fortayconnect

A PRACTICAL TOOLKIT

AI Agents in Financial Services.

FOR SENIOR FINSERV LEADERS

Stakeholders, objections, operating models & what good looks like.

IN PARTNERSHIP WITH

zoom THETALAKE

PUBLISHED
May 2026

What's inside

A working reference for the people inside regulated firms tasked with getting AI agents from idea to live deployment.

01 Why this toolkit exists

If you are reading this, you have already decided AI agents matter for your business. You probably do not need another deck telling you the technology is real, the productivity gains are large, or the competitive risk of inaction is rising.

What you need is help getting an agent into production inside a regulated financial services firm, without the project stalling in a Risk committee, dying in Procurement, or being quietly defunded after the third InfoSec review.

This toolkit is built for that. It assumes you have already done the easy part of building conviction. It is here to help with the hard part: navigating the seven internal stakeholders, twenty-or-so common objections, and three or four governance frameworks that sit between your conviction and a live deployment.

It is opinionated where we think a clear point of view is more useful than a balanced summary, and neutral where you need raw reference material. We have flagged which is which.

> **A note on what this is not.** This is not legal, regulatory, or financial advice. The references here are starting points for your own teams, not substitutes for them. If a regulator asks you a hard question on Friday morning, do not quote this document.

02 The seven stakeholders who decide whether your AI agent ships

Most AI agent projects in regulated firms do not fail on technology. They fail because the project lead underestimates how many internal sign-offs are required, and pitches each stakeholder with the wrong message.

Here are the seven you will need to win over, what they care about, and the one question to answer for them upfront.

Risk

Cares about: enterprise risk exposure, model risk, third-party risk.

Will block on: any deployment that creates a new material risk without a corresponding control.

Answer upfront: what is the worst plausible outcome, how likely is it, and what controls reduce it to acceptable levels?

Compliance

Cares about: regulatory obligations (FCA, PRA, ICO, sector-specific), audit trail, evidenceability.

Will block on: anything that cannot be evidenced to a regulator.

Answer upfront: what record will exist of every agent decision, and how is it retained?

InfoSec

Cares about: data residency, access controls, vendor security posture, attack surface.

Will block on: data flowing to environments they have not approved.

Answer upfront: where does customer and firm data go, how is it protected in transit and at rest, and what is your vendor's last SOC 2, ISO 27001 or pen test result?

Legal

Cares about: contractual liability, IP, data protection law, customer contract terms.

Will block on: terms that put the firm on the hook for the vendor's failures.

Answer upfront: what does the contract say about liability, indemnities, IP ownership of outputs, and exit?

Data Privacy / DPO

Cares about: lawful basis, data minimisation, ROPA, automated decision-making under UK GDPR.

Will block on: undocumented processing or unlawful basis.

Answer upfront: what personal data does the agent process, on what lawful basis, and is automated decision-making in scope of Article 22?

Procurement

Cares about: total cost, supplier viability, contract terms, exit.

Will block on: anything that bypasses the procurement process or creates concentration risk.

Answer upfront: what is the three-year TCO, what is the exit plan, and is the supplier financially stable?

Business Sponsor

Cares about: outcomes, ROI, time to value, accountability.

Will block on: nothing. But if they are not engaged, nothing else matters.

Answer upfront: what business metric will move, by how much, by when, and who is accountable if it does not?

IN SHORT

Seven stakeholders. Seven different questions. Seven different answers.

The mistake we see most often is treating these as one audience and writing one deck. Each of these stakeholders has the power to slow your project down, and most of them will quietly do exactly that if their specific question goes unanswered.

FORTAY'S VIEW

Most teams pitch AI agents to all seven stakeholders the same way. That is the single biggest avoidable mistake we see. Build seven versions of your one-pager (same project, different lens for each audience) before your first formal review.

03 The objections handbook

Ten objections you will hear inside any UK financial services firm, with the response we see leading teams give.

Objection	Response
"What about hallucinations?"	Distinguish generative agents (which can hallucinate) from retrieval and workflow agents (which retrieve from approved sources or follow deterministic logic). Most enterprise agent use cases are the latter. For genuinely generative components, controls include retrieval-grounded responses, confidence thresholds, human-in-the-loop on low-confidence outputs, and post-hoc audit.
"What about data leakage?"	Three layers: contractual (no training on your data, written into the MSA), technical (private endpoints, no data egress to the foundation model provider, customer-managed keys where available), and operational (DLP scanning of inputs and outputs). Ask the vendor for their data flow diagram in writing.
"How do we explain this to a regulator?"	The same way you explain any consequential automated process: documented purpose, documented controls, evidence of testing, evidence of monitoring, evidence of human oversight where required. The FCA has been clear they expect existing frameworks (SM&CR, Consumer Duty, operational resilience) to be applied to AI, not replaced.
"Who is accountable under SM&CR?"	The Senior Manager whose business area owns the agent's function, in the same way they would be accountable for a human team doing the same job. Document it in the Statement of Responsibilities. The agent is a tool; the accountability is the SMF holder's.
"What about Consumer Duty?"	Apply the four outcomes (products and services, price and value, consumer understanding, consumer support) to the agent's interactions. If a human agent doing the same task would need to deliver these outcomes, the AI agent does too. Test for it explicitly.

Objection	Response
"What if it discriminates?"	Bias testing pre-deployment, ongoing monitoring of outcomes by protected characteristic, clear escalation paths for affected customers. The Equality Act 2010 applies to automated decisions the same way it applies to human ones.
"What about model drift?"	Ongoing monitoring against a baseline, scheduled re-evaluation, change control on any model updates from the vendor. Treat it like any other production system that can degrade, because it is.
"Why not just use ChatGPT?"	Three reasons most regulated firms land on enterprise platforms instead: data handling commitments (training, retention, residency), integration with internal systems and identity, and audit and admin controls. The model itself is rarely the differentiator.
"What if the vendor goes bust?"	Standard exit planning, but be specific. Get the vendor's commitment in writing on data return format, model artefact handover (where applicable), and transition support. Apply the same scrutiny you would to any operationally critical supplier under SS2/21.
"How do we know it is working?"	Define success metrics before deployment, instrument for them from day one, and review monthly. Common ones: containment rate (for service agents), accuracy against human baseline, customer outcome metrics, escalation rate. If you cannot define what good looks like, you are not ready to deploy.

04 The operating model question

This is where most AI agent programmes get stuck, and it is the question this event was built around. The technology is solvable. The operating model is harder, because it requires the firm to answer four questions it has usually never had to answer at the same time:

Where does the agent sit?

Inside a business line (e.g. embedded in CX), inside a horizontal function (e.g. a centre of excellence), or as shared infrastructure (e.g. a platform team)? Each model has trade-offs. Embedded models move fastest but create fragmentation. Centralised models govern best but bottleneck delivery. Platform models scale but require maturity to build.

Who owns it day-to-day?

Product? Operations? Engineering? The answer determines the rhythm. Product gives you roadmap discipline, operations gives you outcome focus, engineering gives you technical rigour. Most firms end up with shared ownership; the firms that succeed have one accountable owner.

Who governs it?

Existing model risk committees? A new AI governance body? A subcommittee of an existing risk forum? Net new bodies create work but signal seriousness. Reusing existing bodies is faster but risks AI being underweighted in their agenda.

Who reviews outputs?

For low-risk use cases, periodic sampling. For higher-risk, continuous monitoring with human-in-the-loop. The decision is not technical; it is a risk appetite decision, and it should be made by Risk and Compliance, not the project team.

FORTAY'S VIEW

Pick a model, document it, communicate it, and revisit it every six months. The worst outcome is a default operating model that nobody chose, usually "the team that built the first agent owns everything by accident." That collapses under any kind of scale.

05 A starter business case structure

A business case for an AI agent in a regulated firm needs seven sections. We are not giving you the numbers; those are yours. We are giving you the scaffolding so you do not miss anything that gets queried in committee.

01 Problem statement

Quantify the current state in money or time. "Our team handles 40,000 enquiries a month, average handling time 7 minutes, fully loaded cost £X per minute." Vague problems get vague approvals.

02 Proposed solution

What the agent does, what it does not do, what is in scope for phase one. Be ruthless about phase one scope: most failed deployments tried to do too much.

03 Expected outcomes

Two or three measurable metrics, with baseline and target. Containment rate from X% to Y%. AHT from X to Y. CSAT held flat or improved. Avoid vanity metrics.

04 Cost

Three-year TCO, broken into platform, integration, change management, ongoing operations. Most business cases under-cost the last two by 50%.

05 Risk

A short risk register: top five risks, with likelihood, impact, and mitigation. Include the risks of *not* deploying.

06 Governance

How the deployment is governed, who signs off, who reviews ongoing performance, how exceptions are handled.

07 Decision required

What you are asking for, from whom, by when. If the committee does not know what decision they are being asked to make, they will not make one.

06 The regulatory reference shelf

A starting point, not a substitute for your Compliance team. Each item below should be pulled, current version, by someone who can interpret it for your firm. Dates are publication dates of the most relevant version we are aware of at time of writing.

> **Verify each of these with your Compliance and Legal teams before quoting in internal papers.** Regulatory positions evolve and document codes are superseded. Treat this list as a research starting point, not authoritative source material.

FCA Discussion Paper DP5/22: Artificial Intelligence and Machine Learning

The FCA's foundational paper on AI in financial services, joint with PRA and Bank of England. Sets the direction of travel: existing frameworks (SM&CR, operational resilience, model risk) apply.

FCA / PRA Feedback Statement FS2/23

Follow-up to DP5/22 summarising responses and the regulators' position. Worth reading in full.

FCA Consumer Duty (PS22/9) and ongoing guidance

Applies to all automated interactions with retail customers. The four outcomes (products and services, price and value, understanding, support) apply to AI agents the same as to human ones.

SM&CR: Senior Managers and Certification Regime

Existing accountability framework. AI agents do not change who is accountable; they change what the accountable person is accountable for. Statements of Responsibilities should reflect material AI deployments.

SS2/21: Outsourcing and third party risk management (PRA)

Relevant where the AI agent is materially provided by a third party. Most enterprise AI deployments will trigger this.

UK GDPR and the Data Protection Act 2018

Lawful basis, data minimisation, transparency, ROPA. Article 22 on automated decision-making is particularly relevant where the agent's output materially affects a customer.

ICO Guidance on AI and Data Protection

Practical guidance on applying UK GDPR to AI systems, including explainability and fairness expectations.

EU AI Act

Directly relevant for firms operating in the EU; indirectly relevant for UK firms whose vendors operate in the EU. Risk-tiered framework with phased implementation through 2026 and 2027. Confirm current implementation status with Legal.

DORA: Digital Operational Resilience Act

EU regulation on operational resilience for financial services, including ICT third party risk. Applies extraterritorially in some cases; confirm scope with Legal.

07 What good looks like: a maturity model

Use this to self-assess where your firm is, and to set realistic expectations with your sponsors. Most firms are at stage two, occasionally stage three.

STAGE 1

Exploring

Conversations happening in pockets. Maybe a few workshops, a handful of POCs in safe environments. No coordinated programme, no executive owner, no governance framework. Investment is opportunistic.

Tell-tale signs: the term "AI strategy" being used aspirationally; no named accountable senior leader; nothing in production.

STAGE 2

Piloting

One or two use cases moving toward production. A nominated owner, often a transformation lead or innovation function. Risk and Compliance engaged but not yet comfortable. First objections being worked through.

Tell-tale signs: a named pilot, a steering group, frequent re-litigation of the same Risk questions.

STAGE 3

Deploying

First use case in production, or close to it. Operating model decisions being made (often painfully). Governance framework drafted. Second wave of use cases beginning to queue behind the first.

Tell-tale signs: an agent live with real customers or colleagues; a documented governance framework; clear ownership.

STAGE 4

Scaling

Multiple use cases in production. Governance functioning. A platform or set of patterns enabling reuse. Investment moving from exploration to scaling.

Tell-tale signs: shared infrastructure, repeated patterns, declining marginal cost per new use case.

STAGE 5

Embedded

AI agents are part of how the firm operates. Build vs buy decisions made on each use case rather than each programme. The conversation has shifted from "should we" to "where next."

Tell-tale signs: AI capability is part of operating reviews; new business cases default to considering an AI component; the firm has talent depth, not just a single team.

08 Where to go from here

FORTAY'S VIEW

Most of the value compounds at stage four, but most firms get stuck at stage two. The blocker is rarely technology. It is the operating model and governance work that bridges stage two to stage three. That is where this toolkit is designed to help.

If this toolkit was useful, the most valuable next step is to stop reading about AI agents and start working with one.

OUR OFFER TO YOU

fortayconnect

A free proof of concept, built for your use case.

Pick a use case. We will build a working AI agent in a sandbox environment, mapped to your data, your workflows, and your compliance constraints. No fees, no commitment beyond your time to scope it with us.

> **The fastest way to find out what AI agents can actually do for your firm is to see one running on your problem.**

TO GET STARTED, CONTACT:

Mark Taylor
Fortay Connect
mark@fortayconnect.com · 0161 240 3411 · fortayconnect.com

Thanks for being in the room.