Federal Minimum Learning & Integrity Standard (FMLIS)

A Standards-Grade Governance Framework for Constitutional Reliability

Christine K. Hillier
Kapukai – Governance Lab

2025-12-16

MEMORY Non-erasable, auditable record integrity: prior placements, outcomes, and adverse events must remain visible, traceable, and correctable.

LEARNING Closed feedback loops with measurable triggers: error thresholds, reversals, and adverse patterns must generate logged remediation and deployed safeguards.

INSULATION Redaction-first, identity-minimized decision pathways: double-blind review options, role separation, and independent audit to prevent bias, capture, and perverse incentives.

Contents

1 Executive Summary

FMLIS defines minimum learning and integrity requirements for government and institutional decision systems that can produce severe harm when they fail. The standard is written to be implementable as a governance framework with a reference architecture and test suite: each requirement is measurable, auditable, and designed to reduce constitutional risk and civil-rights exposure by design.

Three Non-Negotiables

- **Memory:** the system must retain accurate, non-erasable records of prior placements, outcomes, and adverse events so failures remain visible and correctable.
- **Learning:** the system must update procedures and safeguards based on outcomes, error rates, and validated adverse-event patterns (closed feedback loops).
- **Insulation:** decision processes must be insulated from identity bias, corruption, and perverse incentives through redaction-first intake, role separation, and independent auditability.

What FMLIS Produces

- A minimum control catalog (**FMLIS-001**–**FMLIS-099** initial draft range).
- A test matrix specifying evidence required to pass controls.
- Procurement language agencies can adopt immediately.
- A phased implementation roadmap for municipal pilots.

2 Problem Statement

High-stakes systems frequently optimize for throughput, administrative closure, and short-term liability avoidance rather than truth, correction, and human safety. When records can be overwritten, evidence cannot be reliably submitted, and adverse events are not retained as durable learning signals, the system compounds error and harms vulnerable populations.

Engineering Claim (falsifiable)

A decision system that lacks: (i) durable record integrity, (ii) evidence intake with receipts, and (iii) correction/appeal loops with measurable SLAs, cannot reliably satisfy constitutional due-process expectations *as applied* in high-risk contexts, because the risk of error becomes both high and persistent.

Goal

Define minimum controls so these systems *learn*, *retain truth with integrity*, and *insulate decision pathways* from corruption and bias.

3 Constitutional and Civil-Rights Basis (Non-Adjudicative)

This section states the governance rationale for minimum safeguards without adjudicating any particular case.

Procedural Due Process (conceptual mapping)

At minimum, high-stakes deprivation systems require: notice, meaningful opportunity to respond, neutral decision pathways, and correction mechanisms commensurate with the risk of error. FMLIS translates these into implementable controls: evidence receipts, record of decision provenance, correction SLAs, and independent audit trails.

Civil-Rights Risk Reduction (conceptual mapping)

Systems that cannot preserve and correct records predictably externalize harm and create repeatable rights-deprivation patterns. Controls here are designed to reduce: arbitrary action risk, discriminatory effects, retaliation risk, and denial-of-access-to-process risk, by making the system auditable and correctable.

4 System Failure Modes (Threat Model)

Observed failure modes in collapsed systems

- **Memory erasure:** adverse events disappear; prior placements/outcomes are not durable.
- **Non-learning loops:** the system repeats known failure patterns; no update triggers.
- **Identity capture:** identity and location become attack surfaces (bias, corruption, targeting).
- **Evidence blockade:** no receipt, no routing ID, no chain-of-custody; submissions vanish.
- **Opaque decisions:** no record-of-decision linking inputs/rules/operators/timestamps.
- **Perverse incentives:** incentives favor removals, closures, or churn over outcomes and safety.

Design stance

FMLIS treats these as engineering defects with measurable mitigations, not as moral narratives.

5 Minimum Learning Standards (Controls Overview)

This section defines learning requirements as controls.

Core controls (starter set)

- **FMLIS-001** Learning Triggering: the system **SHALL** define measurable triggers (error thresholds, adverse-event patterns, reversals, sustained appeals).
- **FMLIS-002** Feedback Closure: each trigger **SHALL** produce a documented change proposal, review outcome, and deployment record.
- **FMLIS-003** Outcome Visibility: placements/outcomes/adverse events **SHALL** be retained as non-erasable learning signals.
- **FMLIS-004** Corrective Action SLA: correction requests **SHALL** have measurable timelines and dispositions.

Control writing rule

Each control must include: requirement, rationale, evidence required, pass/fail criteria, and severity classification.

6 Memory and Record Integrity

Principle

If the system can change a person's life, the system must be able to prove what it did, why it did it, and how to correct it.

Minimum requirements

- **FMLIS-010** Non-erasable Adverse Event Log: adverse events **SHALL** be append-only with immutable audit trail.
- **FMLIS-011** Record of Decision (RoD): every adverse action **SHALL** generate a RoD linking inputs, rules/policies, operator role, timestamps, and outcome.
- **FMLIS-012** Provenance: records **SHALL** retain source, time, and chain-of-custody metadata.
- **FMLIS-013** Litigation Hold Ready: retention **SHALL** support holds and independent export.

Implementation note

"Non-erasable" does not require a public blockchain; it requires immutability under audit, with controlled access and export.

7 Identity, Redaction, and Insulation

Redaction-first intake

- **FMLIS-020** Immediate Redaction: intake **SHALL** default to redaction/anonymization before routing.
- **FMLIS-021** User-Owned Copy: the user **SHALL** receive a receipt and a copy/hash of what was submitted.
- **FMLIS-022** Pseudonymous Case Identifier: the system **SHALL** assign a non-human-readable ID for tracking without exposing identity broadly.

Insulation goal

Reduce bias, corruption, and targeting by separating identity from evaluation pathways wherever possible, while preserving lawful due process and authorized access where required.

8 Double-Blind Governance (Decentralized Expert Review)

Model

A governance operating system where review of system behavior is performed by verified experts through auditable, insurable processes rather than political influence.

Minimum requirements

- **FMLIS-030** Role Separation: intake, evaluation, and audit **SHALL** be separated to prevent capture.
- **FMLIS-031** Verified Expert Pool: reviewers **SHALL** meet track-record criteria and independence requirements.
- **FMLIS-032** Change Board: material changes **SHALL** be reviewed via an engineering change board with logged votes and rationales.
- **FMLIS-033** Double-Blind Review Option: evaluation **MAY** occur with identity/location withheld unless legally necessary.

Note

This section defines governance controls; it does not disclose proprietary selection algorithms or operational IP.

9 Insurance and Accountability

Principle

If a system cannot be insured against foreseeable harms, it is not safe enough for high-stakes decisions.

Minimum requirements

- **FMLIS-040** Insurability Evidence: system operators **SHALL** maintain evidence required for underwriting (audit logs, error metrics, corrective actions).
- **FMLIS-041** Independent Audit Access: auditors **SHALL** have read access to required logs under controlled terms.
- **FMLIS-042** Retaliation Barrier: use of correction/appeal channels **SHALL** not trigger adverse classification.

10 Transparency Pool (Optional, Consent-Based)

Concept

Users may opt into a redacted/anonymized transparency pool that enables independent engineering review of system performance without exposing identity.

Minimum requirements

- **FMLIS-050** Explicit Opt-In: the pool **SHALL** be consent-based with clear scope and revocation terms.
- **FMLIS-051** Redaction Baseline: pool entries **SHALL** be redacted before inclusion.
- **FMLIS-052** Public Utility Orientation: pool governance **SHOULD** be insulated from perverse incentives and conflicts.

11 Reference Architecture (Non-Proprietary)

Components

- Intake Gateway (redaction-first, receipt, routing ID)
- Record Store (append-only logs, provenance)
- Decision Service (RoD generator, policy linkage)
- Correction & Appeal Service (SLA tracking, dispositions)
- Audit Interface (independent export + inspection)
- Metrics Service (error thresholds, triggers)

Design rule

No adverse action without a Record of Decision and a verifiable audit trail.

12 Metrics and Tests

Minimum metrics

- Time-to-receipt for evidence submissions (median, p95)
- Percentage of adverse actions with complete RoD
- Time-to-correction disposition (median, p95)
- Adverse-event retention completeness rate
- Audit log completeness and tamper-evidence status

Test suite concept

Each control has a test: artifacts required, sampling method, pass/fail threshold, and severity.

13 Procurement Language (Immediate Leverage)

Vendor minimum clauses (starter set)

- Vendor **SHALL** provide exportable audit logs and RoD artifacts.
- Vendor **SHALL** support evidence intake receipts and routing IDs.
- Vendor **SHALL** support correction workflows with measurable SLAs.
- Vendor **SHALL** not deploy black-box adverse decisioning without explainable RoD artifacts.
- Vendor **SHALL** support independent audit under controlled access.

Purpose

Make constitutional reliability a purchasing requirement, not a lawsuit afterthought.

14 Implementation Roadmap (Municipal Pilot)

Phase 0 (0–2 weeks): Standard Draft + Test Harness

Deliver: compiled PDF standard, control catalog v0.1, test matrix v0.1, procurement clause set.

Phase 1 (2–6 weeks): Evidence Intake + Receipt Prototype

Deliver: intake receipt template, routing ID workflow, append-only log proof, basic metrics.

Phase 2 (6–12 weeks): Correction Loop + RoD Prototype

Deliver: RoD schema, correction SLA tracking, audit export, minimal dashboard.

Phase 3 (12–24 weeks): Double-Blind Review + Insurance Readiness

Deliver: reviewer process, change board logs, underwriting evidence pack, independent audit procedures.

Phase 4 (6–12 months): Scale + Adoption

Deliver: procurement adoption kit, training/certification, multi-department onboarding playbook.

15 Appendix: Glossary

- **RoD (Record of Decision):** A traceable artifact that explains an adverse action with inputs, rules/policies, operator role, timestamps, and outputs.
- **Append-only log:** A record that can be extended but not rewritten without detection.
- **Redaction-first:** Identity is minimized before routing and review unless legally required.
- **Double-blind review:** Reviewers assess system behavior with identity/location withheld when feasible.
- **Insurability:** Ability to underwrite system operation based on evidence of controls, metrics, and remediation.