INTAI

DOWNLOADABLE GUIDE

ASSURANCE

AI Assurance Guide

Understand how testing, evidence, controls, monitoring and independent review can build greater confidence in AI systems and their outcomes.

INTAI LIMITED
AI Governance & Assurance Consultancy

www.intai.co.uk
2026 Edition

Contents

1 What Assurance Means for AI

2 Building the Evidence Base

3 Testing & Validation in Practice

4 Ongoing Monitoring

5 Independent Review

6 Demonstrating Assurance to Others

AI Assurance Guide

Understand how testing, evidence, controls, monitoring and independent review can build greater confidence in AI systems and their outcomes.

> Confidence in an AI system and assurance over an AI system are not the same thing. Confidence can be a feeling; assurance requires evidence. This guide explains how testing, documentation, controls, monitoring and independent review combine to build genuine assurance over how an AI system operates and what it produces — and how to demonstrate that assurance credibly to regulators, customers or auditors if asked.

CHAPTER 1 · DEFINITION

What Assurance Means for AI

Assurance is the evidence that supports confidence, not the confidence itself. An organisation can feel confident about an AI system for reasons that have little to do with actual evidence — familiarity, the absence of complaints so far, or simply trust in the team that built it. None of these constitute assurance in a meaningful sense.

> *Confidence without evidence is an opinion. Assurance is what turns that opinion into something defensible.*

Genuine assurance means the organisation could produce specific evidence, on request, that a system has been tested appropriately, is monitored on an ongoing basis, and has been reviewed by someone independent of those who built or operate it. This standard applies with more force as a system's risk level increases.

KEY TAKEAWAYS

- Assurance is the evidence supporting confidence, distinct from confidence as a feeling.
- Genuine assurance means specific evidence could be produced on request, not just general reassurance.
- The standard of assurance required should scale with a system's risk level.

CHAPTER 2 · EVIDENCE

Building the Evidence Base

Assurance evidence accumulates over a system's life, but only if it is deliberately captured along the way. Organisations that only think about assurance retrospectively, after a system is already live, often find the evidence they need was never generated in the first place.

EVIDENCE TYPE	WHAT IT DEMONSTRATES
System documentation	What the system does, how, and its intended use boundaries
Data documentation	What data was used to build or operate the system, and its provenance
Test results	How the system performs against defined acceptance criteria
Monitoring records	How the system has actually performed once live
Review records	That independent scrutiny has genuinely taken place

A useful discipline is to ask, at each stage of a system's life, what evidence this stage should leave behind — and to treat the absence of that evidence as a gap worth closing before moving to the next stage.

KEY TAKEAWAYS

- Assurance evidence should be captured deliberately throughout a system's life, not reconstructed after the fact.
- Five evidence types matter most: system documentation, data documentation, test results, monitoring records and review records.
- Treat the absence of expected evidence at any stage as a gap, not an acceptable shortcut.

CHAPTER 3 · TESTING

Testing & Validation in Practice

Testing that only covers convenient, expected scenarios provides limited real assurance. Meaningful validation deliberately includes edge cases, adversarial inputs, and testing across different groups to catch fairness issues that overall accuracy figures would not reveal.

What Good Testing Covers

1. Performance against realistic, representative real-world scenarios
2. Behaviour with incomplete, unusual or adversarial inputs
3. Consistency of outcomes across relevant groups, not only aggregate accuracy
4. Behaviour after any material change to the system or its underlying data

Testing should also be repeated, not treated as a one-time gate. A system validated before launch has not necessarily been validated for how it performs six months later, once real-world data and usage patterns have shifted.

KEY TAKEAWAYS

- Meaningful testing goes beyond convenient scenarios to edge cases and group-level outcomes.
- Aggregate accuracy figures can mask significant performance differences across specific groups.
- Testing should be repeated over time, not treated as a single pre-launch gate.

CHAPTER 4 · MONITORING

Ongoing Monitoring

Assurance does not end at deployment. Ongoing monitoring is what allows an organisation to detect when a system's real-world performance has drifted from what was validated before launch — a common occurrence as data, usage patterns and context evolve.

> *A system that was assured at launch and never monitored afterwards is not currently assured — it was assured once.*

Making Monitoring Useful

Monitoring is only valuable if it is actively reviewed and has defined thresholds that trigger further action. Data collected but never examined provides the appearance of oversight without its substance.

KEY TAKEAWAYS

- Assurance is not a one-time state established at launch — it requires ongoing monitoring to remain current.
- System performance can drift meaningfully after deployment as real-world conditions change.
- Monitoring data must be actively reviewed against defined thresholds to provide genuine value.

CHAPTER 5 · INDEPENDENCE

Independent Review

Independent review means someone without a stake in a system's success assesses it critically. This matters because those who build or operate a system, however well-intentioned, are naturally less likely to identify its weaknesses than an independent reviewer would be.

REVIEW APPROACH	BEST SUITED TO
Internal audit review	Organisations with an established internal audit function
Cross-functional peer review	Smaller organisations without dedicated audit capacity
External third-party review	Higher-risk or public-facing systems warranting outside scrutiny

KEY TAKEAWAYS

- Independent review provides scrutiny that those closer to a system are less likely to provide themselves.
- The right review approach scales with organisational size and system risk level.
- Higher-risk or public-facing systems benefit from external, third-party review specifically.

CHAPTER 6 · DEMONSTRATION

Demonstrating Assurance to Others

Assurance that cannot be demonstrated to an external party — a regulator, a customer, an auditor — provides only partial value. A useful test of whether assurance is genuinely embedded is whether it could be produced credibly and promptly if requested.

1. Could you produce testing and monitoring evidence for your highest-risk AI system within a few working days if asked?
2. Would that evidence demonstrate ongoing assurance, or only a single point-in-time assessment from launch?

> **WHERE TO GO NEXT**
>
> INTAI's AI Assurance Checklist offers a structured way to test whether a specific system has adequate assurance in place, and the AI Risk Management Guide explores how assurance activity connects to wider risk management.

KEY TAKEAWAYS

- Assurance that cannot be readily demonstrated externally provides only partial practical value.
- The ability to produce evidence promptly is a useful real-world test of whether assurance is genuinely embedded.
- Evidence should demonstrate ongoing assurance, not only a historic point-in-time assessment.

INTAI

Want support applying this guide?

INTAI works with organisations across Europe — from SMEs to public sector bodies and larger enterprises — to turn AI governance guidance into a practical, proportionate programme. If you would like support applying the ideas in this guide to your own organisation, our team is glad to help.

www.intai.co.uk | hello@intai.co.uk

© 2026 INTAI Limited. This guide is provided for general guidance purposes only and does not constitute legal, regulatory or professional advice. Organisations should adapt it to their own context and seek independent advice where appropriate.