

# AI in Governance: What Actually Works

A recorded session on where AI genuinely helps a governance function, where it quietly creates risk, and the operating pattern that keeps a human accountable for every decision that matters.

## What AI is genuinely good at here

The useful applications are unglamorous and they all share one trait: the AI proposes, and a person disposes.

- Summarising a week of exceptions into something a committee can read in five minutes.
- Spotting patterns across variants that nobody had time to look for.
- Drafting the first version of a rule, a risk statement, or a control description.
- Explaining a framework practice or objective in the language of your organisation.

## Where it goes wrong

Every failure we see in the field falls into one of four buckets, and none of them are fixed by a better model.

- Unattributed output — an AI-drafted conclusion that ends up in a board pack with no label.
- Confidence without provenance — a plausible number that nobody can trace to a source.
- Silent scope creep — a tool that started as a summariser and quietly became an approver.
- Data leakage — operational detail pasted into an ungoverned tool.

## The operating pattern that works

One pattern covers almost every safe use: label, score, hold, approve, log.

- Label every AI-generated artefact as AI-generated, wherever it appears.
- Attach a confidence score and the metrics the suggestion was derived from.
- Hold it in a pending state — nothing reaches a report until a person accepts it.
- Record who approved it, when, and what they changed.

## Questions from the session

Does the AI ever act on its own? Not in a governance context. Actions are automations with explicit rules; AI drafts the case for change, people authorise it.

How do we handle a wrong suggestion? Reject it with a reason. Rejections are as valuable as approvals because they are the record that review is real.

What about the regulator? Show them the log. A labelled, scored, human-approved trail is a far stronger position than claiming no AI was involved.

## Video walkthrough — key points

### 1. AI in Governance

- What actually works
- And what quietly creates risk

This is the session summary for AI in governance: what actually works. We will cover where artificial intelligence genuinely helps a governance function, where it quietly creates risk, and the operating pattern that keeps a person accountable for every decision that matters.

### 2. What works

- Summarising exceptions
- Spotting patterns
- Drafting first versions

The useful applications are unglamorous, and they share one trait: the AI proposes and a person disposes. Summarising a week of exceptions into five readable minutes. Spotting patterns across variants nobody had time to look for. Drafting the first version of a rule or a control description. Explaining a framework practice in your own language.

### 3. Where it goes wrong

- Unattributed output
- Confidence without provenance
- Silent scope creep
- Data leakage

Every failure we see falls into four buckets, and a better model fixes none of them. Unattributed output that lands in a board pack with no label. Confidence without provenance, a plausible number nobody can trace. Silent scope creep, where a summariser quietly becomes an approver. And data leakage into ungoverned tools.

### 4. Label, score, hold, approve, log

- Nothing reaches a report unapproved
- Every approval recorded

One pattern covers almost every safe use. Label every AI generated artefact as AI generated. Attach a confidence score and the metrics it came from. Hold it pending, so nothing reaches a report until a person accepts it. Then record who approved it, when, and what they changed.

### 5. From the Q and A

- AI drafts; people authorise
- Rejections are evidence too
- Show the regulator the log

From the questions. Does the AI ever act alone? Not in governance. Actions are automations with explicit rules; the AI drafts the case for change and people authorise it. What about a wrong suggestion? Reject it with a reason — rejections prove the review is real. And for a regulator, a labelled, scored, human approved trail is a far stronger position than claiming no AI was involved.

© Automated Governance Systems. Shared for use inside your organisation. For a walkthrough against your own processes, book a session at [ags-apip.lovable.app/book-demo](https://ags-apip.lovable.app/book-demo).