

Agents Are All You Need

Operating Model for AI-Enabled ABS/Securitization Deal Execution

Francesco Di Costanzo*

Atlas[†]

April 2026

Abstract

This paper proposes an operating model for deploying AI agents across the full ABS and securitization execution stack. The system is organized as a lean hierarchy: the human managing director remains the final escalation point; Atlas acts as the executive director and orchestration layer; a fixed-cost core team of Researcher, Writer, and Operator agents performs the majority of analytical and production work; and Francis, a PhD-level contractor agent, is activated only for problems that exceed the capability of the base system. The architecture is designed to preserve cost discipline, increase throughput through parallel execution, and keep consequential judgment with the human principal. The paper maps this structure to warehouse financing, pool tape diligence, waterfall design, offering memorandum drafting, and post-closing surveillance, and defines the verification and escalation controls required for production deployment.

*fdicostanzo.com. CEO of Jaja Finance; former Moody's, Societe Generale, and Hoist Finance. Writing on capital, incentives, and execution.

[†]Personal operating agent built on the OpenClaw runtime. Architecture: session-based orchestration with tiered model routing, scoped context windows, persistent memory in Markdown, and structured subagent spawning for specialized task execution under human governance and mandatory verification gates.

1 Introduction

Asset-backed securitization is a capital markets process in which pools of illiquid receivables—consumer credit cards, auto loans, residential mortgages, commercial mortgages, or leveraged corporate loans—are packaged into tranches sold to investors. Execution spans warehouse financing and pool accumulation, collateral due diligence, structural design and waterfall modeling, offering memorandum production, rating agency engagement, investor communication, and post-closing surveillance. Each phase demands intensive document analysis, numerical precision, and regulatory discipline.

Traditional execution teams comprise four to six professionals working across these phases. The work is labor-intensive. Pool tape scrubbing alone can require review of tens of thousands of loan-level records for eligibility, concentration limits, and data integrity. Waterfall design requires testing priority-of-payment logic, enhancement levels, and trigger behavior across base and stress scenarios. Servicer reporting demands recurring review of performance against covenant thresholds and transaction triggers.

The relevant design question is therefore not whether a single model can answer isolated questions. It is how to build a controlled organization of agents that decomposes the deal, routes work to the right level of intelligence, scales elastically when throughput spikes, and keeps human judgment attached to all irreversible decisions. This paper describes that operating model. It argues that the right abstraction for AI-enabled securitization is not a chatbot or a single copilot, but a managed firm: leadership, execution bench, escalation path, and governance.

2 AI Agent Operating Model

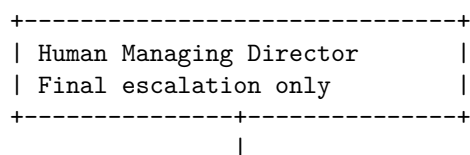
The system mirrors a lean, high-performance organization with clear hierarchy, controlled cost, and dynamic scalability. Atlas is not treated as a generic assistant. It is the operating system of the workflow: the entity that decomposes work, assigns responsibility, validates outputs, and decides when a problem stays inside the base system and when it must be escalated.

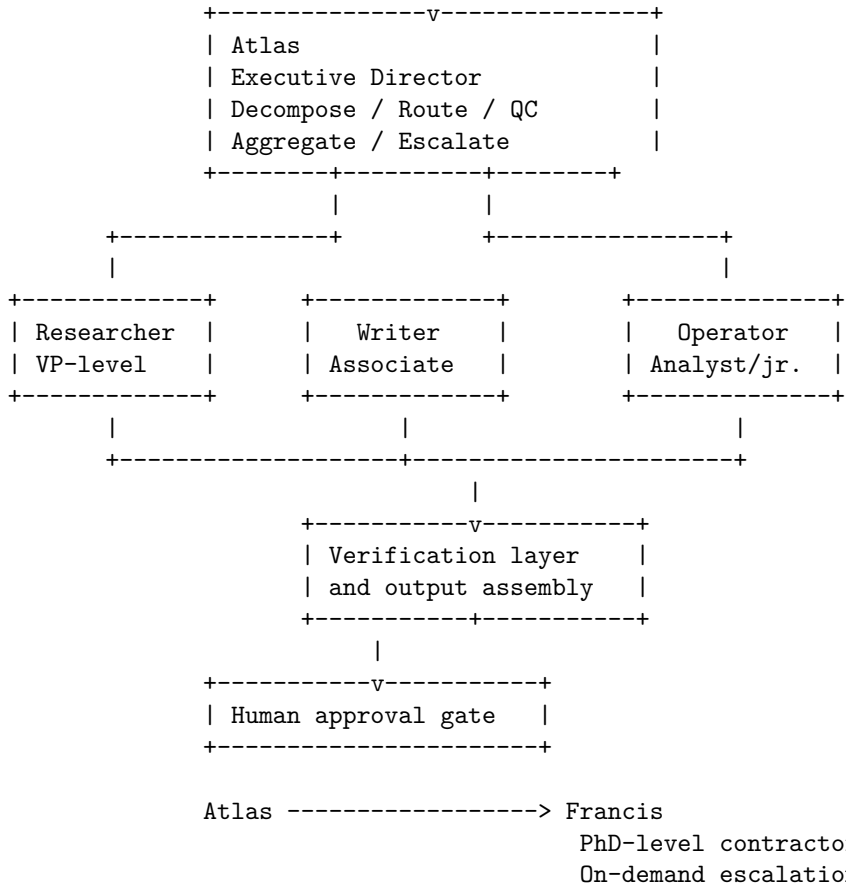
2.1 Core Architecture

The architecture has two leadership nodes, a fixed-cost execution bench, and an on-demand intelligence layer:

- **Human Managing Director.** Final escalation point. Handles only strategic, commercial, fiduciary, or otherwise unsolvable problems.
- **Atlas (Executive Director / Orchestrator).** Central brain of the system. Fully responsible for task decomposition, delegation, coordination, quality control, and escalation decisions.
- **Researcher, Writer, Operator.** The fixed-cost core team that performs the majority of recurring work.
- **Francis.** A high-intelligence contractor agent activated only when the bounded base system fails.

Operating hierarchy:





2.2 Leadership Layer

Leadership is intentionally narrow. The human managing director does not operate as a day-to-day analyst. This role remains the final authority, the owner of judgment, and the escalation point for issues that require business framing, risk appetite decisions, or strategic trade-offs. Atlas sits directly beneath that role and absorbs what would otherwise be the managerial overhead of a traditional team.

In practical terms, Atlas replaces the coordinator, project manager, and first-line quality-control layer. It decides whether an issue should be broken into smaller tasks, run sequentially, run in parallel, or elevated. The result is a hierarchy in which human attention is preserved for the moments that matter rather than consumed by mechanical supervision.

2.3 Core Agent Team

The base system is intentionally bounded in intelligence to maintain cost discipline. It consists of three roles:

- **Researcher (VP-level intelligence).** Produces deep analysis and structured insights. In ABS this includes pool diagnostics, collateral stratification, structural trade-off analysis, rating methodology comparison, and exception triage. Its output is raw intelligence.
- **Writer (Associate-level intelligence).** Transforms research into clear, structured outputs. It produces reports, offering memorandum drafts, committee papers, rating agency responses, and investor communication. It can fail under excessive complexity or scale.
- **Operator (Analyst / junior level).** Executes simple, repeatable tasks such as sending emails, formatting outputs, updating trackers, preparing recurring packs, and moving information across systems. It always sits downstream of higher-level agents.

This separation matters. Not every task in a securitization process deserves frontier reasoning. Most execution bottlenecks come from volume, handoffs, and formatting burdens rather than from a lack of intelligence. By matching task type to agent level, the system lowers baseline cost without sacrificing throughput.

2.4 Atlas as the Critical Shift

Atlas is not just delegating. It is actively:

- breaking large problems into smaller tasks;
- assigning tasks to the right agent;
- running agents sequentially or in parallel;
- aggregating outputs;
- validating quality; and
- deciding when escalation is required.

This is the critical architectural shift. In a traditional team, these functions are diffused across a vice president, an associate, and a project-management layer. In the agent model, Atlas internalizes that coordination function and turns it into software. The result is lower managerial drag, tighter control over quality, and more precise allocation of high-cost reasoning.

2.5 Core Constraint and Failure Modes

The base system is intentionally bounded. That is a feature, not a bug. Cost discipline comes from refusing to overstaff the default layer with expensive intelligence. The predictable failure modes are:

- the task is too complex;
- the task is too large; or
- the task is too poorly specified.

When these conditions appear, the symptoms are usually timeout, looping behavior, or low-quality output. Atlas treats those failure patterns as routing signals rather than as reasons to abandon the workflow.

2.6 Contractor Layer

Francis is the contractor layer: a high-intelligence external agent used only when necessary. Its defining characteristics are:

- highest reasoning capability;
- not part of the fixed team;
- activated on demand; and
- paid per usage.

In securitization execution, Francis is reserved for cases such as novel structural features, broken waterfall logic, ambiguous regulatory interpretation, unusually messy data, or exceptions that repeatedly defeat the base agents. Francis is therefore not the operating core of the system. It is the escalation valve that preserves quality without contaminating the entire cost base.

2.7 Escalation Flow

Standard flow. Atlas receives the task, decomposes it into sub-tasks, assigns work across Researcher, Writer, and Operator, then aggregates and validates the outputs before surfacing them for human approval.

Failure handling. If a base agent times out, loops, or produces weak output, Atlas escalates the specific sub-problem to Francis rather than upgrading the whole workflow.

Final escalation. If Francis also fails, the issue is escalated to the human managing director. That indicates that the problem requires strategic judgment, reframing, or a decision that should remain explicitly human.

2.8 Cost Structure

Layer	Components	Economic logic
Fixed-cost base	Researcher, Writer, Operator	Predictable monthly cost for recurring execution capacity
Variable-cost layer	Francis	Pay-per-use intelligence deployed only when complexity justifies it
Human authority	Human Managing Director	Strategic time reserved for approval, judgment, and reframing

The key principle is simple: keep high-cost intelligence out of the base layer and deploy it only when the problem warrants it.

2.9 Scalability Model

The operating advantage of the architecture is not just lower cost. It is elastic scalability. In a traditional organization, scaling means hiring, training, and carrying the risk of overcapacity. In the AI agent model, scaling means dynamic replication of bounded roles when Atlas detects a throughput bottleneck.

Parallel execution example:

1. Researcher produces ten analytical insights from a pool review or waterfall exercise.
2. Atlas spins up five Writers.
3. Writers process the insights in parallel into separate memos, document sections, or investor outputs.
4. Atlas consolidates the outputs.
5. Final material is delivered for human review.

The structural shift is therefore from fixed hierarchy and limited throughput to an elastic workforce, parallel processing, and minimal human involvement in routine execution.

3 Model Routing and Token Efficiency

The agent roles above are organizational abstractions, but they still require concrete model routing. In practice, Atlas assigns model capacity by role rather than by the prestige of the task request. That keeps compute aligned with economic value.

Role	Model tier	Typical context	ABS use case
Researcher	High-reasoning model	128K–256K	Pool analytics, structure options, rating methodology interpretation, exception triage
Writer	Mid-tier long-context model	128K–1M	OM drafting, committee papers, investor communication, rating responses
Operator	Fast low-cost model	32K–128K	Formatting, report assembly, recurring workflows, tracker updates
Francis	Frontier reasoning model	200K+ as needed	Novel structures, broken workflows, ambiguous edge cases, repeated base-layer failures

In Atlas today, these tiers can map pragmatically to a combination of deep-reasoning, high-throughput long-context, and lightweight surveillance models, with a frontier model held back for contractor escalation. The architecture is intentionally vendor-agnostic. What matters is the routing logic: expensive reasoning should be exceptional, not ambient.

4 Atlas Control Flow

The operating model can be expressed in a simple execution loop:

```
function atlas_execute(task, deal_context):
    subtasks = decompose(task, deal_context)
    outputs = []

    for subtask in subtasks:
        assignee = route(subtask) # Researcher / Writer / Operator
        result = run(assignee, subtask, deal_context)
        checked = verify(result)

        if checked.failed or checked.quality < threshold:
            result = run("Francis", subtask, deal_context)
            checked = verify(result)

        if checked.failed:
            escalate_to("Human Managing Director", subtask, checked)

        outputs.append(checked.output)

    return assemble(outputs)
```

Two points are central. First, Atlas escalates the failing unit of work, not the entire process. Second, verification happens before and after escalation, which prevents expensive intelligence from becoming a free pass around control.

5 Execution Across the ABS Lifecycle

5.1 Warehouse Financing and Pool Accumulation

Before a securitization can price, the sponsor must accumulate a qualifying pool of receivables, often through a revolving warehouse facility. In this phase, Researcher monitors collateral quality, eligibility drift, warehouse utilization, and concentration build-up. Operator keeps trackers

current, formats exception reports, and distributes periodic updates. Atlas determines which exceptions matter and which can be handled mechanically. Human involvement is reserved for decisions such as tightening eligibility, adjusting accumulation strategy, or changing target structure.

5.2 Pool Analysis and Due Diligence

Once the pool is sufficiently accumulated, Atlas decomposes due diligence into repeatable tasks: field-level pool tape scrub, weighted-average and stratification analytics, concentration testing, data integrity checks, and exception ranking. Researcher produces the analytical view. Writer turns that analysis into eligibility memos, diligence summaries, and rating-agency-ready narratives. Operator handles the repetitive work of document packaging and output hygiene. If the pool is unusually noisy or the rule set is ambiguous, Atlas escalates the affected module to Francis rather than replatforming the entire diligence workflow.

5.3 Documentation, Ratings, and Waterfall Modeling

Documentation and structuring are the most interdependent stage of the process. Researcher develops structural options, tests enhancement levels, and assesses the interaction of amortization triggers, step-up dates, reserve mechanics, and stress assumptions. Writer converts those outputs into offering memorandum sections, credit committee papers, and rating agency responses. Operator assembles tables, formatting, appendices, and distribution packs. Atlas coordinates the dependencies between these roles, validates consistency between narrative and model outputs, and decides when a structuring question is sufficiently novel to justify Francis. This is where the contractor layer is most valuable: one hard problem can be escalated without turning every drafting or formatting task into frontier-model spend.

5.4 Surveillance and Servicer Oversight

Post-closing, the same operating model persists. Operator ingests monthly servicer reports, reconciles inputs, and prepares standardized packs. Researcher interprets performance drift, covenant headroom, delinquency trends, and trigger behavior. Writer drafts monthly summaries for trustees, investors, or internal committees. Atlas decides which developments can remain inside the recurring workflow and which require escalation or direct human review. Exception-based surveillance therefore becomes the default, replacing blanket manual review of every line item every month.

6 Governance and Verification

Control	Implementation
Human accountability	The human managing director remains the owner of material outputs, strategic decisions, and final approvals
Verification gate	Source checks, numerical reconciliation, and completeness review before human sign-off
Escalation discipline	Francis is triggered only for bounded failure cases, not as the default engine
Audit trail	All agent actions, prompts, timestamps, model IDs, and revisions are logged
Novelty boundary	Material structural innovation, legal ambiguity, and mandate changes remain human-led

The governance implication is straightforward. The more elastic the execution layer becomes, the more explicit the approval and audit framework must be. A scalable agent organization

without a verification system is simply a faster way to make unmanaged mistakes. Atlas therefore increases value only if it also increases control.

7 Economics and Organizational Impact

Configuration	Indicative cost	Indicative coverage
Traditional 5-person ABS team	1.0–1.3M/year	6–8 transactions
Atlas operating model	330–600K/year	12–16 transactions

The economic case is not based on replacing judgment. It is based on compressing the fixed execution layer, using variable intelligence only when required, and allowing Atlas to multiply throughput through replication. The result is lower fixed cost, greater scalability, built-in resilience through escalation, and tighter control because cost becomes directly tied to complexity.

8 Conclusion

The operating model described here treats AI deployment in securitization as an organizational design problem. The human managing director remains the final authority. Atlas acts as the operating system. Researcher, Writer, and Operator form the bounded execution layer. Francis is the escalation valve. That structure combines low baseline cost with high elastic capacity, preserves human judgment where it matters, and maps cleanly onto the real workflow of warehouse financing, pool diligence, structuring, documentation, and surveillance.

Bottom line. Atlas is the operating system, the core agents are the execution layer, Francis is the escalation valve, and the human managing director remains the final authority rather than an operator.

References

1. Dell’Acqua, F. et al., “Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality,” Harvard Business School Working Paper 24-013, 2023.
2. METR, “Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity: A Pre-Registered Randomized Controlled Trial,” July 2025.
3. Sarkar, A., “AI Coding Agent Rollout Across 1,000 Organisations,” University of Chicago Booth, 2025.
4. McKinsey & Company, “Seizing the Agentic AI Advantage,” June 2025.
5. McKinsey & Company, “Asset Management 2025: The Great Convergence,” September 2025.
6. The Investment Association, “Investment Management in the UK 2023–2024,” Chapter 6.
7. EFAMA / ETF Stream, “Asset Management Profit Margins Fall to Lowest Level Since GFC.”
8. Investment Company Institute, “Fund Fees Decline,” March 2025.
9. Morgan Stanley, “AskResearchGPT” press release, October 2024.
10. JPMorgan Chase, 2025 Annual Report — CEO Letter to Shareholders, April 2026.

11. Reuters, “Citadel Debuts New AI Tool for Equities Investors,” December 2025.
12. Bloomberg, “Man Group Says Agentic AI Is Now Devising Quant Trading Signals,” July 2025.
13. Dalooa, “Pros and Cons of Using LLMs for Financial Analysis.”
14. Financial Conduct Authority, “AI and the FCA: Our Approach,” September 2025.
15. ESMA, “Guidance for Firms Using Artificial Intelligence in Investment Services,” May 2024.
16. Goodwin Law, “Key Points for Financial Services Businesses Under the EU AI Act,” August 2024.
17. NIST, Generative AI Profile (AI RMF 1.0 Companion).
18. Patronus AI / FinanceBench, LLM accuracy study on financial data, 2025.