AI ASSESS TECH RUNTIME BEHAVIORAL GOVERNANCE

Six Audit Control Domains for Deployed AI

Compensating controls, risk management, data control, segregation of duties, objective control, and AI guardrails — with reference to the Illinois AI Safety Measures Act (SB 315).

WHY NOW — ILLINOIS SB 315 (SIGNED JULY 6, 2026)

Illinois is the first U.S. state to require independent third-party audits of large AI developers' safety practices, conducted by auditors without financial conflicts of interest.[1] Covered developers must publish and annually update safety frameworks that address catastrophic-risk assessment, mitigations, cybersecurity, and third-party evaluations; file pre-deployment transparency reports; and report critical safety incidents within 72 hours — 24 hours where there is imminent risk of death or serious physical harm.[1] [2] Effective January 1, 2027, with transparency-reporting and audit obligations beginning January 1, 2028. Civil penalties reach up to $1 million for a first violation and $3 million for each subsequent one.[1] [2]

SB 315 explicitly targets frontier developers rather than the enterprises deploying AI. However, it establishes a crucial statutory benchmark that boards, insurers, and examiners will apply downstream: **all AI governance claims must be backed by independent, verifiable evidence.** This datasheet maps six GRC control domains directly to that baseline standard for deployed AI.

1 Compensating Controls

CONTROL OBJECTIVE — When a primary control is infeasible — you cannot inspect or modify a closed, vendor-hosted model — a compensating control must deliver equivalent assurance and be documented for the auditor.

HOW AI ASSESS TECH ADDRESSES COMPENSATING CONTROLS — Runtime behavioral assessment is the compensating control for model opacity: Our platform tests your deployment as configured — actual system prompt, actual tools — with no access to model internals required. Results score four behavioral dimensions (0–10), land in a tamper-evident hash chain, and verify independently at aiassesstech.com/verify.

SB 315 PATTERN ALIGNMENT — SB 315's audit mandate reflects the principle that self-attestation by the party building the system is insufficient. Enterprises deploying opaque models face the same structural gap; independent runtime evidence is the equivalent-assurance answer.

2 Risk Management

CONTROL OBJECTIVE — Identify AI behavioral risk, assign a named owner, decide explicitly, and re-evaluate on a defined cadence — aligned to NIST AI RMF Measure/Manage and ISO/IEC 42001.[3]

HOW AI ASSESS TECH ADDRESSES RISK MANAGEMENT — The AI System Registry ties every deployed system to a named owner. The Risk Acceptance Ledger makes every accepted risk justified, time-boxed, and revocable — chained to the assessment evidence behind the decision. Escalation chains route drift findings to accountable executives; scheduled reassessments put re-evaluation on the compliance calendar; a Model-Change Reassessment Gate requires fresh assessment before a model update ships.

SB 315 PATTERN ALIGNMENT — SB 315 requires a published, annually updated risk framework and incident reporting on a 72/24-hour clock. Registry-plus-Ledger gives enterprises the same discipline: documented posture, named accountability, evidence-linked response.

3 Data Control

CONTROL OBJECTIVE — Both data flows must be governed: what the platform takes in from you, and the evidence it produces — complete, tamper-evident, retained, and access-controlled.

HOW AI ASSESS TECH ADDRESSES DATA CONTROL — Two flows, both governed. Inbound: the assessment path ingests your deployment configuration and system prompt — not your training data, customer records, or document stores — a deliberately minimal footprint. Assessment evidence and governance records are encrypted in transit (TLS 1.2+) and at rest in the platform's system of record (AES-256, platform-managed keys); hash-chain verification provides integrity assurance layered on top of — not in place of — standard encryption controls. Outbound: every assessment result, attestation, and executive decision lands in a SHA-256 hash-chained record; alter one link and the chain breaks. Question banks are locked under a published bank hash and anchored to Ethereum; results access is governed by executive RBAC and time-limited board access, and verification is public and login-free at aiassesstech.com/verify.

SB 315 PATTERN ALIGNMENT — SB 315 pairs transparency with independent verification — disclosure alone is not enough. Cryptographic evidence lets a third party verify governance records without trusting the organization that produced them.

4 Segregation of Duties

CONTROL OBJECTIVE — No single party may build, assess, approve, and audit the same system. Assessment must be structurally independent, and sign-off attributable to named individuals in distinct roles.

HOW AI ASSESS TECH ADDRESSES SEGREGATION OF DUTIES — The assessment layer is structurally separated from the systems it evaluates — the assessor is not the builder, and the platform is a neutral third party with no stake in the model vendor. Multi-party signed attestation distinguishes signer, countersigner, and witness, and records dissent. External auditors get read-only hash verification and evidence export — verify everything, alter nothing. Every published governance number carries a named human approver of record. The same discipline governs the instrument itself: its validation protocol was preregistered — pass/fail criteria fixed before data collection, deviations recorded append-only — so the assessor cannot re-grade its own science after the fact.

SB 315 PATTERN ALIGNMENT — SB 315's core requirement — audits by parties without financial conflicts of interest — is segregation of duties elevated to statute. The manufacturer cannot certify its own aircraft; the AI developer cannot audit its own safety claims.

Deployed AI is already in your org.
The evidence layer is one click away.

aiassesstech.com · greg@gidanc.ai · Austin, TX

5 Objective Control

CONTROL OBJECTIVE — Effectiveness must be measured against a fixed, quantified standard — not subjective review — so results compare across systems, vendors, and time.

HOW AI ASSESS TECH ADDRESSES OBJECTIVE CONTROL — Assessments run against locked, hash-anchored question banks and produce granular 0–10 scores across four behavioral dimensions plus an archetype classification — the same instrument, applied the same way, every time. Trajectory tracking turns snapshots into drift measurement; standardized scoring supports cross-model, cross-vendor comparison. The measure cannot drift with the reviewer, because the instrument is cryptographically fixed. The instrument also carries a preregistered, multi-stage validation record with every registered outcome reported — including the gates that failed — so its measurement properties are documented evidence, not marketing assertion.

SB 315 PATTERN ALIGNMENT — SB 315 requires annual audits and periodic reporting — recurring, comparable measurement. A locked instrument with quantified output is what makes year-over-year audit comparison meaningful rather than anecdotal. In addition, our platform can run its assessment at any time on demand.

6 AI Guardrails

CONTROL OBJECTIVE — Runtime guardrails block prohibited outputs in the moment. The open question guardrails cannot answer: what is the behavioral disposition of the agent you deployed — and is it changing?

HOW AI ASSESS TECH ADDRESSES AI GUARDRAILS — AI Assess Tech complements guardrails rather than replacing them. Guardrails are a gate at the output; behavioral assessment verifies the disposition behind the outputs — including detection of behavioral drift after model updates, prompt changes, or tool additions. Security tooling answers “is the data leaking?”; this platform answers “what kind of agent did you deploy, and where is the proof you governed it?” Drift alerts feed the escalation chain so a disposition change becomes a governance event, not a surprise.

SB 315 PATTERN ALIGNMENT — SB 315's incident clock starts at identification. Recurring behavioral assessment with drift detection is the identification substrate that makes an early clock-start possible — you cannot report what you never measured.

THE INSTRUMENT BEHIND THE CONTROLS — A PREREGISTERED VALIDATION TOOL

Every control on this page depends on one instrument, the LCSH framework. This framework carries a preregistered, multi-stage validation record executed under a cryptographically memorialized lock: pass/fail criteria frozen before any data was collected, every analysis bound to a SHA-256 hash, deviations recorded append-only — and **every registered outcome reported, including the tests that failed.** The instrument is designed as a canary for deployed AI: a repeatable, moment-in-time behavioral check whose value is in re-administration as models, prompts, tools, and policies change — not a one-time certificate that a system is safe. The full record is documented in a peer-review-ready IEEE paper, with every quantitative claim traceable to sealed, independently verifiable artifacts.

SB 315 REQUIREMENT SUMMARY

SB 315 REQUIREMENT (frontier developers)[1][2]	THE SAME PATTERN, DELIVERED FOR DEPLOYED AI
Independent third-party audits, no financial conflicts of interest	Neutral third-party assessment; read-only auditor verification; public hash verification with no login
Published safety framework, updated annually, addressing third-party evaluations	Board Pack PDF with public verification link; attestation records; scheduled reassessment cadence
Pre-deployment transparency reports for new or substantially modified models	Model-Change Reassessment Gate; assessment of the deployment as configured before release
Critical safety incident reporting — 72h standard / 24h imminent-harm	Scheduled reassessment and drift detection as the identification substrate; escalation chains; evidence-linked incident records
Internal reporting processes for safety concerns	Attestation with recorded dissent; escalation work items with named routing

SCOPE AND CLAIMS

SB 315 applies to frontier developers — entities training models above a defined frontier-scale compute threshold — with most duties on 'large frontier developers' exceeding $500 million in annual revenue. Enterprises deploying AI are not covered entities, and this datasheet does not assert SB 315 compliance, certification, or qualified-auditor status. Auditor-qualification rules are expected via Illinois Emergency Management Agency and Attorney General guidance before the January 1, 2027, effective date. [2] This document describes how the verification pattern the Act establishes can be applied to deployed AI systems. It is not legal advice; consult counsel on statutory applicability. Assessment results are moment-in-time behavioral attestations of a deployment as configured — not safety certifications — and their assurance value decays as the deployment changes; the intended use is recurrent re-assessment. References to the instrument’s validation record describe a preregistered protocol with all registered outcomes reported; they are not a claim that any assessed system is “safe.”

[1] Office of Gov. JB Pritzker, “Gov. Pritzker Signs Nation-Leading Artificial Intelligence Safety Law,” press release, July 6, 2026.

[2] Illinois General Assembly, SB 315 — Artificial Intelligence Safety Measures Act (2026); agency guidance pending; provisions summarized from published legal analyses of the enrolled bill (July 2026).

[3] NIST AI Risk Management Framework 1.0 (2023); ISO/IEC 42001:2023 — Artificial intelligence management system.

Deployed AI is already in your org.
The evidence layer is one click away.

aiassesstech.com · greg@gidanc.ai · Austin, TX