

When the regulator says "show us it works", what do they actually look at?

The FCA will not write an AI rulebook. It will supervise the AI you already run through the frameworks you already hold. That moves the burden of proof onto you. This is what the proof looks like.

A control that is written down is documentation. A control that demonstrably changes what a person or a system does at the point of decision is governance. The distance between the two is where firms are now exposed.

THE SHIFT, IN ONE LINE

Principles-based supervision means there is no checklist for the examiner to tick and none for you to hide behind. Consumer Duty, SM&CR and operational resilience all resolve to a single question a supervisor can ask on any given day: **show us your AI governance working in practice, not on paper**. Most firms can produce a control library. Far fewer can show that one control in it ever changed an outcome.

DOCUMENTATION — WHAT FAILS THE QUESTION

- ✗ A policy that names a control but cannot show it firing
- ✗ Sign-off that is a click, at speed, near 100% approval
- ✗ Model risk owner on the org chart, absent from decisions
- ✗ Guardrails described in a deck, not enforced at runtime
- ✗ An audit trail assembled after the examiner asks for it

GOVERNANCE — WHAT SURVIVES IT

- ✓ A control with a log of the times it stopped something
- ✓ Review that rejects substantively and routinely
- ✓ A named owner who can be shown deciding, with authority
- ✓ Constraints supplied to the AI at the moment it acts
- ✓ Provenance captured as work happens, not reconstructed

THE TRAP MOST FIRMS ARE IN

The Treasury Committee has already accused the regulators of a wait-and-see approach. The regulators have pushed the same test straight back onto firms. Capability has a leaderboard. Whether a competent check can stop the wrong thing from shipping does not. Adoption has raced ahead of anyone's ability to evidence it, and the evidence is now the thing being examined.

WHAT TO LOOK FOR

Five questions a supervisor can ask that documentation cannot answer

Run these against any AI-touching process before an examiner does. If the honest answer to any of them is a document rather than an event, that is the gap.

1

Where did this output come from, and can you prove it?

For any AI-assisted decision that reached a customer, can you reconstruct which model, which prompt, which data and which human hands touched it, from records captured at the time?

Weak signal: provenance is stitched together retrospectively, or lives in someone's memory.

2

Who is accountable, and can you show them exercising it?

SM&CR needs a named person who understands and controls the risk. Not a box on a chart. Can you show that person making or overturning a decision, with the authority to do so?

Weak signal: an owner exists on paper but never appears in the decision record.

3

What is the AI allowed to do, and how do you know it stayed inside that?

Scope drifts on demos, excitement and seniority. Is the boundary of what the AI may decide defined, and is every expansion of it backed by evidence rather than a meeting?

Weak signal: scope grew because it worked in a pilot, with no reversal ever recorded.

4

When a human reviews the AI, is the review substantive or a signature?

Review is where governance becomes real or quietly gives up. A reviewer who approves everything is not a reviewer. Is review time protected, and does rejection actually happen?

Weak signal: approval rate near 100%, reviews complete in seconds, reviewers also on delivery.

5

Do your controls reach the AI at the moment it acts, or only the policy shelf?

The rules the firm cares about, fair value, treating customers fairly, the constraints your team knows by heart, are they supplied to the system at runtime, or written down somewhere the system never reads?

Weak signal: constraints live in tribal knowledge; violations are caught by humans, sometimes, after the fact.

BEFORE THE EXAMINER DOES

A self-check you can run this sprint

Pick one AI-assisted process that touches customers. Answer honestly. Every box you cannot tick with an **event** rather than a **document** is a place where you hold documentation, not governance.

- ☐ **Evidence.** I can produce, from records made at the time, the full provenance of one live AI-assisted customer decision.
- ☐ **Accountability.** I can name the accountable person and show a specific decision they made or reversed.
- ☐ **Scope.** The AI's decision boundary is written down, and the last time it changed there is evidence behind the change.
- ☐ **Review.** Reviewers have rejected AI output in the last sprint, and review time is protected in planning.
- ☐ **Constraint.** The firm's key rules reach the system at runtime, not just the policy library.
- ☐ **Stop.** There is a point at which a competent check can halt the thing from shipping, and it has been used.

THE ONE THAT MATTERS MOST

If you cannot point to a moment where a competent check actually stopped the wrong thing from shipping, you do not yet have that control. You have its description. A supervisor asking to see governance in practice is asking to see that moment.