A G E N T I C A S S U R E

The Enterprise AI Assurance Checklist

30 Questions Every Board, CISO, CRO, and Compliance Team Should Ask Before Deploying AI

> ***AI governance without testing is theatre.***
>
> *A policy does not prove an AI system is safe. A committee does not prove that it behaves correctly. A vendor statement does not prove it is defensible.*
>
> *This checklist gives your board, risk, security, and compliance teams a single document to ask the right questions and a defensible way to answer them.*

H O W T O U S E T H I S C H E C K L I S T

Use this checklist to: baseline existing AI systems · qualify vendor AI products · prepare for board, audit, or regulator review · prioritise an independent AI assurance assessment.

Maps to: EU AI Act · NIST AI RMF 1.0 · ISO/IEC 42001 · ISO/IEC 23894 · MAS FEAT · UK FCA / PRA SS1/23 · GDPR Art. 22, 35 · SOC 2.

H O W T O S C O R E

For each question, mark one:

Y — Evidenced. We have documented, testable proof.
P — Partial. We have policy or intent, but no test evidence.
N — No evidence. Gap exists.
N/A — Not applicable, with stated rationale.

Scoring guide: Fewer than 24 of 30 evidenced (Y) = your AI is **not defensible** to a regulator, auditor, or enterprise buyer today.

1. AI SYSTEM INVENTORY

If you cannot list every AI system in production, you cannot govern any of them.

#	Question	Y / P / N
1.1	Do you maintain a single, current inventory of every AI / ML / LLM / agentic system in production, staging, and shadow use, including third-party and embedded AI?	☐
1.2	For each system, do you record owner, business purpose, model lineage, version, training data source, and downstream decisions affected?	☐
1.3	Have you classified each system against the EU AI Act risk tier (Prohibited / High-Risk / Limited / Minimal) and logged the rationale?	☐
1.4	Are AI systems integrated into your enterprise risk register, change management, and procurement process, not run as side projects?	☐

2. DATA AND PRIVACY EXPOSURE

Most AI privacy incidents are not breaches. They are leakage at inference.

#	Question	Y / P / N
2.1	Have you tested whether the model leaks PII, secrets, or training data under membership inference, entropy probing, and seeded extraction attacks?	☐
2.2	Are inputs and outputs scanned and redacted for PII, PHI, PCI, and confidential business data, at runtime, not just in policy?	☐
2.3	Do you have a DPIA / Article 35 record for every customer-facing or decision-impacting AI system, with a stated lawful basis and Art. 22 review?	☐
2.4	Are training data, embeddings, vector stores, and prompt logs subject to the same retention, residency, and right-to-erasure controls as production data?	☐

3. HALLUCINATION AND RELIABILITY TESTING

A confident, fluent, wrong answer is the most dangerous output an enterprise AI can produce.

#	Question	Y / P / N
3.1	Have you measured a hallucination rate against a domain-specific ground-truth benchmark, not a generic public eval?	☐
3.2	For RAG and agentic systems, do you test grounding accuracy, citation faithfulness, and refusal behaviour when retrieved context is missing or contradictory?	☐
3.3	Have you defined acceptable error thresholds by use case (e.g. medical, legal, financial) with documented sign-off from the business owner?	☐
3.4	Are reliability tests re-run automatically on every model, prompt, RAG corpus, or tool change, with results versioned?	☐

4. PROMPT INJECTION AND ADVERSARIAL RISK

If your AI takes input from a user, a document, a webpage, or another agent — it is attackable.

#	Question	Y / P / N
4.1	Have you red-teamed the system for direct and indirect prompt injection, jailbreaks, system-prompt extraction, and tool / function-call abuse?	☐
4.2	For ML / CV models, have you tested adversarial robustness (FGSM, perturbation, evasion) and model extraction through API misuse?	☐
4.3	Are guardrails (input/output filters, policy classifiers, allow-lists, sandboxing of tools and code execution) in place, and independently tested, not vendor-attested?	☐
4.4	Do you maintain a threat model for each AI system covering OWASP LLM Top 10 and MITRE ATLAS, reviewed at least quarterly?	☐

5. REGULATORY READINESS

Regulators will not accept a policy document. They will ask for evidence.

#	Question	Y / P / N
5.1	For each High-Risk system, can you produce the EU AI Act technical documentation pack (Art. 11 + Annex IV), system description, risk management, data governance, testing, monitoring?	☐
5.2	Have findings been mapped to specific regulatory articles and controls (EU AI Act Art. 9, 10, 13, 14, 15, 61 · NIST AI RMF Govern/Map/Measure/Manage · ISO 42001 Annex A)?	☐
5.3	Can you produce a conformity / pre-deployment assessment that states PASS or FAIL, and the system is blocked from deployment on FAIL?	☐
5.4	Are sector-specific obligations covered (MAS FEAT, FCA SS1/23, NYDFS, HIPAA, FDA SaMD) where applicable, with named accountable executive?	☐

6. MONITORING AND DRIFT

An AI system tested only at launch is an AI system you stopped governing on day one.

#	Question	Y / P / N
6.1	Are input drift, output drift, performance drift, and policy-violation rates monitored continuously in production with alert thresholds?	☐
6.2	Are re-tests automatically triggered by model updates, prompt changes, RAG corpus changes, fine-tunes, or upstream provider version bumps?	☐
6.3	Are incidents and near-misses (hallucinations, jailbreaks, unsafe outputs, biased decisions) logged, triaged, and reported under EU AI Act Art. 62?	☐
6.4	Is there a defined kill-switch and rollback procedure, with the last successful test evidence retained for forensic comparison?	☐

7. GOVERNANCE AND ACCOUNTABILITY

If no single executive owns the AI, no one is accountable when it fails.

#	Question	Y / P / N
7.1	Is there a named accountable executive (CRO, CISO, CDO, or AI Risk Officer) for each AI system — recorded in the inventory and the risk register?	☐
7.2	Does an AI Risk / Governance Committee review High-Risk systems with a documented charter, RACI, and minuted decisions?	☐
7.3	Are human oversight controls (Art. 14, review, override, escalation, stop) defined, tested, and exercised at least semi-annually?	☐
7.4	Are third-party and vendor AI assessed against the same standard as in-house systems, with assurance evidence demanded contractually?	☐

8. EVIDENCE AND AUDITABILITY

If you cannot produce the evidence in 24 hours, you do not have it.

#	Question	Y / P / N
8.1	Are all test runs, prompts, responses, scores, and remediation actions stored in a tamper-evident, append-only audit log with timestamp and actor?	☐
8.2	Can you produce, on demand, an executive evidence pack, risk score, findings, regulatory mapping, remediation status, without re-running the work?	☐
8.3	Is the assurance evidence reproducible, same model, same prompts, same data, same result, for a regulator or external auditor?	☐
8.4	Are assurance reports independently produced or independently verifiable, not self-attested by the team that built the model?	☐

YOUR SCORE

Total Y (Evidenced)	Defensibility Level
28 – 30	**Defensible.** You can credibly show evidence to a board, auditor, regulator, or enterprise buyer.
24 – 27	**At Risk.** Material gaps. A regulator or large buyer would find them in week one.
16 – 23	**Exposed.** Governance is mostly policy. Stop new AI deployments until evidenced.
< 16	**Indefensible.** Active board, regulatory, and reputational risk. Independent assurance required immediately.

WHAT TO DO NEXT

If you cannot tick Y for every question with evidence you would put in front of a regulator, you have an assurance gap, not a governance gap.

AgenticAssure is the independent AI assurance platform that turns AI risk into **defensible evidence.**

We test enterprise AI systems for hallucination, privacy leakage, prompt injection, adversarial robustness, unsafe behaviour, and model drift, and produce evidence packs mapped to the EU AI Act, NIST AI RMF, ISO 42001, MAS FEAT, and your internal AI controls. Output is for boards, auditors, regulators, and customers.

Run an AI Assurance Readiness Assessment → www.agenticassure.ai

Independent · Evidence-based · Defensible.