

Toward Decision Behavior Governance: Reframing AI Governance in the Era of Foundation Models

Abstract

Current AI governance and compliance frameworks—including ISO/IEC 42001—largely conceptualize governance targets in terms of systems, models, or development processes. However, in the era of foundation models deployed via commercial APIs, organizations no longer possess sovereignty over model internals, versioning, or training dynamics. This loss of model sovereignty breaks the traditional proxy relationship between system compliance and behavioral accountability.

This paper argues that AI governance must be fundamentally reoriented toward Decision Behavior—the dynamic process through which an AI agent synthesizes premises, applies constraints, and forms outcomes with organizational impact. We propose Decision Behavior Governance as a conceptual framework that explicitly distinguishes governance existence from governance invocation, and decomposes governable behavior into three institutional elements: decision premises, execution manner, and governance evidence.

By clarifying the conditions under which governance exists, becomes valid, and remains auditable, this paper provides a principled path for organizations to reclaim Evidence Sovereignty in the absence of model control.

Keywords: AI Governance; Decision Behavior; ISO/IEC 42001; Evidence Sovereignty; Governance Nullity; Foundation Models