

A SUPERWISE® LEARN KIT · OCTOBER 2026

# Real governance or theater: a 10-point scorecard

Score one AI workflow against ten controls drawn from NIST, ISO, and the EU AI Act, using evidence from the last 90 days. Every zero becomes a task with an owner and a date.

# Governance is a loop that leaves evidence

The three frameworks a firm hears about most describe the same thing: a management loop that runs on a schedule and leaves dated records behind it. A written policy is where that loop starts.

| FRAMEWORK                     | WHAT IT DESCRIBES IN OPERATION                                                                                                                                                                                                   |
|-------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>NIST AI RMF 1.0</b> [1, 2] | Risk management that continues across the life of a system. Systems are tested before deployment and monitored in production, and measured risks drive treatment and monitoring.                                                 |
| <b>ISO/IEC 42001</b> [3, 4]   | An AI management system run on Plan-Do-Check-Act, with recurring risk assessment, treatment, and continual improvement. Cited from ISO's public description; the standard itself is paywalled and was not reviewed for this kit. |
| <b>EU AI Act</b> [5]          | For high-risk AI: a continuous risk-management system, testing, logs, human oversight, a quality-management system, and post-market monitoring (Articles 9, 12, 14, 17, and 72).                                                 |

The EU AI Act applies by role and by where a system's output is used, and its duties phase in over several years (Articles 2 and 113) [5]. A US firm serving a European client should ask counsel what applies. The scorecard uses the Act for what it describes, which is a useful picture of operating controls for any firm.

## The test underneath all three

Ask a named person to show four things for one AI workflow: that the process ran, that it found something, that a decision followed, and that someone checked again afterwards. A policy binder answers the first question at best. The scorecard measures all four, control by control.

## How to score

|                                                                                   |                                                                                           |                                                                                                  |
|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| <b>0</b><br><b>No evidence.</b> Nobody can produce a record, a name, or a sample. | <b>1</b><br><b>On paper, or partly.</b> A document, an owner, or a partial sample exists. | <b>2</b><br><b>Operating.</b> The control runs on a defined cadence and produces dated evidence. |
| <b>16 to 20</b><br>Operating governance                                           | <b>10 to 15</b><br>Incomplete                                                             | <b>0 to 9</b><br>Policy-led theater                                                              |

**RULE OF THUMB** The bands are a practical guide drawn from the research behind this kit. No standard sets them, and a score proves no conformity with ISO/IEC 42001 or any other framework.

### PLAIN ENGLISH • OPERATING CONTROL

Something that acts on the AI work itself and leaves a record when it does: a setting that blocks a data type, a reviewer who signs off and can stop a release, an alert someone answers. An operating control can prevent, detect, escalate, constrain, or correct a risk. You can point to the last time it did.

# Ten controls, scored on evidence

Pick one approved AI workflow, such as drafting client letters in ChatGPT or summarizing records in Copilot. Ask the owner for the two-point evidence on each line, and score only what you are shown.

|           |                  |                 |
|-----------|------------------|-----------------|
| WORKFLOW  | AI TOOL AND PLAN | BUSINESS OWNER  |
| SCORED BY | DATE             | EVIDENCE WINDOW |

|          |                                                                                                                                                                                                                                                                                                 |                                                                                                     |                                           |
|----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|-------------------------------------------|
| <b>1</b> | <b>Inventory and boundaries</b><br><b>A 2 looks like:</b> a current list of AI systems with owners, users, vendors, data, purpose, and prohibited uses, plus a way of finding tools nobody approved.<br><small>NIST GOVERN 1.6, MAP 1 • ISO 42001 scope • EU Art. 9</small>                     | <b>SCORE</b><br><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/><br>0 1 2 | <b>EVIDENCE SEEN</b><br><hr/> <hr/> <hr/> |
| <b>2</b> | <b>Named accountability</b><br><b>A 2 looks like:</b> one business owner and one control owner for the workflow, with authority to stop its use, and records of the decisions they made.<br><small>NIST GOVERN 2 • ISO leadership and accountability • EU Art. 17</small>                       | <b>SCORE</b><br><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/><br>0 1 2 | <b>EVIDENCE SEEN</b><br><hr/> <hr/> <hr/> |
| <b>3</b> | <b>Risk assessment before use</b><br><b>A 2 looks like:</b> a dated assessment of this workflow: the people affected, the data, the ways it can fail, and the decision to approve, limit, or reject.<br><small>NIST MAP 1, MAP 2, MAP 5 • ISO risk assessment and treatment • EU Art. 9</small> | <b>SCORE</b><br><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/><br>0 1 2 | <b>EVIDENCE SEEN</b><br><hr/> <hr/> <hr/> |
| <b>4</b> | <b>Vendor and data due diligence</b><br><b>A 2 looks like:</b> contract and technical evidence for retention, training use, subprocessors, access, deletion, incident notice, and how you would leave.<br><small>NIST GOVERN 6, MAP 4 • EU Art. 26 (deployers)</small>                          | <b>SCORE</b><br><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/><br>0 1 2 | <b>EVIDENCE SEEN</b><br><hr/> <hr/> <hr/> |
| <b>5</b> | <b>Tested safeguards</b><br><b>A 2 looks like:</b> test cases showing that high-risk inputs or actions are blocked, escalated, or constrained, run before release and again since.<br><small>NIST MEASURE 2.3, MEASURE 2.6 • ISO Plan-Do-Check-Act • EU Arts. 9 and 15</small>                  | <b>SCORE</b><br><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/><br>0 1 2 | <b>EVIDENCE SEEN</b><br><hr/> <hr/> <hr/> |

|           |                                                                                                                                                                                                                                                                                                      |                                                                                                         |                                           |
|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|-------------------------------------------|
| <b>6</b>  | <b>Human review with authority</b><br><b>A 2 looks like:</b> reviewers who know what to inspect and can override or stop the system, and sampled records showing that they did.<br><br><small>NIST GOVERN 2, MANAGE 4.1 · EU Art. 14</small>                                                         | <b>SCORE</b><br><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/><br>0   1   2 | <b>EVIDENCE SEEN</b><br><hr/> <hr/> <hr/> |
| <b>7</b>  | <b>Logs and attribution</b><br><b>A 2 looks like:</b> records that connect input, output, model, rule result, user or agent, reviewer, and final action, with defined access and retention.<br><br><small>NIST MEASURE 2.4, GOVERN 4.3 · EU Arts. 12 and 26(6)</small>                               | <b>SCORE</b><br><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/><br>0   1   2 | <b>EVIDENCE SEEN</b><br><hr/> <hr/> <hr/> |
| <b>8</b>  | <b>Live monitoring and incidents</b><br><b>A 2 looks like:</b> defined measures, alert thresholds, an incident route, and evidence that alerts were investigated and the cause corrected.<br><br><small>NIST MEASURE 2.4, MANAGE 4.1, MANAGE 4.3 · ISO checking and improvement · EU Art. 72</small> | <b>SCORE</b><br><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/><br>0   1   2 | <b>EVIDENCE SEEN</b><br><hr/> <hr/> <hr/> |
| <b>9</b>  | <b>Training verified in behavior</b><br><b>A 2 looks like:</b> role-specific training that is completed, tested, and refreshed, including for contractors, and sampled against how people use the tool.<br><br><small>NIST GOVERN 2.2 · ISO competence · EU Arts. 4 and 26</small>                   | <b>SCORE</b><br><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/><br>0   1   2 | <b>EVIDENCE SEEN</b><br><hr/> <hr/> <hr/> |
| <b>10</b> | <b>Review and improvement</b><br><b>A 2 looks like:</b> management reviews results on a cadence, assigns fixes, confirms they closed, and changes controls after incidents or material changes.<br><br><small>NIST GOVERN 1.5, MANAGE 4 · ISO continual improvement · EU Arts. 9, 17, and 72</small> | <b>SCORE</b><br><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/><br>0   1   2 | <b>EVIDENCE SEEN</b><br><hr/> <hr/> <hr/> |

TOTAL, OUT OF 20

\_\_\_\_\_

☐ 16 to 20 · Operating    ☐ 10 to 15 · Incomplete    ☐ 0 to 9 · Theater

Sources for each basis line: NIST [1, 2], ISO [3, 4], EU AI Act [5]. Every zero goes to the fix list in part 5.

# Five records an operating system can produce

Short on time? Ask the workflow owner for these five, and write down the date of each. A policy binder cannot produce any of them. A control that runs produces all five as a matter of course.

| ASK FOR                                                 | LAST DATE | WHERE IT IS KEPT |
|---------------------------------------------------------|-----------|------------------|
| The last risk review of this workflow.<br>Items 3, 10   |           |                  |
| The last blocked or escalated event.<br>Items 5, 8      |           |                  |
| The last human override of an AI output.<br>Item 6      |           |                  |
| The last incident or false - positive review.<br>Item 8 |           |                  |
| The ticket that shows the fix closed.<br>Item 10        |           |                  |

## Reading the result

- **Five dates inside 90 days.** The loop runs. Score the full card to find the weak controls.
- **Some dates, some blanks.** Parts of the loop run. The blanks show where the scorecard will land zeros.
- **No dates.** The workflow runs on the policy alone. Start the fix list with items 2, 5, and 7: an owner, a tested safeguard, and a record.

**Score from evidence, every time.** “We do that” earns a 1 at most. A 2 needs a dated record from the evidence window, shown to the person scoring. The same rule applies when the workflow belongs to a vendor: ask for their record, then check it against your own.

## Where the written rule did not reach the work

Each case comes from a regulator's complaint, a safety board's findings, or a tribunal's decision. Three cases show the pattern; they say nothing about how often it happens.

US FEDERAL TRADE COMMISSION • COMPLAINT FILED DECEMBER 19, 2023

### Rite Aid, facial recognition in stores

|                         |                                                                                                                                                                                                                         |
|-------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>The written rule</b> | Store employees who operated the system were to receive about one to two hours of training.                                                                                                                             |
| <b>What operated</b>    | The FTC's complaint alleges that in nearly all cases Rite Aid did not verify or keep any record that the training happened, and that the system produced false matches that harmed customers (paragraphs 66 to 72) [6]. |
| <b>Scorecard items</b>  | 9, training verified in behavior; 8, monitoring and incidents; 10, review.                                                                                                                                              |

US NATIONAL TRANSPORTATION SAFETY BOARD • FINDINGS ADOPTED NOVEMBER 19, 2019

### Uber ATG, automated-driving test vehicle

|                         |                                                                                                                                                                                                                                                                                         |
|-------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>The written rule</b> | Uber ATG had written policies and procedures for its vehicle operators, held in the NTSB docket [8].                                                                                                                                                                                    |
| <b>What operated</b>    | The operator was distracted by her personal phone throughout the trip. The NTSB named inadequate risk assessment, ineffective oversight of operators, and no adequate mechanism for automation complacency as contributing causes, and recommended a full safety management system [7]. |
| <b>Scorecard items</b>  | 3, risk assessment; 6, human review with authority; 8, monitoring.                                                                                                                                                                                                                      |

BRITISH COLUMBIA CIVIL RESOLUTION TRIBUNAL • DECIDED FEBRUARY 14, 2024

### Air Canada, the bereavement-fare chatbot

|                         |                                                                                                                                                                                                                                                                                         |
|-------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>The written rule</b> | Air Canada's published fare policy on its website.                                                                                                                                                                                                                                      |
| <b>What operated</b>    | The airline's chatbot told a customer he could apply for the fare after travel, which the policy did not allow. The tribunal held Air Canada responsible for the information on its site, chatbot included, and rejected the argument that the chatbot was a separate legal entity [9]. |
| <b>Scorecard items</b>  | 5, tested safeguards; 6, human review; 8, monitoring.                                                                                                                                                                                                                                   |

Air Canada is a Canadian decision, included because it is a direct public ruling on the gap between a written policy and an AI answer. The Rite Aid complaint states the FTC's allegations.

### What the three have in common

Each organization had the rule in writing, and nothing checked that the rule held in the work. Training went unrecorded, an operator went unwatched, and a chatbot answered with nobody comparing it against the published policy. A small firm has the same gap when staff sign an AI policy that says "no client data in ChatGPT" and nothing ever looks at a prompt.

- **Item 5 closes it at the request.** A tested safeguard acts on the prompt, whoever sends it.
- **Item 7 shows whether it held.** A record of what was sent turns a signed policy into something you can check.
- **Item 9 confirms people follow it.** Sample real use against the training, and keep the sample.

[illegible]

- ☐ **Item 2: name the owners.** One business owner and one control owner, written on the scorecard, each with the authority to pause the workflow.
- ☐ **Item 5: write three test cases.** A made-up client record, a pasted credential, and a request the policy forbids. Run them, and keep what happened.
- ☐ **Item 7: export one record.** Pull one log entry for the workflow and check it names the user, the time, and what was sent.
- ☐ **Item 10: book the next review.** Put the re-score on the owner's calendar, 90 days out.

7

## Evidence for three of the ten items

---

Items 5, 7, and 8 need something that acts on each request and keeps a record. A gateway on the PC can supply part of that for the AI traffic sent through it. The other seven items are people and process, and stay with the firm.

### What SUPERWISE Sentinel does

As described on [superwise.ai/sentinel/](https://superwise.ai/sentinel/), read October 8, 2026 [10]:

- **Item 5, a safeguard you can test.** SUPERWISE® Sentinel replaces emails, phone numbers, account and card numbers, government IDs, and API keys with placeholders before a prompt reaches ChatGPT or Claude. Type a made-up client record and see what the model receives.
- **Item 7, a record.** Every inspected interaction lands in an Activity Log: what was checked, what was redacted, and when. Check that the entry carries the user identity item 7 asks for.
- **Item 8, rule triggers you can see.** The same log shows where a rule fired and what it caught, on a screen your own people open.

### What it leaves to you

- **Item 1 needs its own inventory.** Sentinel sees the traffic sent through it, so a tool used outside that path stays off its record.
- **Its coverage has limits.** It catches the data types it lists, in Claude Code, Claude Desktop, and, through a browser extension in preview, ChatGPT and Claude in Chrome and Edge on Windows.
- **It certifies nothing.** A higher score with a gateway in place is still an internal diagnostic. Sentinel makes no tool or workflow compliant.

**Trying it.** Sentinel runs on the machine or in your own cloud, and nothing is routed through SUPERWISE. It installs on Windows 10 and 11 without admin rights, free for 30 days with no card, then \$10, \$25, or \$50 a month per deployment, as shown on [superwise.ai/pricing/](https://superwise.ai/pricing/) on October 8, 2026 [10].

# Sources

---

Read October 6, 2026, unless marked. Standards and rules change, and the EU AI Act phases in; check the current version before you rely on one.

- 1 National Institute of Standards and Technology (January 2023), *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1. Subcategory wording checked October 8, 2026. [nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf](https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf)
- 2 NIST AI Resource Center, *AI RMF Core*. Continuous risk management, testing before deployment and in operation. [airc.nist.gov/airmf-resources/airmf/5-sec-core/](https://airc.nist.gov/airmf-resources/airmf/5-sec-core/)
- 3 International Organization for Standardization, *AI management systems*. ISO's public description of ISO/IEC 42001 and Plan-Do-Check-Act. [iso.org/artificial-intelligence/ai-management-systems](https://iso.org/artificial-intelligence/ai-management-systems)
- 4 International Organization for Standardization, *ISO/IEC 42001:2023* overview, and the ISO publication on competence and operation. The full standard is paywalled and was not reviewed; this kit claims no clause-level coverage. [iso.org/standard/81230.html](https://iso.org/standard/81230.html) · [iso.org/publication/PUB200420.html](https://iso.org/publication/PUB200420.html)
- 5 Regulation (EU) 2024/1689, the *Artificial Intelligence Act*. Articles 2, 4, 9, 12, 14, 15, 17, 26, 72, and 113. [eur-lex.europa.eu/eli/reg/2024/1689/oj](https://eur-lex.europa.eu/eli/reg/2024/1689/oj)
- 6 Federal Trade Commission v. Rite Aid Corporation, complaint, Case 2:23-cv-05023 (filed December 19, 2023), paragraphs 66 to 72. Paragraph 67 checked October 8, 2026. [ftc.gov/system/files/ftc\\_gov/pdf/2023190\\_riteaid\\_complaint\\_filed.pdf](https://ftc.gov/system/files/ftc_gov/pdf/2023190_riteaid_complaint_filed.pdf)
- 7 National Transportation Safety Board, *Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian*, HWY18MH010 (adopted November 19, 2019). Probable cause and recommendation H-19-52, checked October 8, 2026. [ntsb.gov/investigations/Pages/HWY18MH010.aspx](https://ntsb.gov/investigations/Pages/HWY18MH010.aspx)
- 8 National Transportation Safety Board, public docket HWY18MH010, including Uber ATG's operator policies and procedures. [data.ntsb.gov/Docket?NTSBNumber=HWY18MH010](https://data.ntsb.gov/Docket?NTSBNumber=HWY18MH010)
- 9 *Moffatt v. Air Canada*, 2024 BCCRT 149 (February 14, 2024), British Columbia Civil Resolution Tribunal. [canlii.org/en/bc/bccrt/doc/2024/2024bccrt149/2024bccrt149.html](https://canlii.org/en/bc/bccrt/doc/2024/2024bccrt149/2024bccrt149.html)
- 10 SUPERWISE, *Sentinel and Pricing*, read October 8, 2026. Every product statement in part 6. [superwise.ai/sentinel/](https://superwise.ai/sentinel/) · [superwise.ai/pricing/](https://superwise.ai/pricing/)

## Further reading on superwise.ai

- **The five questions behind the scorecard.** [superwise.ai/learn/ai-governance-vs-theater/](https://superwise.ai/learn/ai-governance-vs-theater/)
- **Policy and control.** [/learn/enforced-or-written-down/](https://superwise.ai/learn/enforced-or-written-down/) and the AI acceptable use policy template at [/learn/ai-acceptable-use-policy/](https://superwise.ai/learn/ai-acceptable-use-policy/)

This scorecard is an internal diagnostic and general information. It is not legal advice, an audit, or a certification of conformity with any standard. Have counsel advise on the rules that apply to your firm.