

PKOS v1 — A Constitutional Membrane for Accountable AI Execution

A structural response to justificatory collapse under scale

Abstract

As artificial intelligence systems increasingly participate in decision-making, a structural failure has emerged: actions persist, but their justifications do not.

PKOS (Personal Knowledge OS) proposes a constitutional substrate for AI–human collaboration. It introduces an enforceable membrane between intention and execution, ensuring that no action with real-world consequence can occur without a complete, auditable justification.

Authority must be computable.

The Crisis: Justification Collapse Under Scale

Modern AI systems operate probabilistically. They predict tokens based on patterns across large datasets.

They do not retain durable internal intention or jurisdiction.

Structural failures include:

- Justificatory Amnesia
- Context Dissolution
- Hallucinated Authority
- Accountability Void
- Post-Action Fragility

These are architectural failures — failures to enforce linkage between intention, validation, execution, and memory.

Design Principle

No execution is permitted without a validated justification bundle.

Execution is structurally separated from validation and memory.
Authority cannot emerge without justification.

Bi-Hemispheric Architecture

Body (Execution Layer): The only interface to external tools and actions.

Soul (Validation Layer): Evaluates whether an action is lawful within declared intent and authority scope.

Ledger (Binding Layer): Append-only system binding intention, validation, execution, and consequence.

The Justification Bundle

Every execution request must include:

1. Objective Declaration
2. Provenance Bundle
3. Delegation Proof (if applicable)
4. Semantic Anchors

No anchor, no authority.

Repair Doctrine

PKOS rejects deletion as a primary response to error.

Instead, Repair Receipts preserve causality:
original event → repair event → restored state

Memory remains the highest court.

Research Invitation

This paper invites critique and validation of:

- The enforceability of the execution membrane
- The completeness of the Justification Bundle model
- The semantics of append-only ledger binding
- The viability of repair-based accountability
- The Promotion Gate as a creativity-preserving governance mechanism

PKOS is presented as a testable constitutional engineering proposition.

Covenant

Authority must be computable.
Justification must persist.
Repair must be additive.
Memory must remain the highest court.