

AI4LAW WORKSHOP · ICML 2026

Governance-as-Code: Translating EU AI Act Technical Requirements into Executable Compliance Pipelines for Generative AI Systems

Rudrendu Kumar Paul · Sourav Nandy

Boston University University of Texas at Austin

High-risk obligations bind on Aug 2, 2026, but Articles 8-15 were drafted for predictive AI

THE DEADLINE AND THE STAKES

On **August 2, 2026**, the high-risk provisions of the EU AI Act take full effect. Non-compliance carries penalties of up to **€15M or 3% of global annual turnover**.

Articles 8-15 presume bounded output spaces, deterministic inference, tractable data provenance, and interpretable decision boundaries. Generative systems violate all four at once.

Point-in-time conformity audits go **blind exactly where generative systems drift**.

A FAILURE THE RULES NEVER CATCH

A retrieval-index rebuild shifts the embedding distribution. Domain hallucination climbs **4% → 9%**.

Model weights are unchanged, so it never trips Article 43's **"substantial modification"** trigger. No re-assessment fires.

The risk management system Article 9 requires is built around a single audit, so quality drift between audits is invisible to it.

Seven technical gaps where Articles 8-15 break on LLM-based systems

WHERE THE REGULATION MEETS GENERATIVE AI

Gap	Article
G1: Output non-determinism	Art. 10
G2: Training-data provenance	Art. 11 / 53
G3: Continuous conformity	Art. 9
G4: Oversight at speed	Art. 14
G5: Open-ended robustness	Art. 15
G6: Emergent risk	Art. 9
G7: Generative fairness	Art. 10

GETTING THE LAW RIGHT: THE VALUE CHAIN

A framing error in technical treatments: demanding per-sample documentation of a web-scale corpus from a downstream deployer. The Act does not ask for that.

Under **Article 25 / Chapter V**, the deployer relies on the upstream provider's **Article 53(1)(d) training-data summary** and documents only the layers it controls: fine-tuning data, alignment data, and the retrieval corpus.

Governance-as-Code: 43 machine-checkable criteria, declared not buried

SIX MODULES AS REGO POLICY FOR OPEN POLICY AGENT

43 acceptance criteria run across CI / CD / runtime; every pass emits JSON-LD evidence indexed by Article.

Module	Criteria
Data Lineage	7
Output Monitoring	5
Oversight Hooks	7
Robustness Testing	9
Risk Registry	8
Audit Logging	7

DECLARED, NOT BURIED

Robustness threshold:

$$\tau_{\text{deploy}} = \max(\tau_{\text{floor}}, \tau_{\text{baseline}} - \delta)$$

Every number is written to the evidence store with its source.

Framing bias becomes **8 measurable proxies** tested by counterfactual demographic probing, so "possible biases" is a per-feature test with an effect size, not a vague claim.

Benchmarked against a manual expert audit, the right baseline

TWO ENTERPRISE DEPLOYMENTS

12k

queries/day, high-risk advisory
chatbot (Annex III 5(b))

3,500

items/day, limited-risk content
generator (Art. 50)

GaC reproduced **all 12 manual-audit findings** (8 chatbot + 4 generator), including **3 penalty-triggering** issues. Documentation artifacts flagged none of them.

The baseline is a **manual expert audit**, not Model Cards or FactSheets. Those describe a system; they do not gate a deployment.

AUDIT LABOR: GAC VS. MANUAL EXPERT AUDIT

**~75%**

audit-labor reduction

43/43

Art. 8-15 criteria covered

Effort shifts from gathering evidence to interpreting flagged failures.

WHAT TO CARRY AWAY

- Vague standards ("appropriate levels," "possible biases") must become **explicit, recorded numbers** before any check can verify them
- Get the **value chain** right: the deployer relies on the provider's Article 53 summary (Art. 25) and documents only what it controls
- Benchmark against a **manual expert audit**, the actual status quo, not documentation artifacts built to communicate rather than enforce
- **43 criteria, six modules**, reproducing the manual audit's findings at roughly a quarter of the labor

THE CORE IDEA

Every operationalizing choice is surfaced as recorded, contestable evidence, so a passing result **invites scrutiny rather than foreclosing it**.

TAKEAWAY

Compliance you can run, and audit. As the August 2, 2026 deadline arrives, teams need executable checks whose judgments are visible.

Thank you. Join us at the AI4Law workshop session.