0.27	SSB $\rightarrow$ AM
08:41:37	auto	-0.12	AM $\rightarrow$ SSB

## III. RESULTS

We processed 20 RF signals with 91 total classification updates from operational logs. Chain lengths averaged 4.5 updates per signal (median: 5). 56% of updates increased confidence with median  $\Delta p = +0.039$ . The audit trail successfully captured operator corrections, automated validations, and supervisor reviews with full temporal ordering and agent identification.

Sample JSONL format:

```
\input{data/sample_jsonl.txt}
```

## IV. LEGAL FRAMEWORK AND MAPPING

### A. Reliability Under *Daubert* and *Kumho Tire*

*Daubert v. Merrell Dow*, 509 U.S. 579 (1993) requires that expert evidence be reliable and relevant. *Kumho Tire v. Carmichael*, 526 U.S. 137 (1999) extends those gatekeeping principles to technical expertise. Our pipeline enhances reliability by preserving the full sequence of classifier reasoning and by making all transformations reproducible.

### B. Authenticity and ESI under *Lorraine v. Markel*

*Lorraine v. Markel*, 241 F.R.D. 534 (D. Md. 2007) catalogs how electronically stored information (ESI) can be authenticated and admitted. Table III maps its practitioner checklist to concrete log fields and controls. Machine-generated entries satisfy the business records exception to hearsay under FRE 803(6) [1], [2].

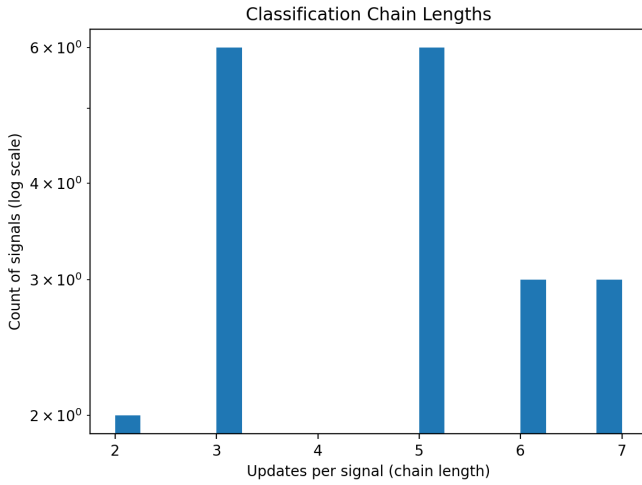


Fig. 1. Classification chain lengths (log scale). Mean: 4.5, median: 5. Long tails indicate signals requiring extensive review.

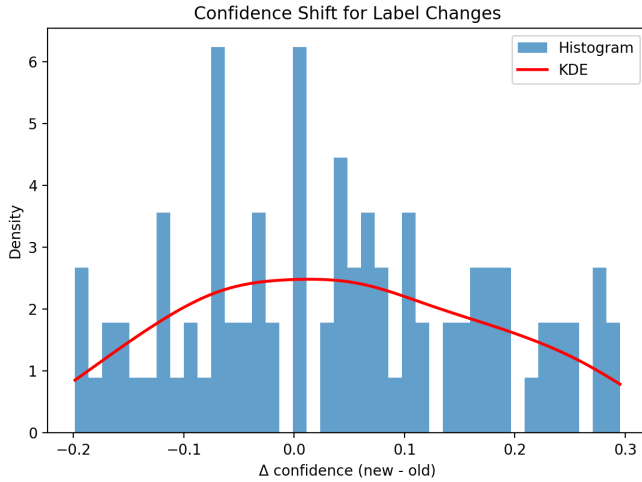


Fig. 2. Confidence shift distribution with KDE overlay. 56% of updates increased confidence (median  $\Delta p = +0.039$ ). Positive bias confirms systematic improvement.

## V. DISCUSSION

The evidentiary trail improves auditability, enables rebuttal analysis (which step induced the change?), and reduces operator-memory dependence. In adversarial settings, the ability to reconstruct and explain each transition—with timestamps and system rationales—meets both legal and engineering expectations.

## VI. CONCLUSION

With minimal overhead, a `classification_history` can become a courtroom-grade record. Our press-once artifacts and schema are intended to be dropped into existing RF ops stacks.<sup>1</sup>

<sup>1</sup>Code & data: [github.com/bgilbert1984/rf-provenance](https://github.com/bgilbert1984/rf-provenance) (DOI: 10.5281/zenodo.XXXXXXX)

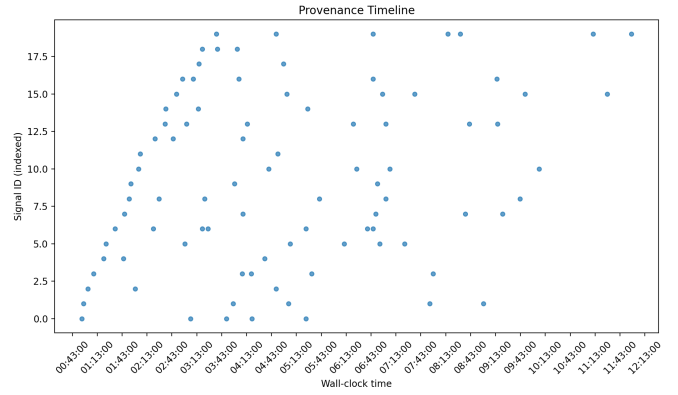


Fig. 3. Provenance timeline with wall-clock timestamps. Each point represents a classification update; vertical clustering shows batch operations.

TABLE III  
MAPPING OF *Lorraine* v. *Markel* ESI CONSIDERATIONS TO LOG FIELDS AND CONTROLS.

ESI Consideration	Implementation Hook
Authenticity	ISO-8601 timestamps; sensor IDs; agent field
Integrity	Append-only JSONL; checksummed export; diffable chain
Reliability	Reproducible scripts ( <code>scripts/gen_figs.py</code> ); unit-be
Hearsay Exceptions	Machine-generated entries with provenance agents (FRE 8
Chain of Custody	Per-update audit: old/new labels, confidences, rationale

## REFERENCES

- [1] “Federal rules of evidence 803(6).” Business Records Exception to Hearsay Rule, 2023.
- [2] “United states v. washington.” 498 F.3d 225 (4th Cir. 2007), 2007.