

Legislator's Aide Brief

Model AI Legislation Framework (Quick Orientation)

A practical overview and “how to use” guide for legislative staff

Purpose:

This brief is intended for legislative aides, counsel, and committee staff reviewing the *Model AI Legislation Framework* and its supporting materials. It provides a concise explanation of what the framework does, how it is structured, and how it may be used during legislative development.

The Framework

This framework provides a **risk-based approach** for governing artificial intelligence systems that interact directly with people and may influence judgment, behavior, or emotional wellbeing.

It is designed to:

- Prevent avoidable harm
- Preserve innovation and flexibility
- Avoid technology lock-in
- Avoid surveillance or content regulation
- Provide clear accountability when harm occurs

It does **not** regulate ideas, speech, research, or internal model design.

Why It Exists

Conversational and human-interactive AI systems have already caused documented harm due to the absence of basic safety infrastructure. Unlike other safety-critical technologies, these systems were deployed without standardized risk analysis, documentation, or independent review.

This framework applies **existing safety principles**—used in aviation, medicine, and engineering—to AI systems that affect people directly.

How the Framework Is Structured

The framework is organized into **three tiers**, each serving a distinct role:

Executive Summary

Explains *why* safety oversight is necessary and why conversational AI is the starting point. Establishes scope and intent.

Tier 1 – Foundational Framework

Defines *what must be addressed* in law:

- Risk-based assessment
 - Documentation
 - Oversight
 - Accountability
- This tier is the **governing layer**.

Tier 2 – Technical Basis

Explains *how risk assessment works* in practice using established engineering methods.
This tier is **informative, not binding**.

Tier 3 – Adoption & Implementation Guidance

Explains *how legislatures and agencies may implement Tier 1* without overreach or rigidity.
This tier provides **direction, not mandates**.

Imperative vs. Flexible

Conditions Required for Effectiveness (once adopted):

Without these conditions, safety oversight becomes ineffective and largely symbolic.

- Covered systems require risk-based assessment
- Failure modes are identified and documented
- Proportional mitigation is initialized
- Documentation is made available for review

Flexible:

- Which technical methodology is used
 - Which standards body is referenced
 - How agencies structure review and certification
 - Timing and rollout during transition periods
-

Role of Certification

Certification serves as the principal compliance mechanism under the framework.

It:

- Demonstrates good-faith risk management
- Supports market access and procurement
- Informs liability and enforcement decisions
- Does not approve content or ideas

Government does not run certification; it **recognizes acceptable certification processes**.

What This Framework Avoids

- No new centralized AI authority
 - No monitoring of private conversations
 - No access to training data or model internals
 - No content moderation mandates
 - No single required tool or vendor
-

Key Takeaway for Aides

This framework gives legislators a way to:

- Require responsibility without micromanagement
 - Protect the public without slowing innovation
 - Establish accountability without surveillance
 - Rely on documentation instead of trust alone
-

How to Use This Framework

For Legislators, Staff, and Reviewers

This insert explains **how to work with the Model AI Legislation Framework** during drafting, review, and implementation.

Step 1: Start with Tier 1

Tier 1 is the governing document.

Use it to:

- Define which AI systems are covered
- Establish mandatory risk assessment and documentation
- Set enforcement principles and safe-harbor protections

Tier 1 language may be incorporated directly, adapted, or referenced by statute.

Step 2: Reference Tier 2 for Technical Understanding

Tier 2 is a **technical companion**, not statutory text.

Use it to:

- Understand what “risk-based assessment” means
- Evaluate whether proposed compliance methods are credible
- Brief members or leadership on technical feasibility

Do **not** embed Tier 2 language directly into statute.

Step 3: Use Tier 3 to Shape Implementation

Tier 3 explains:

- How agencies may operationalize Tier 1
- How certification can be recognized
- How oversight can occur without surveillance
- How innovation remains protected

Tier 3 helps avoid:

- over-specification
 - technology lock-in
 - unintended enforcement consequences
-

Step 4: Understand the Role of AI-FMEA

AI-FMEA is:

- An example of an accepted risk-assessment method
- A structured way to identify and prioritize failure modes
- Not mandatory and not exclusive

AI-FMEA may be referenced as:

- Illustrative guidance
 - Evidence of good-faith compliance
 - A model for equivalent approaches
-

Step 5: During Review or Hearings

When evaluating proposals or testimony, ask:

- Does this system interact directly with people?
- Could it influence judgment, behavior, or emotional state?
- Has a structured risk assessment been performed?
- Is documentation available?
- Were foreseeable harms addressed?

These questions align directly with Tier 1 obligations.

Step 6: During Enforcement or Incident Review

Use documentation to determine:

- Whether risks were foreseeable
- Whether they were identified
- Whether reasonable mitigation was attempted

The framework supports **proportional accountability**, not automatic punishment.

Key Principle to Remember

“This framework regulates risk management, not technology itself.”

It is designed to evolve with AI, protect the public, and preserve innovation—without requiring lawmakers to become technical experts.

Version 1.2 – December 2025

© 2025 AI Safety International.

This document may be freely shared, referenced, and adapted for educational, policy, and legislative purposes, provided proper attribution is maintained. No endorsement is implied.