Claude deployment for businesses

A practitioner's framework

Extending Anthropic's playbook for regulated and risk-conscious organizations

Bianca Janse van Vuuren

May 2026

Executive summary

Anthropic's recent industry-specific guides, including the financial services deployment guide published in May 2026, are useful introductions to the Claude product surface and the customization layer of connectors, skills, and plugins. The guides recommend a three-phase adoption roadmap built around motivated champion teams.

For businesses operating in regulated industries, this recommended approach leaves four meaningful gaps. The discovery method assumes champion teams will surface use cases, which works in motivated functions but breaks down in compliance, operations, finance control, and back-office work. Risk is treated as a parallel concern rather than the sequencing principle that orders deployment. Governance instruments such as the delegation of authority matrix sit separately from the AI deployment, creating a parallel regime. Audit trail design is left implicit, which leads to expensive IT-based audits later.

This framework closes those gaps with seven layers, organized into a five-phase deployment lifecycle. The discovery and sequencing layers replace champion-led use case selection with a waste-focused process inventory and a multi-axis risk tier. The adoption, knowledge capture, and deployment methodology layers handle the activation of non-champion functions, the encoding of senior judgment into reusable skills, and the conversion of generic agent templates into entity-specific tools through observed practice. The governance and accountability layers fit Claude managed agents into existing controls and design audit trails that pass external review without inflating IT audit costs.

The framework is built on deployment experience in financial services, the hardest case for AI governance. It generalizes to insurance, healthcare, utilities, and any business that has to defend its AI programme to a board, an internal audit committee, a model risk committee, or an external regulator.

How to use this document. Read this summary and the seven-layer overview on the next page. Use Section 2 to assess Anthropic's approach against your organization's controls environment. Move to Section 3 to design your own framework layer by layer. Section 5 provides a worked example of what a sequenced deployment looks like in practice.

The seven-layer framework

Each layer addresses a specific question that an enterprise rollout must answer. The layers run in roughly the order they become relevant. Several operate in parallel and reinforce each other.

01	**Waste-focused process mapping** Inventory candidate processes by activity-level observation and waste category.	DISCOVERY
02	**Multi-axis risk tiering** Sequence deployment by compliance, data sensitivity, and decision consequentiality.	SEQUENCING
03	**Demonstration-led adoption** Show before asking in functions that lack natural AI champions.	ACTIVATION
04	**Skills as senior-knowledge capture** Encode senior judgment in a versioned skill register, with senior-staff KPIs.	KNOWLEDGE
05	**Delegation matrix integration** Name managed agents in existing controls instruments to avoid a parallel regime.	GOVERNANCE
06	**Audit trail design** Workflow-level logging that auditors follow without examining model mechanics.	ACCOUNTABILITY
07	**Cowork-first, agent-second** Build entity-specific tools through Cowork observation before deploying managed agents.	DEPLOYMENT

The five-phase deployment lifecycle in Section 4 organizes these layers into Foundation, Discovery, Customization, Governed deployment, and Scale, with success signals at each phase.

Foreword

This document sets out a practical framework for businesses deploying Anthropic's Claude across knowledge work. It builds on Anthropic's own guidance, specifically the May 2026 publication “Claude for the Financial Industry: A Practical Deployment Guide,” and extends it for regulated firms and any organization that has to defend its AI deployments to a board, internal audit, or external regulator.

Anthropic's guide is a strong starting point. It explains the product surfaces, introduces the customization layer of connectors, skills, and plugins, presents ten pre-built agent reference architectures for financial workflows, and offers a three-phase adoption roadmap. Where the guide is strong, this document acknowledges it and builds on it. Where the guide is light, the framework here proposes an alternative. The gaps addressed are five: how to surface use cases in functions that lack natural AI champions; how to sequence deployment by risk rather than enthusiasm; how to encode senior judgment into reusable artifacts; how to fit autonomous agents into existing governance instruments; and how to design audit trails that pass external review without inflating IT audit costs.

The financial services industry serves as the primary illustrative case throughout. The reasons are practical. The deepest examples exist there, and the disciplines required in finance (process discipline, risk tiering, delegation governance, audit trail rigour) are the same disciplines that less-regulated industries are adopting as their AI deployments mature. A framework built for the harder case generalizes downward more easily than the reverse.

This document is written from a practitioner's perspective rather than a vendor's. It is intended for executives, AI governance leads, transformation programme owners, and compliance leads deciding how to deploy Claude in a way they can defend and scale.

1. The Anthropic starting point

Anthropic's guide frames Claude as three product pillars. Chat for conversational drafting and research. Cowork for multi-step work that crosses files and applications with humans in the loop. Code for engineering and quantitative work. Alongside the pillars sit Claude for Microsoft 365 (add-ins for Outlook, Excel, PowerPoint, and Word, plus a connector to SharePoint, OneDrive, and Teams), the Claude Platform for firms building their own applications, and Claude Managed Agents, a cloud-hosted runtime for agents deployed as managed services with scoped permissions, credential vaults, and execution tracing built in.

On top of the products sits a customization layer. Connectors give Claude access to firm data over the Model Context Protocol. Skills are reusable, encoded workflows that teach Claude to perform specific tasks consistently. Plugins bundle skills, subagents, and connectors into installable packages. Anthropic has also shipped ten pre-built agent reference architectures for financial work (pitch builder, meeting preparer, earnings reviewer, model builder, market researcher, general ledger reconciler, KYC screener, month-end closer, statement auditor, and valuation reviewer), which firms customize to their own templates, data sources, and review steps.

The guide proposes a three-phase adoption roadmap. Lay the foundation with access setup, governance scoping, champion teams, and connectors. Pilot with real workflows, defined success metrics, weekly check-ins, and time saved as the primary measure. Scale through admin-provisioned plugin marketplaces, skills that compound across teams, and faster onboarding. Industry use cases are illustrated across investment banking and private equity, wealth and asset management, retail and commercial banking, and insurance. Customer references include AIG (underwriting review timelines compressed more than 5x, data accuracy from 75 percent to over 90 percent), Commonwealth Bank of Australia (fraud prevention and customer service), IG Group (70 hours per week saved in the analytics function, full ROI within three months), and Moody's (credit memo preparation reduced from 40 hours to 2 minutes).

This is a useful, accurate, well-organized starting point. The product matrix is clean. The customization framing is sound. The case studies are credible. The cold-start insight, that pre-configured plugins must produce something useful within the first ninety seconds of a user's first session, is one of the sharpest pieces of practical advice in the document. Any business considering Claude should read the guide before going further. The remainder of this framework assumes that grounding.

2. Where the Anthropic approach needs strengthening

Anthropic's guide is an introductory product playbook. It is not, and does not claim to be, a deployment framework for regulated industries. Four gaps emerge on careful reading. None is fatal. All matter for any business moving from interesting pilots to enterprise scale.

Gap 1: Discovery is bottom-up by default

The guide proposes that champion teams surface use cases that matter to their function rather than having priorities prescribed from above. The reasoning is sound, and in motivated, AI-curious functions it works. It breaks down in three settings. In administrative and back-office work where AI literacy is low and stability is rewarded over experimentation. In heavily regulated functions where employees may perceive AI involvement as a personal risk rather than an opportunity. In any firm that needs a documented, defensible logic for why it started where it started. Champion-driven discovery also produces a deployment portfolio biased toward enthusiastic teams rather than high-value processes; the two correlate but are not the same thing.

Gap 2: Risk-based sequencing is missing

The guide treats risk as something to manage in parallel with rollout, scoping model risk management, SEC/FINRA alignment, and data-privacy review alongside access setup. It does not propose a structured way to rank candidate processes by risk and deploy in increasing risk order. For regulated firms the sequencing question is decisive. The credibility of the entire AI programme rests on the first few deployments being defensible to compliance, internal audit, the board, and the operating committee.

Gap 3: Governance instruments sit separately from existing firm controls

The guide describes scoped permissions, managed credential vaults, and execution tracing as features of Managed Agents. It does not address how those agents fit into the delegation of authority matrices, model risk management frameworks, and three-lines-of-defence structures that regulated firms already operate. Treating AI governance as a parallel regime creates duplicate machinery, slower decision-making, and an audit trail that compliance functions have no reason to trust.

Gap 4: Audit trail design is left implicit

The guide notes that Managed Agents include execution tracing. It does not address how those traces should be turned into audit-grade records that external auditors and regulators can examine without an IT-based audit of the underlying model behaviour. Firms that defer this design choice end up paying for repeated, expensive audits that produce findings about model drift, version control, and non-determinism, none of which are the right questions to ask. The audit trail design is the single decision that separates clean external review from costly remediation later.

> *Anthropic's guide is the right starting point. It is not yet a sufficient framework for regulated deployment. The seven layers that follow are designed to close those four gaps without contradicting the product guidance.*

3. The framework: seven layers

The framework here treats Claude deployment as a layered system. Each layer addresses a specific question that an enterprise rollout must answer. The layers run in roughly the order they become relevant, though several operate in parallel and reinforce each other.

Layer 1: Waste-focused process mapping

Before selecting use cases, the firm needs to inventory the processes that are candidates for Claude support. The most defensible way to do this in a knowledge-work organization is activity-level process mapping drawn from Lean practice and from William Heitman's work on knowledge-work productivity. The discipline asks the firm to observe what people actually do at the activity level rather than at the role level, and to categorize the time spent against the eight categories of waste: rework, waiting, over-processing, defects, motion (context-switching between systems), overproduction (drafting documents nobody reads), transportation (moving information between systems and people), and underutilized talent (senior people doing junior work).

This matters because Claude's capability surface maps almost directly onto these waste categories. Drafting compresses overproduction and motion. Reconciliation captures defects and rework. Search and synthesis collapse waiting. Skill-encoded senior judgment liberates underutilized talent. When the process mapping output is framed as a waste inventory rather than a process inventory, the firm gets a prioritization signal that is hard to argue with. This activity wastes X hours per analyst per week, here is the Claude-shaped intervention, here is the projected recovery. Compliance, finance, the board, and the operating committee can all read that artifact the same way.

Anthropic's guide proposes that champions surface use cases. The framework here is more deliberate. The firm produces a documented waste inventory across in-scope functions, and that inventory becomes the candidate list. Champions are still recruited, but they drive deployment of vetted candidates rather than invent the candidates themselves. The start is slower; the resulting portfolio is much easier to defend.

Layer 2: Multi-axis risk tiering

Once the candidate list exists, it needs to be sequenced. Risk tiering is the sequencing principle. A single risk axis is insufficient. Regulated firms need at least three dimensions running in parallel.

Compliance and regulatory exposure captures fair lending, fiduciary duty, MNPI handling, suitability, and model risk. Data sensitivity captures PII, client confidential information, and market-moving material. Decision consequentiality captures whether Claude is drafting work product that a human reviews, or

producing outputs that are acted on directly. A fourth axis, human-in-the-loop necessity, is sometimes treated separately and sometimes embedded as a control on consequentiality. Either framing works, provided the firm applies it consistently.

The output of the tiering exercise is a tiered backlog. The top tier (low-risk, repeatable, high-waste) is where deployment starts. These activities are also where Claude's value is easiest to measure, with clear before-and-after time savings, and where governance overhead is lightest. Proving value at this tier builds the governance muscle the firm will need to deploy further up the risk ladder with confidence.

Ownership matters. If the tiering sits in technology alone, compliance will reopen every decision later. If it sits in compliance alone, the rankings drift conservative and nothing ships. A working group combining first-line operations, second-line compliance and operational risk, internal audit representation, and a technology lead is the workable form. The tiering is a living artifact. Processes move up tiers as the firm's confidence in Claude grows on adjacent work.

Layer 3: Demonstration-led adoption for non-champion functions

Champion-driven discovery works where motivation exists. In many administrative and back-office functions, motivation is absent. People have built deep expertise in specific workflows over years and have no prior reason to believe Claude can help them. Asking those teams to imagine AI applications they have never seen produces confusion or quiet resistance.

The alternative is to show before asking. The firm builds a working demonstration from the team's own pain points, drawn directly from the waste inventory produced in Layer 1, and runs the demonstration on their own data, in their own tools, on their own screens. A reconciliation analyst sees Claude in Excel investigating a real break from last month. A KYC officer watches Claude in Cowork assemble an entity file from documents that are already on the share. The shift from “look at this new AI tool” to “watch what happens when I ask your spreadsheet this question” removes the cognitive cost of imagining how AI fits into the day. Champions emerge from the people who lean forward during these sessions.

Two design notes. The demonstration must feel like a working session rather than a vendor pitch. The moment it looks curated, the audience disengages. And the sequence is dictated by the waste inventory's highest-pain items, not by what looks easiest to demo. Solving a real problem the team lives with every day is the trigger that converts skepticism into engagement.

Layer 4: Skills as senior-knowledge capture

Skills are reusable, encoded workflows. The Anthropic guide describes them accurately as a way to make Claude perform specific tasks consistently. Their organizational potential goes further than the guide

develops. Skills are the first artifact I have seen that has a real chance of cracking the knowledge-transfer problem that has defeated wikis, SOPs, and training documents for decades. They are invoked at the moment of work rather than read in advance. They capture checklist logic rather than describing the activity. They are exercised constantly, so staleness surfaces in weeks rather than years.

The framework here treats skill authorship and refinement as an explicit performance objective for senior practitioners. Most knowledge-management initiatives fail on incentives. Senior people are rewarded for personal output. Making the team around them better is rarely measured, so it rarely gets done. Reframing knowledge capture as a leadership deliverable, with measurable adoption and outcome metrics rather than volume metrics, produces a different result. Four measures hold up in practice: skills authored, active adoption by junior staff, quality of work produced with versus without the skill, and refinement cadence in response to edge cases.

Each skill needs a named senior owner, a review cadence, and sign-off before any material change goes live. The firm maintains a skill register, a catalog of skills with owners, adoption metrics, last refinement date, and links back to the process the skill operationalizes and the controls it sits inside. Over time the register becomes a strategic asset. The firm's distinctive deployment of Claude is the skill library it has built. The underlying product is the same one every other firm has access to.

> *Pre-authored skills become the channel through which senior expertise reaches junior staff at the moment of decision. This is institutional learning at the speed of software.*

Layer 5: Delegation matrix integration

Regulated organizations already operate delegation of authority matrices that specify who can perform, approve, and review what activities at what thresholds. These matrices are board-approved, audited, and written in the language regulators understand. The framework here uses them.

Where a Claude managed agent performs a delegated activity (meaning it actually performs the activity rather than producing work product for a human to act on), the agent is named in the matrix as the responsible actor. Its scope of authority is spelled out in machine-checkable terms, with escalation rules and evidence requirements attached. The existing controls infrastructure absorbs the new actor type. The same audit, sign-off, and exception-handling machinery already in place extends to cover agent activities. A regulator or external auditor reading the matrix sees the agent's authority and limits in the format they already know how to interpret.

The framework here distinguishes two patterns. Agents that produce work product reviewed and acted on by a human are governed by the human's existing delegation; no new matrix entry is required. Agents

that perform an activity within their own scope (month-end reconciliation that posts journal entries within tolerance, valuation review that approves comparables within methodology) need explicit matrix entries with named owners, escalation paths, and tied controls. The risk tier from Layer 2 determines which pattern applies.

One practical note. Matrix amendments in most regulated firms require committee or board approval, which is intentionally slow. The workable pattern is a one-time amendment at the framework level that recognizes “managed AI agent” as a valid actor type with defined tiers and required controls. Individual agent deployments within an existing tier are then authorized by a delegated AI governance committee rather than the full board. The principle has board approval; the operational decisions move at the speed of the programme.

Layer 6: Audit trail design

The single decision that separates AI deployments that pass external audit cleanly from those that get tied up in expensive review cycles is how the audit trail is designed. The principle is simple. The audit trail describes the activity and its controls. The mechanism stays out of it. Auditors of human-performed processes never ask how the analyst's brain works; they ask whether the procedure was followed, whether the review took place, and whether the evidence is retrievable. Managed agent audit trails should follow the same pattern.

The log entries that matter for audit are workflow-level. What activity was performed. What inputs went in. What output came out. Who or what reviewed it. What escalations triggered and how they were resolved. Which version of the skill governed the activity. Where the entry sits in the delegation matrix. Token-level mechanics (prompts, model versions, temperature, latency) should exist somewhere for engineering troubleshooting, but they belong in a separate stratum from the audit trail proper, and ideally are not visible to the external auditor at all. Conflating the two creates the IT-based audit trap, in which external reviewers spend their time examining model behaviour and producing findings the firm cannot easily remediate.

Designing this upfront forces decisions that are expensive to make later. Retention period and storage location need to match the firm's existing records-retention schedule. Access controls on the audit log itself need to be tighter than access to the underlying process, because the log aggregates sensitive information that would not otherwise sit in one place. There is also a useful separation between the internal audit trail used by the firm's controls function, which can be rich, and the external view shared with regulators or external auditors, which is a structured summary demonstrating control operation without exposing every input and output. The Managed Agents execution tracing Anthropic ships provides the raw substrate. What the firm builds on top of it is what makes the trail audit-grade.

The end state is a chain that can be walked end to end: process, risk tier, delegation matrix entry, skill version, audit trail, evidence. A regulator sampling a delegated decision can move along that chain without needing a guide.

Layer 7: Cowork-first, agent-second deployment methodology

Anthropic's ten pre-built agent reference architectures are templates that firms are expected to customize. The guide says little about how the customization should happen, which is exactly the phase where firms either over-build (months of upfront analysis with no value flowing) or under-build (deploying the template as-is and producing generic output that loses credibility with the business). The framework here closes that gap with a sequencing principle. Cowork first. Agent second.

The principle is that a managed agent codifies a process, so the firm has to know the process first, including its firm-specific quirks, exception paths, escalation logic, and judgment calls that no generic template captures. The right discovery environment is Cowork with the relevant plugin, where senior practitioners run the actual process on actual cases with Claude assisting in real time. Every correction the practitioner makes to Claude's output is signal. Every piece of context the practitioner adds (which document gets pulled in, which rule applies in this scenario, which exception goes to which approver) becomes raw material that turns the generic template into the firm-specific tool.

The progression that follows is the autonomy ladder. Human performs the process with Claude assisting in Cowork (no delegation, knowledge building). Human performs with a maturing skill (workflow taking shape, still human-driven). Managed agent performs with human approval on every output (limited delegation, evidence accumulating). Managed agent performs with human sampling (broader delegation, only after the performance distribution is well understood). Each step has a defensible governance position because empirical evidence accumulates along the way.

Two refinements. The discovery phase needs to be time-boxed, with thirty to fifty Cowork-driven runs as a reasonable target, or it becomes a permanent excuse not to deploy autonomously. And the human drivers during the Cowork phase need to be senior practitioners rather than juniors, because the goal is to encode senior judgment into the agent. Juniors driving the discovery produce an agent that captures routine work and misses the judgment work that justified building it in the first place. The senior practitioner running Cowork sessions is, in effect, authoring the skill in real time, which should count toward the knowledge-multiplication objective from Layer 4.

There is an under-appreciated benefit. The Cowork discovery phase is itself revenue-generating work. The team is doing their day job with Claude assistance, capturing time savings immediately while building the discovery dataset. The phase is fundable in a way that a separate AI pilot would not be, and it surfaces edge cases on real cases with real stakes.

4. The deployment lifecycle

The seven layers are conceptually independent but operationally sequenced. The lifecycle below organizes them into five phases. Each phase has an outcome the firm should be able to point to before moving on. None of the phases is one-and-done; the artifacts produced in earlier phases continue to evolve as the programme matures.

Phase	Purpose	Active layers	Key artifacts	Success signal
Foundation	Establish governance frame, technical access, and the controls envelope	Layers 5, 6	Approved AI policy; governance committee; matrix amendment recognizing managed agents; audit trail standard	Board sign-off on framework; technical access live; controls envelope documented
Discovery	Build the candidate list grounded in evidence rather than enthusiasm	Layers 1, 2	Waste inventory; risk-tiered backlog; tiering working group decisions	Documented backlog of candidates with measurable waste estimates and risk tier
Customization	Build entity-specific tooling and the knowledge artifacts that go with it	Layers 3, 4, 7	Skill register entries; Cowork session logs; demonstration recordings; champion network	Senior practitioners authoring skills; time savings documented per session; emerging champions
Governed deployment	Move agents up the autonomy ladder with evidence at each step	Layers 5, 6, 7	Matrix entries per agent; audit log substrate populated; tier-up decisions documented	Agents performing delegated activities with evidence; clean external audit on first review
Scale	Compound the skill library and propagate the framework across functions	All layers compounding	Plugin marketplace; expanded skill register; onboarding integration; reuse metrics	Adoption metrics across functions; skill reuse across teams; declining support burden

Foundation

Foundation is heavier in regulated firms than in unregulated ones. The technical access work (Claude on the first-party API or via Amazon Bedrock, Google Vertex AI, or Microsoft Foundry, with SSO, SCIM, audit logs, and custom data retention configured) is the lighter lift. The governance work is the longer pole. Three artifacts need to be in place before significant deployment activity begins. The framework-level

matrix amendment recognizing managed agents as a valid actor type. The appointment of a delegated AI governance committee. The audit trail standard that all subsequent deployments will follow. Without these, the programme accumulates governance debt that becomes very expensive to pay down later.

Discovery

Discovery is where the framework's most distinctive choice happens. Rather than recruiting champions to surface use cases, the firm produces a waste inventory across in-scope functions and tiers the resulting candidates by risk. The exercise typically takes six to twelve weeks per business line if done thoroughly. The output is a prioritized backlog that the deployment team will draw from for the next twelve to twenty-four months, with a refresh cycle to capture new processes and reassess existing tiers as confidence grows.

Customization

Customization is where most of the day-to-day work of the programme happens. Demonstration sessions run in non-champion functions to convert skepticism into engagement. Skills are authored by senior practitioners using the Cowork-first methodology, with each session adding to the institutional knowledge captured in the skill register. The pre-built Anthropic agents are reverse-engineered into Cowork-driven workflows, exercised with real cases, and progressively encoded into firm-specific skills.

Governed deployment

Governed deployment moves activities up the autonomy ladder. The first move from Cowork-with-human to managed agent with human approval on every output attracts the most scrutiny, because it is the first time the agent acts within its own delegated scope. Subsequent moves to sampling-based review follow a documented tier-up decision supported by evidence from the previous tier.

Scale

Scale is the longest phase and the one Anthropic's guide describes most clearly. Plugins move into an admin-provisioned marketplace. The skill library compounds across teams as skills built for one function are adapted for adjacent work. New hires onboard onto encoded workflows from day one. The framework continues to operate (new processes enter the waste inventory, get tiered, get deployed through the same lifecycle), and the marginal cost of each new deployment drops as the substrate matures.

5. Worked example: financial services

The framework is best understood through a worked example. Consider a mid-sized asset manager planning to deploy Claude across investment research, compliance, and operations. The firm has read Anthropic's guide and is considering the ten pre-built agents, but compliance has flagged that direct deployment of the templates is unlikely to clear internal review. The framework below provides the deployment programme. The figures used in this example are illustrative.

Foundation

The firm stands up Claude on Amazon Bedrock, configures SSO and SCIM through its identity provider, and sets data retention to align with its existing records management schedule. In parallel, a working group drafts an amendment to the delegation of authority matrix recognizing “managed AI agent” as a valid actor type with three tiers. Assist (Claude produces work product for human review). Limited delegation (agent performs activity within tightly scoped parameters with human approval on every output). Supervised autonomy (agent performs activity with periodic human sampling). The audit trail standard specifies workflow-level logging, retention matching the activity's records schedule, and tighter access controls on the log itself. The board approves the framework. An AI governance committee is appointed, composed of business operations, compliance, operational risk, internal audit, and technology.

Discovery

Process mapping is conducted across investment research, compliance, and operations. The waste inventory surfaces several high-pain candidates. Earnings transcript review absorbs ninety minutes per company per quarter on the equity team. KYC entity file assembly takes three to six hours per new entity in compliance, depending on jurisdiction. Month-end portfolio reconciliation takes operations forty hours per cycle to resolve breaks. Each candidate is tiered. Earnings transcript review lands at low-risk, repeatable, high-waste, top of the queue. KYC sits at medium risk with high data sensitivity, second wave. Month-end reconciliation is high-waste but the decision consequentiality is significant (journal entries are posted) and is deployed under limited delegation only.

Customization

Demonstration sessions run in operations and compliance, where champions have not emerged naturally. The compliance team watches Claude in Cowork assemble a real entity file from documents in their share, completing in ten minutes what their team has been doing in three hours. Two members of the compliance team volunteer to drive the Cowork phase. The earnings transcript reviewer agent (one of Anthropic's pre-built templates) is reverse-engineered into a Cowork workflow run by two senior analysts on the equity research team. Over forty sessions they refine the firm's specific approach to flagging thesis-

relevant changes. A skill is authored in parallel and registered. Both senior analysts have skill authorship as a Q3 objective.

Governed deployment

The earnings transcript skill graduates to a managed agent operating in tier 2 (limited delegation, human approval on every output). The matrix is amended to name the agent as the responsible actor, with the senior analyst as the named approver. After ninety days the agent moves to tier 3 (supervised autonomy with weekly sampling) based on documented evidence of approved outputs without material change and no flagged compliance issues. KYC follows the same pattern at tier 2 but does not graduate to tier 3 in the first year. Month-end reconciliation remains in tier 2 indefinitely.

Scale

The earnings transcript skill is adapted for fixed-income credit research, where the structure of the work is similar. The KYC skill is extended to a periodic-review variant used by client servicing. A plugin bundling the firm's reconciliation, KYC, and earnings review skills is admin-provisioned to all relevant functions. New analysts onboarding into investment research find pre-encoded workflows for the routine work, which shifts their first six months away from process mechanics and toward judgment development.

> *Result: Claude deployed across three functions with evidence-backed governance. None of its agents triggered an IT-based audit during its 2027 external review. The skill register holds entries across the three functions, all owned by named senior practitioners. The deployment portfolio is documented and defensible.*

6. When help adds value

Many components of this framework can be executed by capable internal teams. Procuring Claude through the Anthropic API or via a hyperscaler is straightforward. Configuring SSO, SCIM, and data retention is standard enterprise IT work. Reading Anthropic's guide, installing the pre-built finance plugins, and running internal experiments in Chat and Cowork is something motivated teams will do without external support. None of that requires a consultant.

The phases where external support most often adds value are the ones where the framework departs most sharply from a vendor's product playbook. Three areas are worth naming.

Waste-focused process mapping

The methodology is borrowed from industrial engineering and is not native to financial services or other knowledge-work organizations. Done poorly, process mapping produces a wall of detail that nobody acts on. Done well, it produces a waste inventory that immediately suggests the deployment sequence. The difference is method and discipline: observation protocols, activity decomposition, waste categorization, and the facilitation skill to keep an interview from drifting into an aspirational description of the process rather than the actual one. This is where external facilitators typically pay for themselves several times over.

Demonstration design for non-champion functions

Demonstrations that convert skeptics into champions look easy from the outside and are difficult in practice. They require working examples built from the team's own pain points, on their own data, in their own tools, delivered in a way that feels collaborative rather than scripted. They also require enough familiarity with Claude's capability surface across products to choose the demonstration that lands. This is craft work, and it is the highest-leverage activity in the programme. A well-designed demonstration in compliance or operations can move a function from “not interested” to “when can we start” in a single afternoon.

Governance integration

Integrating Claude managed agents into existing delegation matrices, model risk management frameworks, and audit trail design is conceptually clear but operationally complex. It requires someone who understands both the Claude product surface and the firm's controls environment well enough to translate between them, with the credibility to work with compliance and internal audit without being treated as a vendor advocate. Most firms benefit from external support during the framework-level matrix amendment and the first one or two governed deployments, after which the internal team usually has enough pattern recognition to continue independently.

Procurement is the easy part of Claude deployment. The product is excellent, well-supported, and well-documented. The hard parts (choosing what to deploy, in what order, with what governance, supported by what knowledge artifacts) are not problems Anthropic can solve for the firm. The answers depend on the firm's processes, risk appetite, controls environment, and culture. That is where this framework, and targeted external support, earns its keep.

Closing note

Anthropic has built a remarkable product surface and a thoughtful introductory deployment guide. The framework here is not a critique so much as a continuation: a more structured and more honest account of what a regulated business has to do to deploy Claude well.

The seven layers are designed to reinforce each other. The discovery layer produces the candidate list. The tiering layer sequences it. The demonstration layer activates non-champion functions. The skills layer captures the firm's distinctive expertise. The governance and audit layers make the deployment defensible. The deployment methodology turns generic templates into entity-specific tools through observed practice rather than upfront design.

Businesses that approach Claude deployment this way end up with two things they did not have before. A programme they can show to a board, an internal audit committee, a regulator, or an external auditor, and have those audiences understand it. And a skill library owned by senior staff whose expertise is being multiplied across the organization rather than locked in their heads.

Bianca Janse van Vuuren · Claude deployment framework · May 2026

References and further reading

Anthropic, “Claude for the Financial Industry: A Practical Deployment Guide,” May 2026. The introductory product guide this framework extends. Covers the product pillars, customization layer, ten pre-built financial-services agent reference architectures, and Anthropic's three-phase adoption roadmap.

William F. Heitman, “The Knowledge Work Factory,” McGraw-Hill, 2019. The methodology underpinning Layer 1. Applies industrial-engineering process discipline to knowledge work, with particular attention to surfacing the invisible waste that accumulates in professional-services and back-office functions.

Anthropic press releases and customer stories cited in the Anthropic guide: AIG underwriting transformation; Commonwealth Bank of Australia strategic partnership; IG Group analytics function case study; Moody's MCP application for credit and compliance. These are useful as proof points but should not be confused with deployment guidance. They describe outcomes, not method.

U.S. Federal Reserve Board, SR 11-7, “Guidance on Model Risk Management,” 2011 and subsequent updates. The foundational U.S. regulatory framework for model risk in banking. Relevant to Layer 5 (delegation matrix integration) and Layer 6 (audit trail design) for any firm with U.S. banking exposure.

EU AI Act (Regulation 2024/1689). The European Union's framework for AI governance, with risk-tiered obligations that map onto Layer 2. Relevant for any firm with EU operations or EU customers.