The AI investment firm

How leading firms are rebuilding around AI

Claude

Table of contents

Foreword	3
Advancing your firm's AI maturity	4
The AI investment firm in practice	6
The new operating model	8
Key implementations by strategy	10
How the work is controlled	22
The Claude stack on the buy side	24
Resources	27

Foreword

This guide is for investment and trading firms: hedge funds, systematic managers, market makers, digital-asset firms, private equity and private credit firms, venture capital firms, wealth managers, and financial advisors. The firms furthest along on the AI adoption curve have moved from piloting AI at the edges of their processes to rebuilding their operating model: how the work is divided, who does it, and how it is controlled.

When Claude takes the first pass on routine, time-consuming work like updating models after earnings or reconciling positions and cash, people can spend more of their time on decisions that determine returns and on carrying them out. Getting there is a people and process change as much as a technology change. Employees can take on work they used to delegate or never had time for. Firms add roles they didn't need before, such as embedded AI engineers who sit with the investment teams and build agents for the work analysts want to hand off, and owners for the skills Claude follows and the evals that test its work

In this guide, we draw from our work with buy-side firms to outline best practices for AI adoption, and the work Claude takes on for each team.



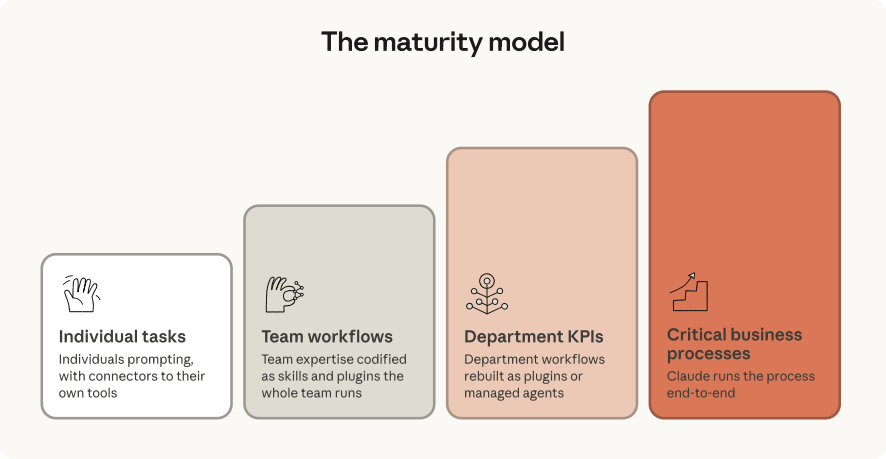
Chapter 1

Advancing your firm's AI maturity

Chapter 1

Advancing your firm's AI maturity

Two investment firms can equip the same number of people with access to Claude and see different results. What an analyst gets from Claude depends on what the firm has built around it: which systems Claude can read and write, how much of the work it is trusted to finish, and whether the team's best methods are written down where everyone can run them. In the deployments we see, impact grows with the complexity of the work Claude can do, and that complexity grows with the access and trust the firm has extended to Claude. The maturity model below makes the progression visible: where your firm stands today, and what it needs to do to get to the next level.



Individual tasks. People prompt Claude on single tasks, with connectors to their own tools. To move up, senior people write the team's methods down as skills.

Team workflows. The team's expertise is written as skills and plugins that everyone runs. To move up, the firm connects Claude to the team's systems, records baselines, and builds evals from the team's real work.

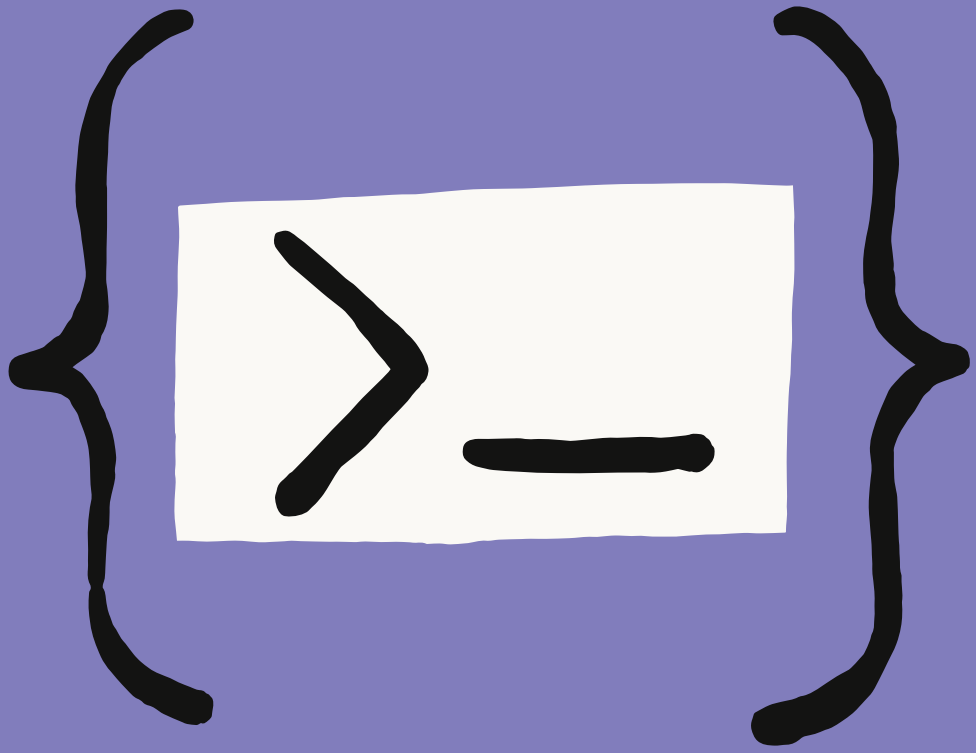
Department KPIs. Departments rebuild workflows as plugins or managed agents tied to a KPI, and Claude does the work while a person checks it. To move up, the evals show the agent meets the owner's standard, and governance covers who publishes skills, what data each workflow reaches, and who owns the results.

Core business processes. Claude runs the process, and a person reviews the exceptions and approves the results before they're acted on.

Firms hand work to Claude in stages. Once the evals indicate that the agent's work has met the standard set by the owner, each process moves to the next stage. In every example we outline in this guide, the assumption is that a person reviews and approves the work before it goes to a client, gets filed, or is acted on.

Most firms start by giving employees a chat assistant, which they can use for drafting emails or doing research. The firms that move up the ladder connect Claude to the systems their teams use and trust it with more complex work. Employees build fluency, from using Claude on single tasks to running several workflows in parallel. At the higher levels, departments rebuild workflows as plugins or managed agents tied to a KPI. The end state is Claude running a process end to end, with a person checking the exceptions and a sample of the rest.

Firms need to establish governance that supports who can publish skills, what data each workflow can reach, and who owns the results. This investment continues throughout the life of the program.



Chapter 2

The AI investment firm in practice

Chapter 2

The AI investment firm in practice

Most investment firms now use AI. The firms getting the most from it organize their work around six best practices.

A firm already has an edge in its proprietary data, house methods, and the judgment of its senior people. But senior people only have so many hours. When they write down how they work, Claude can apply their methods to more names, deals, and decisions than they could cover alone, and draw on more sources for each one. To do that, the firm needs connected data, methods written as skills, evals that score the agent's work, and a memory of past work. All of it belongs to the firm, and teams refine it as they work. Each new model starts from everything the firm has built.

Practices at leading firms

Record baselines

Before an agent takes on a core process, the firm records how long the work takes the people doing it today and how often they get it right, with measures like days to close the books or hours to get from an earnings print to a position.

Test agents on the team's work

For each core process, the team builds a test set from its real work and scores the agent against a human baseline or a desired output. These scored tests, called evals, show when an agent can be trusted with more and which model should run a task. When a new model comes out, the same evals show whether it's ready to adopt.

Write down methods as skills

Senior people write down how they build a model or clear a break as a skill (a set of instructions that Claude follows much like an employee would). Once a skill is written down, the whole team can work from it and the firm can improve it over time. Because the skills encode the firm's best practices, junior employees don't have to start from zero and senior ones spend more of their time on judgment calls and coaching than on checking their work.

Redesign roles for human-agent work

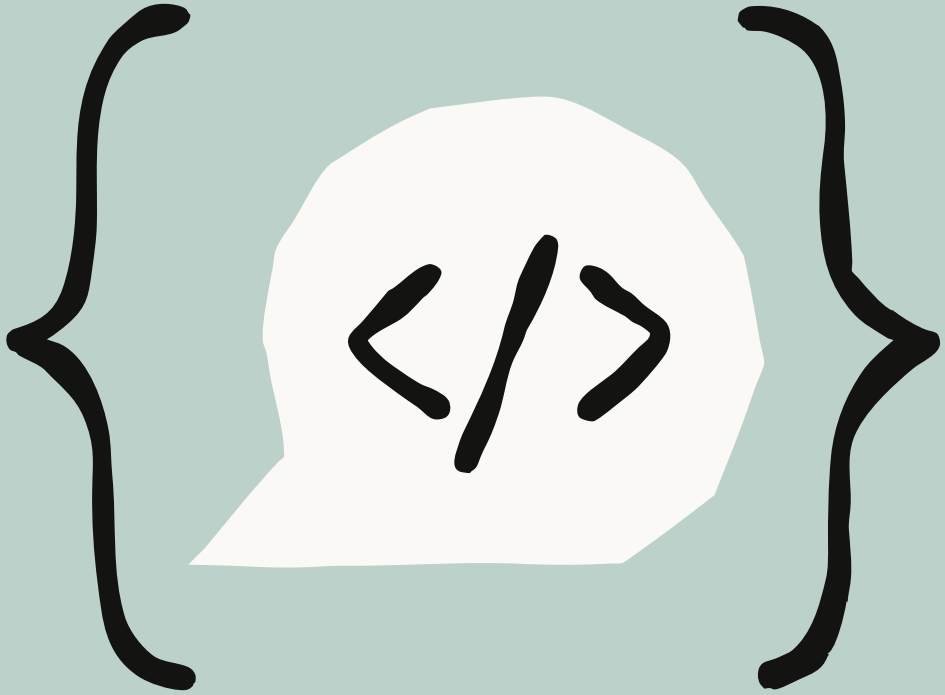
Leading firms expand existing roles and add new ones, such as skill authors, eval owners, and agent operators, and they rethink how junior staff learn the job. Because an agent can take a task through several steps, handoffs between people, within a team or between teams, become simpler.

Have Claude write the code for calculations that must be exact

Some tasks are rule-based calculations with exactly one right answer, such as whether a position breaches a risk limit or how much an investor owes in fees. These tasks run as deterministic code, so the same data always gets the same answer, and the logic can be tested and audited. Claude can write that program and the tests that check it. When the program flags a problem, such as a limit breach, Claude looks into it and explains which trades caused it.

Switch to a new model once it matches the current one on evals

When a new model is released, firms run the same evals and compare its work with the current model's. They keep skills current and connect Claude to their data providers and internal systems, so these evals run smoothly. If the scores hold, they upgrade to the new model; if the scores drop, they adjust the instructions and skills and test again until the new model meets or exceeds the performance of its predecessor.



Chapter 3

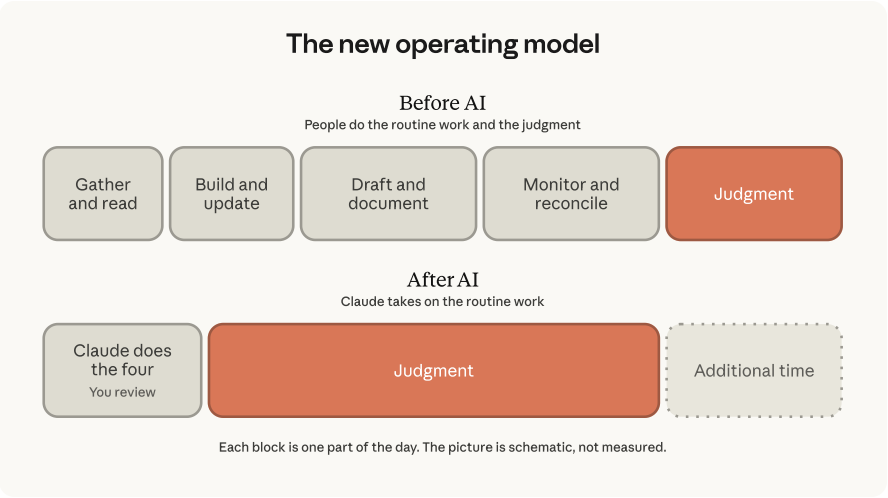
The new operating model

Chapter 3

The new operating model

The same working day, before and after Claude takes on routine work

For any team in a financial firm, the working day includes four kinds of routine work and decisions that need human input. Before AI, people had to do both. Now, Claude can take on the routine work, while people focus on decision-making and work that requires judgment.



A day at any seat at the firm

FRONT OFFICE

SEAT	CLAUDE TAKES ON	STAYS WITH THE TEAM
Research and diligence	Reading filings, model updates, first drafts of notes	Thesis, conviction, management access
Quantitative research	Data exploration, signal code, backtests	Coming up with new ideas to test
Portfolio management and deals	Morning brief, what-ifs, diligence checks	Judgment, market view, which deals to do
Trading and execution	Pre-market prep, parsing dealer runs, fill forensics	Pricing, hedging, when to step in

MIDDLE OFFICE

SEAT	CLAUDE TAKES ON	STAYS WITH THE TEAM
Risk management	Daily commentary, breach notes, scenarios	Which risks matter, when to escalate
Compliance and legal	Alert case prep, first-pass reviews	The call on each case, negotiation
Operations	Break investigation, inboxes, statement checks	Which breaks matter

BACK OFFICE

SEAT	CLAUDE TAKES ON	STAYS WITH THE TEAM
Fund finance and treasury	Close prep, shadow NAV tie-outs, fee checks	Valuation calls, the administrator
Investor relations	Questionnaires, letters, meeting prep	Investor relationships and messaging
Engineering, data, and security	Migrations, tests, code review, alert triage	Architecture, what ships



Chapter 4

Key implementations by strategy

Chapter 4

Key implementations by strategy

The team pages that follow apply to every kind of firm. This page shows what is specific to each kind of firm.

Where to start with Claude depends on where a firm's people spend their hours. Each kind of firm also has processes the others don't.

STRATEGY	WHERE TO START WITH CLAUDE	DISTINCTIVE PROCESSES
Fundamental long/short and multi-manager	Coverage models in Excel, overnight earnings work, the portfolio manager's morning brief.	Skills and memory are kept separate for each pod, so each team's context stays its own.
Quantitative and systematic	Claude Code inside the research codebase.	Agent fleets that propose, implement, and score signals. Papers replicated on internal data.
Market making and high-frequency trading	Engineering, incident investigation, post-trade analysis.	Forensics over logs and tick data. Agents stay out of the latency-critical path.
Digital assets	Engineering, smart-contract review, exchange-versus-chain reconciliation.	On-chain data treated like any other market data. Unsigned transactions decoded and simulated, so the signers can see what they're approving.
Event contracts	Contract resolution, cross-venue pricing, newsflow.	Reading each contract's rules and resolution source.
Private equity	Data-room reading, the deal model, the investment-committee memo.	Portfolio monitoring across companies, and bringing the firm's playbook into the portfolio companies.
Private credit	Financial spreading, the credit model, the credit memo.	Covenant and borrower monitoring across the book, with each amendment and waiver request read against the original terms.
Venture and growth	Sourcing, founder and market diligence, portfolio KPI roll-ups.	Turning sparse, messy company data into a comparable view.
Wealth management and financial advisors	Client review prep, portfolio reporting, meeting notes, and follow-ups.	A brief on each household before the meeting, drawn from holdings, the financial plan, and the CRM.

Front office

Research and diligence

Analysts and deal teams work with Claude as the first reader across their coverage universe, including filings, transcripts, broker notes, expert calls, data rooms, and the firm's own research history. Claude can update models, use connectors to pull the data it needs, and run scheduled tasks overnight.

Main use cases

Updating and building models in Excel

Analysts can ask Claude to roll a coverage model forward from the 10-Q, rebuild a broker's hard-coded sheet as live formulas, or add a sensitivity tab. Deal teams can ask for leveraged-buyout models, credit models, or cap tables, built from confidential information memorandums (CIMs) and data rooms. Claude edits workbooks directly and cites the cells behind each answer.

Processing events overnight

On earnings night, a scheduled agent reads the release, transcript, and filings through connectors. It compares the results with consensus and with the team's model, updates the model, and drafts a first-take note, so the analyst starts the day with a marked-up model and a note to edit. Anthropic's financial services plugins are a starting point for these steps.

Researching the team's archives

Because Claude Enterprise connects to internal document search, vendor data, and the research management system, teams get answers that cite multiple sources. Connectors that each person signs in to use that person's existing permissions, so Claude reaches only the documents that person has access to. Private-markets teams ask Claude to get them up to speed on a new investment and to track portfolio performance.

Conducting deep research, diligence, and red-teaming

For open questions, such as sizing a market or assessing a founding team, an agent breaks the question into subquestions, searches independent sources for evidence, and returns a cited report or a diligence checklist with findings. Before a pitch or an investment-committee meeting, Claude can follow a red-team skill to argue the bear case against the team's own thesis.

Outcome: Analysts and deal teams can cover more names and deals. After an earnings release or another company event, Claude updates the model and drafts a first-take note for the analyst to review before the market opens. Analysts can search the team's past notes and models.

Front office

Quantitative research

Quantitative research is a natural fit for Claude Code because most of a researcher's work is writing and running code, data is structured, and the feedback loop is a backtest on historical data. Researchers work with Claude Code against internal libraries and backtesters, hand it papers to replicate, and increasingly run fleets of agents that generate, implement, and score signals.

Main use cases

Exploring and backtesting signals in the firm's data and tools

Claude Code runs inside the research codebase with a standing brief that describes the data catalog, the backtester, and the coding conventions. At the start of a session, Claude checks a new dataset for gaps, survivorship, and point-in-time integrity. Then it explores feature-engineering ideas, codes the signal, writes the data-quality checks, runs the backtest, and reports Sharpe, drawdown, and turnover in the tear sheet. Researchers often run several sessions in parallel, asking Claude Code for quick answers, as well as for long, careful analysis.

Replicating papers on internal data

When researchers give Claude Code a paper as a PDF, it reconstructs the methodology against internal data, tries to reproduce the result, runs out-of-sample and robustness checks, and writes a verdict with the code ready for review. Researchers used to put off replication until they had bandwidth to take it on; with Claude Code, they receive a first pass the same day they hand over the paper.

Running the research loop with agents

A planner agent proposes hypotheses and researcher agents write and test signal code in isolated copies of the codebase. An evaluator agent scores each signal against the firm's criteria, using multiple-testing controls and a holdout set that the agents can't access. The researchers review the scored signals and decide which to pursue.

Querying an analyst assistant in plain English

Some firms have built in-house analyst assistants on Claude. Analysts ask questions in plain English, and the assistant writes the Python, builds the charts, and iterates on the analysis. Systematic managers pair it with batch pipelines that turn their archives of filings, transcripts, and news into features.

Outcome: Researchers spend more of their time judging which ideas deserve capital, because Claude handles more of the data exploration, feature engineering, and testing. Every trial is logged, so the reported statistics are adjusted for the number of tests run.

Front office

Portfolio management and deals

Portfolio managers, and the deal partners who lead transactions at private-markets firms, work with Claude as a chief of staff for the book or the portfolio. Claude ranks the morning's news by exposure, answers questions about positions and portfolio companies, and runs scenarios in the analyst's model, so sizing and pricing decisions are based on the team's numbers.

Main use cases

Generating a morning brief, ranked by exposure

A scheduled task runs overnight, compiling market updates that affect the book: news, broker commentary, pre-market moves, portfolio-company updates, and model changes. It ranks them by position size and move, and sends a one-page brief to the team's preferred channel, such as email, Slack, or a dashboard, all before the day starts.

Answering questions about positions and portfolio companies

In Claude Enterprise, portfolio managers and deal partners can ask questions in plain English, such as "What's my beta-adjusted net in European cyclicals?" or "Which portfolio companies are behind plan this quarter?" Claude answers using context from the firm's position, portfolio, and risk systems through connectors, inside a project scoped to that team.

Running what-if scenarios in the analyst's model

For a scenario like "gross margin down 200 basis points and the buyback halves" or "exit at 9x in year four," Claude flexes the drivers in the analyst's Excel model and lists exactly which cells changed. The team can check the math cell by cell and focus the discussion on whether the assumptions hold.

Checking diligence files for completeness

Before an idea or a deal reaches committee, Claude reads the associate's diligence file against the firm's checklist and past deals and lists what's missing, thin, or unsupported. The partner's review starts from that list.

Creating dashboards on demand

Portfolio managers ask Claude for interactive dashboards built on the firm's connected data, such as cross-asset moves around an event or a roll-up of portfolio-company metrics.

Outcome: Portfolio managers and deal partners spend the first hour of the day on the brief's top-ranked positions and companies, and Claude answers scenario questions in the analyst's model.

Front office

Trading and execution

On the desk, traders run their standard analyses through skills in a sandbox, agents with read-only access investigate execution anomalies, and parsers turn dealer chatter into structured data.

Main use cases

Running trader analyses in a sandbox

Traders get Claude Code or a lightweight agent app inside a sandbox, with skills for the desk's standard analyses, such as transaction cost analysis (TCA), markouts, P&L explain, event studies, and venue comparison. Claude produces the notebook and the charts.

Investigating execution problems

When fills, latencies, or P&L look wrong, an agent with read-only access to logs, tick data, and configuration gathers the evidence and ranks likely causes. It then proposes a fix as a pull request for an engineer to review.

Answering cross-asset questions through market-data connectors

Traders can ask Claude cross-asset questions in plain language. Claude pulls prices through market-data connectors, runs the calculations in code with the pricing tools, and writes up the result for the trader. For example, a fixed-income desk can ask for bond relative value under a rate shock, a digital-asset desk can ask about funding rates and on-chain liquidity, and an event-contract desk can compare pricing across venues.

Parsing dealer runs, broker email, and chat into structured quotes

Smaller Claude models parse dealer runs, axes, chat, and broker emails into structured quotes and market color, which feed research and the desk's quote history.

Outcome: Traders get same-day answers to cross-asset questions. When fills or latencies look wrong, agents gather evidence from logs, tick data, and configuration, and rank the likely causes, so engineers can diagnose incidents faster.

Middle office

Risk management

Risk teams are building agents that keep context from one day to the next. A typical agent works with the risk managers on the firm's data and helps them explain each day's changes in risk across asset classes.

Main use cases

Drafting daily risk commentary

Each morning, an agent reads positions, sensitivities, and market data, along with the results of the firm's coded checks against its risk limits, investment guidelines, and concentration limits. It explains what changed by desk and asset class, flags positions close to a limit, and drafts the sign-off commentary. It answers risk managers' follow-up questions with the context of prior days, and it logs the reasoning behind everything it produces.

Investigating limit breaches

A program checks the firm's positions against each risk limit and flags any breach or near-breach. For each flag, Claude identifies the trades that caused it, gives a first view of how serious it is, and drafts the escalation for the risk manager to review.

Designing and testing new stress scenarios

Agents propose new coherent scenarios, run each one through the firm's risk engine, and explain which positions drive the P&L. Risk committees review these scenarios next to the standard set that replays past market events.

Validating models and monitoring counterparties

Claude Code, with read-only access, reproduces model results and compares the documentation against the code for validation reports. Scheduled agents watch ratings, spreads, and news on prime brokers, exchanges, custodians, and other counterparties through credit-data connectors.

Outcome: Risk managers review drafted commentary before they sign off, and can trace each explanation to the data as it stood that morning.

Middle office

Compliance and legal

Compliance teams work with Claude to prepare alert queues, review expert-call transcripts, answer policy questions, run a first pass on marketing materials, and track regulatory change. Counsel ask it to review trading and deal agreements and side letters, and to draft answers to the business's questions about fund documents for counsel to review.

Main use cases

Reviewing surveillance alerts and expert-call transcripts

Surveillance systems flag trades and messages that might break a rule, and most flags turn out to be false alarms. Claude reads each flagged item, rates how likely it is to be a real problem, and drafts a recommendation to close or escalate it, with the reasons why. The compliance officer makes the call from that write-up. Claude also flags possible material nonpublic information (MNPI) in expert-call transcripts for the compliance officer to review.

Relaying pre-clearance decisions and answering policy questions

Employees ask Claude, “Can I trade this?” The firm’s pre-clearance system checks the restricted list and holding-period rules and makes the decision, and Claude relays the outcome and logs the exchange. The same assistant can also answer questions about the compliance manual, citing its sources.

Reviewing agreements and tracking investor obligations

Claude checks trading and deal agreements against the firm’s playbook, flags deviations, and drafts redlines. Anthropic’s open-source legal plugin for Claude and Claude Code includes a ready-made command for this review. For side letters and limited partnership agreements (LPAs), Claude extracts the terms into an obligations matrix, which the team uses for most-favored-nation elections, fee checks, and investor reporting.

Keeping records of AI use, reviewing marketing materials, and tracking regulatory change

The firm captures prompts and responses in its compliance archive, through its cloud logging or Anthropic’s Compliance API (on Enterprise plans). Claude gives outbound decks and letters a first-pass check against the SEC Marketing Rule and maps new regulatory releases to the compliance calendar and the policies they affect.

Outcome: Compliance officers can focus on deciding cases instead of assembling them, counsel can spend more of their time on negotiation, and the firm has a record of its AI use.

Middle office

Operations

Operations work suits agents because it's high-volume and rule-bound, and teams already measure it. Agents investigate breaks and watch shared inboxes, and the operations team directs their work. Reconciliation is the usual place to start.

Main use cases

Reconciling positions and investigating breaks

A scheduled agent pulls files from prime brokers, administrators, exchanges, and custodians, or on-chain balances, and code matches the positions, trades, and cash. For each break, the agent checks how similar breaks were resolved before, opens a case, and drafts the follow-up.

Handing shared-inbox triage to an agent

The operations team has an agent watch its shared mailboxes. The agent sorts and routes incoming requests, answers routine internal ones from systems of record, and escalates anything close to its deadline to the team.

Checking reports, notices, and confirms against the firm's records

Claude compares broker and administrator statements against internal numbers and cites each discrepancy by page. It turns notices, confirms, and portfolio-company reporting packs into structured terms, checks them against the firm's records, and flags mismatches for the team.

Briefing the team before market open

By the time the desk arrives, Claude has drafted a one-page brief covering overnight breaks, fails, margin calls, escalations, and the follow-up for each break.

Outcome: The operations team works through the breaks that need judgment, while agents handle routine work like inbox triage and statement checks.

Back office

Fund finance and treasury

In fund finance, good places to start are the month-end close and a shadow net asset value (NAV) calculation checked against the administrator's. Because the work follows set rules in Excel and gets audited, it suits Claude for Excel (alongside the rest of the Claude for M365 suite) with fund-administration skills.

Main use cases

Preparing the month-end close for review

The finance team saves its month-end checklist as a skill. Claude works through it using the firm's books or shadow books and the administrator's extracts, and prepares the accruals, roll-forwards, allocations, draft journal entries, variance commentary, and close report for the team to review. Anthropic's finance plugin is a starting point for journal entries, reconciliations, and variance analysis.

Checking the administrator's NAV

Code reconciles positions, prices, accruals, and fees against the administrator's NAV, and agents propose a root cause for each break. For private holdings, agents assemble the valuation support pack, with comparables and a memo, for the valuation committee. Before distribution, Claude ties every capital-account statement to the NAV pack and flags mismatches.

Recalculating fees, carry, and waterfalls

With side-letter and LPA terms in a structured table, code recomputes management and incentive fees and carry from the fund documents, investor by investor, and Claude explains any differences from the administrator's figures.

Tracking margin and cash, and supporting regulatory filings

In treasury, Claude explains day-over-day margin moves, compares financing terms across brokers, tracks balances across banks, brokers, and wallets, and drafts the daily cash note. For regulatory filings such as Form PF and Form ADV in the US, or Annex IV reports under the AIFMD in the UK and EU, agents assemble the data and draft a comparison with the prior filing for the preparer.

Outcome: The close takes less time, and the team can run shadow calculations every period. Auditors can trace each figure to its source.

Back office

Investor relations

Investor relations and fundraising teams work with Claude to draft due-diligence questionnaires and investor letters, to prepare for meetings, and to produce materials.

Main use cases

Drafting questionnaires from a governed answer bank

Claude drafts each answer to due-diligence questionnaires and requests for proposal into the allocator's template, using approved language and prior responses and citing a source for each answer. It routes any questions without an approved answer to subject-matter experts. Claude handles questions about the firm's own AI governance the same way.

Drafting investor letters and quarterly commentary

Claude drafts investor letters using performance, attribution, portfolio-company updates, and portfolio managers' notes from the reporting pack. It then follows a Marketing Rule skill to give the draft a first-pass check.

Building decks and factsheets

Claude Design (beta) builds factsheets and one-pagers on the firm's design system, and Claude Slides (beta) drafts pitchbooks and annual-meeting decks. Both can be exported to PowerPoint or PDF. Claude can also refresh slides in the existing template directly from the Excel pack.

Preparing for meetings and handling investor requests

Before an allocator meeting, an agent compiles the relationship history, recent communications, and likely questions. Afterward, it logs the notes and follow-ups in the CRM. When investors request exposure cuts, look-through data, or capital-account reports, Claude prepares them in the investor's template using systems of record.

Outcome: Questionnaires and data requests go out sooner, in consistent, sourced language. Teams produce decks and factsheets in-house.

Back office

Engineering, data, and security

Engineering and data teams build the systems the other teams' agents depend on, and they work with Claude more than any other team. They write code with it and rebuild the firm's systems and data so that agents can work with them too.

Main use cases

Running Claude Code under the firm's controls

Developers and quants work in Claude Code with per-user attribution, sandboxing, and a shared library of approved skills. Analysts, operators, and support staff are starting to work in it too.

Building systems for agents as well as people

Most systems of record were built for a person looking at a screen. Now, data teams make tick stores, the security master, positions, the deal pipeline, portfolio-company metrics, and data warehouses available to agents as tools, with access set by team and business unit. They connect Claude to these systems through the Model Context Protocol (MCP), and each agent gets only the access its task needs. New systems are designed so that an agent can work with them as easily as a person can.

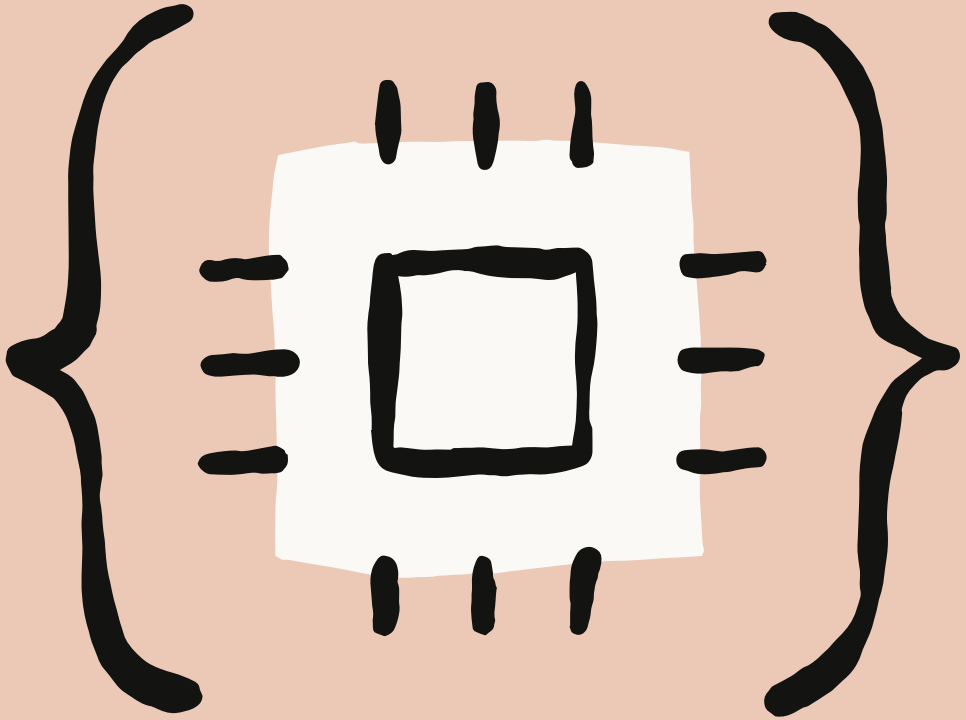
Modernizing legacy systems

Agents work best on systems they can read, so modernization starts with Claude mapping a legacy system and pulling out the business rules buried in its code. From there, the firm can upgrade the system in place, port it module by module, or rebuild it on a modern architecture. On each path, the firm runs the old and new versions side by side to confirm that the outputs match. Refactors can run overnight as fleets of agents, each in an isolated workspace with a budget cap. Anthropic's code modernization plugin packages these steps. More ready-made plugins are in the plugin directory and the financial services and knowledge work marketplaces.

Working with Claude across the software lifecycle

Engineers work with Claude at each stage of the software lifecycle. It writes code along with unit and smoke tests, reviews changes, and checks them for security issues. Security teams work with it in red-team, blue-team, and purple-team exercises, and agents enrich and triage alerts for the security operations team.

Outcome: Engineering teams build more, and junior engineers start from the house standard. The security team can inspect the controls around Claude.



Chapter 5

How the work is controlled

Chapter 5

How the work is controlled

Access follows existing permissions

Connectors that each person signs in to use that person's existing permissions, so Claude reaches only what that person can. Each agent gets only the access its task needs, set by team and business unit. At multi-manager firms, skills and memory are kept separate for each pod.

Exact calculations run as code

Limit checks, fee calculations, and reconciliations run as deterministic code that can be tested and audited. Claude writes the code and the tests, and explains what the code flags.

Agents stay out of production paths

Agents that investigate execution problems work with read-only access and stay out of the latency-critical path. Fixes are proposed as pull requests for an engineer to review.

A person signs off

Claude doesn't provide investment, legal, tax, or compliance advice, and each firm remains responsible for its decisions and regulatory obligations. A person reviews and approves the work before it goes to a client, gets filed, or is acted on. Evals show when a process is ready for the next stage, and when a new model can take over.

Everything is recorded

Prompts and responses go to the firm's compliance archive through its cloud logging or Anthropic's Compliance API (on Enterprise plans). Agents log the reasoning behind what they produce, and Claude Code runs with per-user attribution and sandboxing.

Where to start

Pick one process.

Choose one the team already measures. Reconciliation, the month-end close, and the earnings-night model update are common first choices.

Record the baseline.

Write down how long the work takes today and how often it's right.

Build the eval and write the skill.

Build the test set from the team's real work, and have the people who do the work best write down how they do it.

Connect the systems and name the owners.

Give Claude access to the systems the process touches, and name an owner for the skill, the eval, and the results.

Move up a stage when the evals say so.

Then pick the next process.



Chapter 6

The Claude stack on the buy side

Chapter 6

The Claude stack on the buy side

Different teams or groups within a firm work with different Claude products: Claude Code for everyone who writes code, the Claude app and Claude for Microsoft 365 for investment and operations staff, Claude Enterprise with financial-data connectors for research, and the Claude Agent SDK (software development kit) for the agents the firm builds itself.

Claude products

Claude Code

Engineers work with Claude Code in the firm's codebase from the terminal, a code editor, the browser, or Slack. They ask it to map a repository or modernize legacy code, and it can turn a ticket into a pull request with changes across many files.

Claude and Claude Enterprise

In Claude Enterprise, teams ask questions across the firm's research, data, and documents and get answers with citations. Analysts ask Claude to get them up to speed on a name or draft a first take. Claude can also take on longer tasks as an agent when someone gives it a goal, such as the overnight earnings run or a due-diligence questionnaire. It shows each step as it works across files and the browser, and it can run tasks on a schedule. On Enterprise plans, an owner turns on cloud sessions for the teams that need them, so scheduled tasks can run overnight with no one at a desk. Anthropic's financial services plugins add finance skills and connectors for data providers such as FactSet, S&P Global, LSEG, Morningstar, PitchBook, and Moody's, though each provider may require its own subscription. Other connectors link Claude to Snowflake, Databricks, and the firm's own systems through the Model Context Protocol (MCP).

Claude for Microsoft 365

Claude for Microsoft 365 works inside Excel, PowerPoint, and Word. It builds and edits models while keeping formula relationships intact, cites the cells behind its answers, and traces errors to their root cause. When an analyst moves from the model to a memo or a deck, Claude keeps the context.

Claude Design and Claude Slides

From a person's description, Claude Design builds a first draft of a one-pager or prototype on the firm's design system. For presentations, Claude Slides drafts the deck from notes, reports, or earlier work in a conversation. Drafts from either can be exported to PowerPoint or PDF. Both are in beta, and Enterprise admins choose when to turn them on.

Claude Agent SDK and Claude Managed Agents

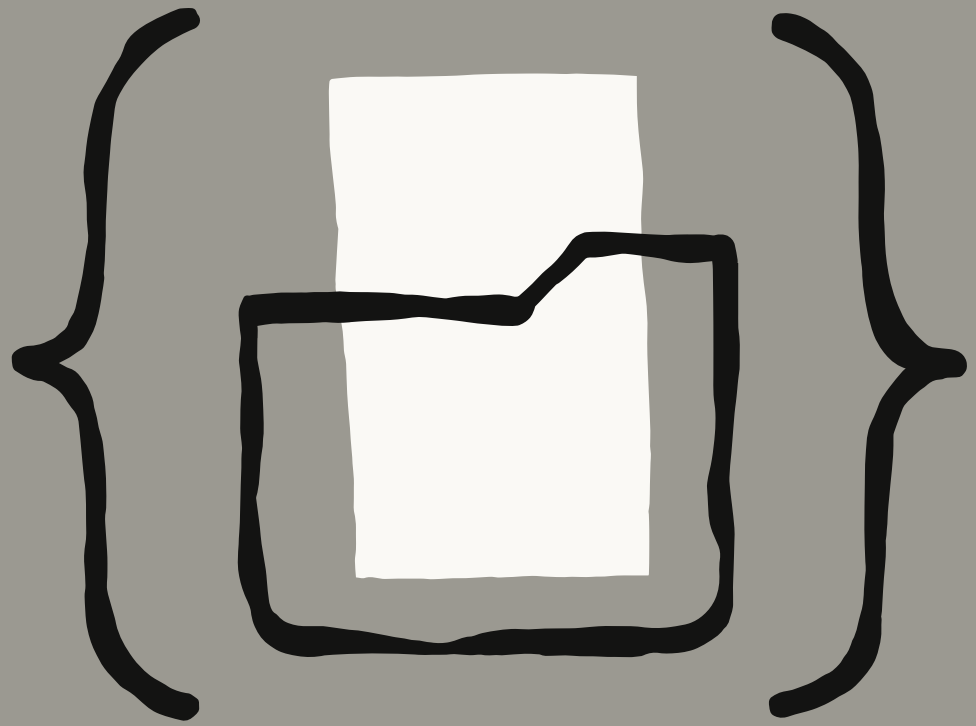
The Claude Agent SDK packages the agent loop behind Claude Code as a library, so the firm's developers can build their own agents for research, risk, reconciliation, or surveillance that work with the firm's data and follow its methods. The firm can run those agents in its own environment or have Anthropic host them through Claude Managed Agents (beta).

Claude Tag: @Claude in Slack

Claude Tag (beta) brings Claude into Slack threads, where several people can work with it at once. Engineers can tag in Claude and the principal engineers involved to triage a break, and an investment team can work through an idea together. Claude reads the thread, answers with the firm's data, takes on longer tasks, and posts updates in the thread.

Open-source plugins

Anthropic publishes plugins for financial services (equity research, financial analysis, private equity, and fund administration), for legal and finance teams, and for code modernization. Each one is a starting point that the firm can adapt to its house methods.



Chapter 7

Resources

Chapter 7

Resources

Webinars and walkthroughs for investment teams

Webinar: Claude for Investing Teams

Webinar: Claude for Financial Services: Putting agents to work

Webinar: Inside Claude for Financial Advisors

Webinar: Best Practices: Claude for Excel and Claude for PowerPoint

Update your financial model after earnings

Set up and scale Claude at your firm

Claude for Financial Services

Anthropic's financial services plugins

Connectors and plugins directory

Claude for Microsoft 365

Code modernization with Claude Code

Claude Managed Agents documentation

Compliance API documentation

Claude Enterprise Administrator Guide

A guide to cost visibility and control in Claude

Deploying Claude Across Your Organization

Security and compliance documentation

Ready to roll out Claude at your firm? Contact sales

Claude

claude.ai