



# The Judgment Layer.

---

AI is removing the people who used to build judgment, but not the people who are liable for it.

# Capturing human sign-off in AI-assisted workflows and compounding it into institutional judgment

The Judgment Layer is the approval system for AI-generated work. It replaces the existing sign-off interface with a structured decision system that captures rationale, enforces accountability, and feeds decisions back into the workflow. Its wedge is in high-liability workflows like audit, law, and finance, where humans need to stand behind decisions.

## WHAT THIS LOOKS LIKE IN PRACTICE - AUDIT

■ Judgment moments the Layer captures



# AI is eating professional services work in two ways

## AT THE MARKET LEVEL: REPLACING THE FIRM

**AI-native services (“AINS”) firms absorb the work** and the liability with it, replacing the incumbent outright. The wave is well-capitalized, accelerating, and now consensus among top-tier VCs.



AI-native TPA. A dozen engineers process claims more accurately than 1,000-person incumbents.



Rebuilds fund admin stack so AI can operate natively. Replaces armies of humans.



NDA's as outcomes, not legal hours. Replaces external counsel for standard work.

*This works where the liability **can** move with the work. In the highest-stakes professional work, it **can't**.*

## INSIDE THE FIRM: REPLACING THE MAKER

**AI takes the junior's work, but the senior must still sign.** The maker-checker model (the maker does the work; the checker reviews and signs) that built judgment is collapsing.

### BEFORE AI

**JUNIOR** does the work

↓ *repetition builds judgment*

**MID-LEVEL** reviews + signs

↓ *earned competence*

**SENIOR** signs with conviction

### AFTER AI

**AI** does the work

*MID-LEVEL: no reps*

↓ *must still sign*

**SENIOR** signs without reps

AI removes the maker. The checker still signs — *but the checker's competence came from having been a maker first.*

# But, there are firms AI cannot replace; humans still need to sign, and the risk is personal

## Audit

**AI NOW DOES** document extraction, work-paper drafting, risk scoring, anomaly detection.

**600 → 30 hr**

*Grant Thornton audit testing procedures*

### HUMAN STILL SIGNS

The partner signs the opinion. PCAOB inspections cite insufficient documentation.

### WHY IT'S PERSONAL

The partner's name is on the opinion. A PCAOB finding is public, career-defining, and uninsurable.

### PURCHASING TRIGGER

Inspection findings, quality review deficiencies, professional liability

## Law

**AI NOW DOES** drafting, redlining, research, due diligence, contract review.

**16 hr → 4 min**

*Complaint drafting at leading law firms*

### HUMAN STILL SIGNS

Personal duty of competence — the signature cannot be delegated to a model.

### WHY IT'S PERSONAL

Duty of competence is personal and non-delegable. The bar cites the attorney, not the firm.

### PURCHASING TRIGGER

Malpractice insurance requirements, bar association guidance, client audit rights

## Finance

**AI NOW DOES** model development, credit decisioning, transaction monitoring.

**400+ AI models**

*JPM uses 400+ AI models, each requiring named human owners*

### HUMAN STILL SIGNS

Regulations require a named human to own every model. The signature is personal liability.

### WHY IT'S PERSONAL

A supervisory finding lands on an individual's record, not the institution's.

### PURCHASING TRIGGER

SR 11-7 requirements, OCC exam findings, internal audit deficiencies

# This personal liability combined with regulatory tailwinds and upstream instrumentation mean the moment is now

**01**

## **Regulators are citing the gap**

PCAOB inspection findings now flag insufficient documentation of AI supervision in audit engagements. The Fed and OCC require banks to document AI model oversight under SR 11-7. These are current findings, not future risks, driving pressure for infrastructure like this.

**02**

## **Upstream instrumentation is already commercializing**

Businesses like Flow State have built the workforce operating system for enterprise AI spend, with paying customers like Confused.com, RAC, and RVU. Enterprises are already buying the idea that AI usage must be instrumented and made legible. So, AI spend has been instrumented at the usage layer, the next instrumentation point is the sign-off moment.

**03**

## **Liability insurers are repricing AI risk**

Legal malpractice carriers now ask about AI review processes at renewal. D&O policies are scrutinizing AI governance at the board level. Professional liability insurance for audit firms has begun referencing AI-assisted work specifically. Increasing insurance bills are likely a leading indicator.

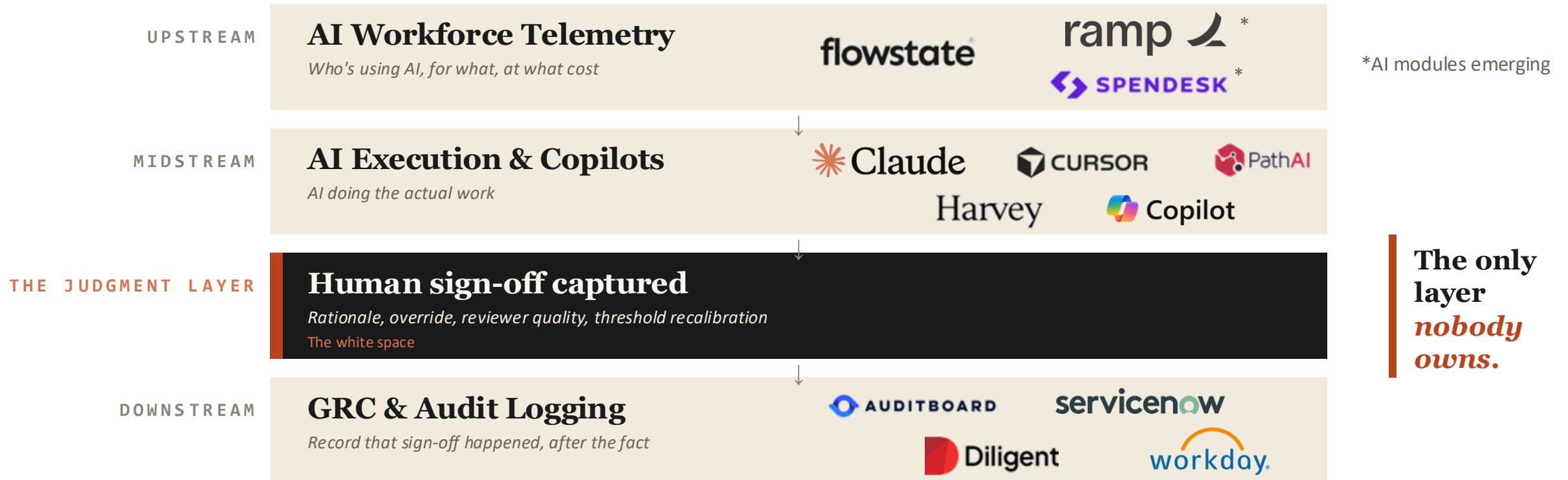
# A meaningful opportunity exists already across the high-liability sectors

Illustrative sizing based on potential client spend – directional, not a full TAM

|  |   |   |
|--|---|---|
| <p><b>Audit</b><br/><b>25 firms</b></p> <p><i>Big 4 + Next 20 PCAOB-registered</i></p> <hr/> <p><b>BIG 4 ACV</b><br/>\$10–12M per firm</p> <p><i>mission-critical infra across '00os engagements</i></p> <p><b>NEXT 20 ACV</b><br/>\$2–3M per firm</p> <p><i>mid-tier firms with AI deployment underway</i></p> <hr/> <p><b>BLENDED</b><br/>~\$4M average ACV</p> <p><i>conservative vs. Workiva and audit workflow comps</i></p> <hr/> <p><b>OPPORTUNITY</b><br/><b>~\$100m</b></p> | <p><b>Law</b><br/><b>200 firms</b></p> <p><i>AmLaw 200</i></p> <hr/> <p><b>ACV RANGE</b><br/>\$500k–2M per firm</p> <p><i>scales with firm size and AI deployment depth</i></p> <p><b>AVERAGE ACV</b><br/>~\$1M per firm</p> <p><i>modest vs. existing legal workflow software spend</i></p> <hr/> <p><b>OPPORTUNITY</b><br/><b>~\$200m</b></p> | <p><b>Finance</b><br/><b>~550 banks</b></p> <p><i>50 large (&gt;\$30B) + 500 mid (\$1–30B)</i></p> <hr/> <p><b>LARGE BANK ACV</b><br/>\$750k–1M per institution</p> <p><i>within existing MRM and GRC software budgets</i></p> <p><b>MID-TIER ACV</b><br/>\$150–200k per institution</p> <p><i>SR 26-2 extends obligation down-market</i></p> <hr/> <p><b>BLENDED</b><br/>~\$360k average ACV</p> <p><i>conservative vs. model risk management spend</i></p> <hr/> <p><b>OPPORTUNITY</b><br/><b>~\$200m</b></p> |
|--|---|---|

**Spend is likely to be smaller initially, but expand materially as AI penetration increases within each workflow**

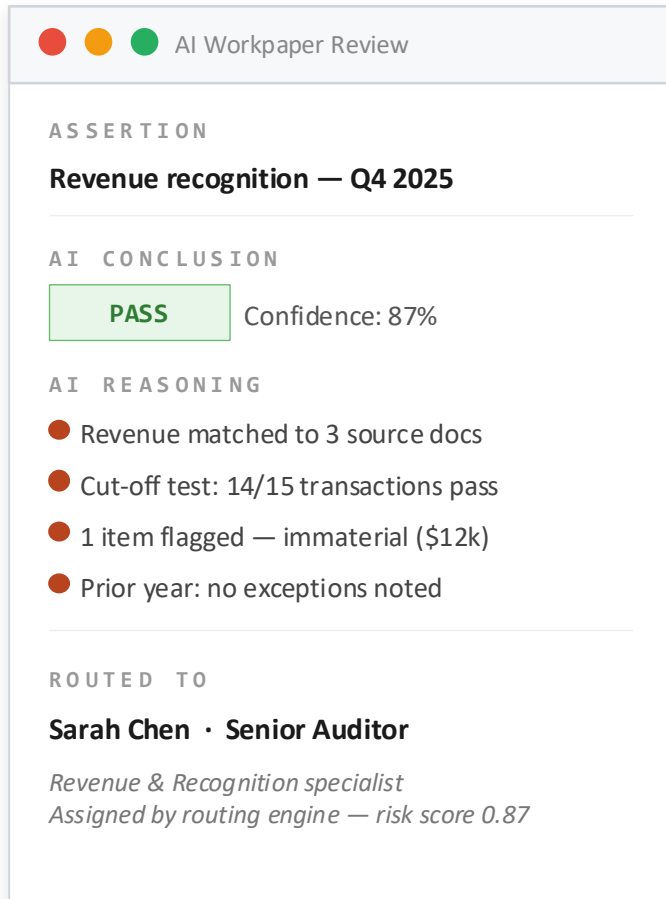
# There is a layer within the AI workflow that nobody owns; this is the judgment Layer



Telemetry watches AI usage. Copilots do the work. GRC records what happened.

Nothing sits in the sign-off moment itself — capturing the human judgment as it is applied.

# The Judgment Layer product is a decision layer that replaces the approval UI and controls the routing logic behind it



AI Workpaper Review

ASSERTION

Revenue recognition — Q4 2025

AI CONCLUSION

**PASS** Confidence: 87%

AI REASONING

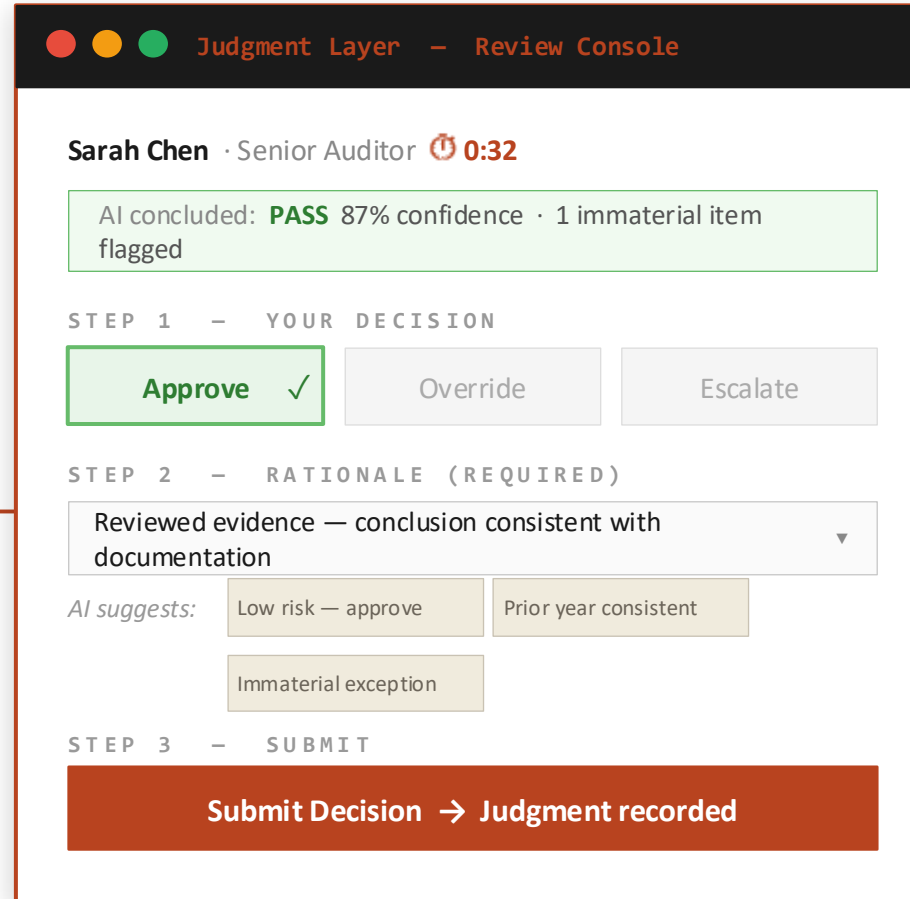
- Revenue matched to 3 source docs
- Cut-off test: 14/15 transactions pass
- 1 item flagged — immaterial (\$12k)
- Prior year: no exceptions noted

ROUTED TO

**Sarah Chen · Senior Auditor**

Revenue & Recognition specialist  
Assigned by routing engine — risk score 0.87

① AI drafts conclusion and routes to reviewer



Judgment Layer — Review Console

Sarah Chen · Senior Auditor ⌚ 0:32

AI concluded: **PASS** 87% confidence · 1 immaterial item flagged

STEP 1 — YOUR DECISION

**Approve** ✓ Override Escalate

STEP 2 — RATIONALE (REQUIRED)

Reviewed evidence — conclusion consistent with documentation

AI suggests: Low risk — approve Prior year consistent Immaterial exception

STEP 3 — SUBMIT

**Submit Decision → Judgment recorded**

② Reviewer decides in 3 clicks  
Rationale captured in < 10 seconds



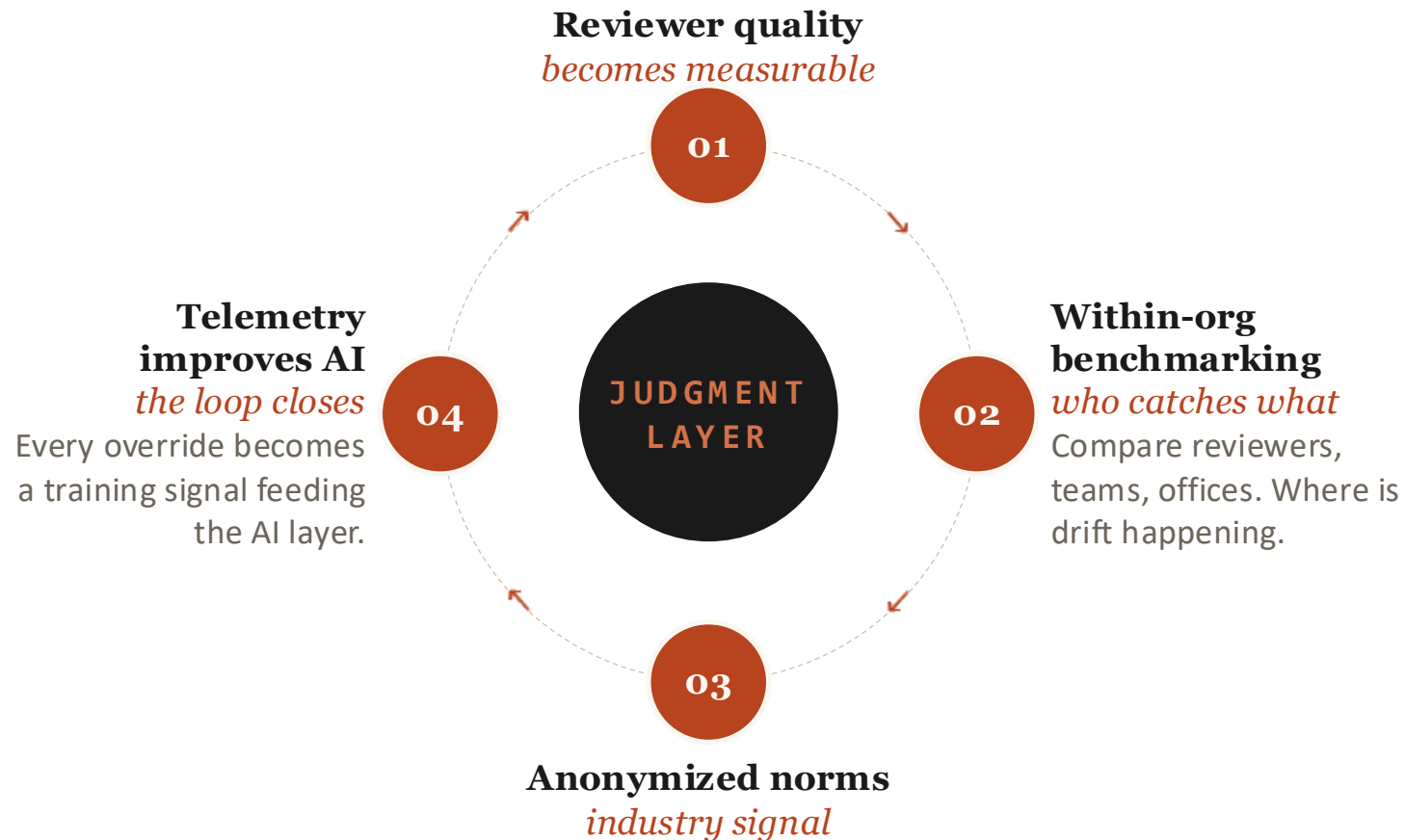
```
judgment_record.json

"reviewer_id": "sc_4421"
"decision": "APPROVE"
"rationale_code": "R07"
"rationale_text": "Reviewed evidence — consistent"
"ai_conclusion_seen": "PASS"
"confidence_at_review": 0.87
"time_on_task_sec": 8
"escalation_path": null
"timestamp": "2026-04-21T18:44Z"
"exported_to": ["AuditBoard", "ServiceNow"]
"workflow_unblocked": true
```

③ Judgment record written  
Workflow unblocked · GRC notified

# Every sign-off in the Judgment Layer makes the next one smarter, which compounds into a flywheel















The system of record for human judgment



## WHY IT'S VENTURE SCALE

- Data moat: structured decision dataset
- Workflow moat: embedded in approval path
- Switching cost: tied to compliance and audit trails

# This is a category primed for a startup to win, not incumbents

| CATEGORY  | VERDICT   | WON'T BUILD   | WOULD ACQUIRE   |
|---|---|---|---|
| <b>GRC Platforms</b><br><i>AuditBoard, ServiceNow, Diligent</i>   | They run the approval workflow but capture timestamps, not rationale. Adding structured judgment capture would be a product philosophy shift. Active M&A programs — most likely acquirer.                   |    |    |
| <b>Big 4 Internal Tools</b><br><i>Omnia, Halo, Clara, Helix</i>   | They have the data and the users, but services firm DNA means they build in-house rather than acquire. Internal deployment is large; commercial acquirer is unlikely.                                       |    |    |
| <b>AI Model Monitoring</b><br><i>Arize, WhyLabs, Fiddler</i>      | They watch the AI, not the human. Different buyer (ML Ops vs CRO/GC), different data model. Pivoting into judgment capture would mean rebuilding around the accountability surface.                         |    |    |
| <b>Execution Layer AI</b><br><i>Harvey, Copilots, Vertical AI</i> | Already embedded in the moment where work is executed and reviewed. Positioned as copilots or agents today, but can extend into approval and capture. Well-capitalized players may choose to acquire.       |    |    |
| <b>ERP / System of Record</b><br><i>SAP, Oracle, Workday</i>      | They log that approvals happened, not why or how. Could add reason codes, but business priorities sit elsewhere. Acquire for scale, rarely for capability.  |    |    |
| <b>AI Workforce Telemetry</b><br><i>Flow State</i>                | Sits one layer above the Judgment Layer. Different buyer (CTO/CFO for spend vs CRO/GC for liability), different data model. Complementary, not competitive.   |    |    |
| <b>Foundation Models</b><br><i>Anthropic, OpenAI</i>              | Sell horizontal model capability, not vertical workflow infrastructure with regulatory integration, named customer success, and audit-ready logging. Will power a Judgment Layer product, but won't be one. |  |  |

## THREE STRUCTURAL REASONS NO ROW ABOVE SHIPS THIS

### 01 Data cuts across systems

Sign-off spans AI output, human decision, rationale, and outcome. No single vendor controls the full surface.

### 02 No write-path to routing

Incumbents treat approval as an event, not a dataset and control surface. To compete, they would need to redesign their systems around decision quality, not workflow tracking.

### 03 Different buyer, different sale

Incumbents sell to compliance, IT, ML Ops. The Layer sells to CRO, GC, Chief Auditor.

# How I'd underwrite a company in this space

| CRITERIA                              | WHAT GOOD LOOKS LIKE   | HOW I'D ASSESS  |
|---------------------------------------|--|---|
| <b>1 Buyer with downside</b>          | The person writing the check personally owns the risk — a CRO, General Counsel, or Chief Auditor. Not a compliance manager or IT director. | <i>Customer ref calls confirm the budget owner. Contracts list a named executive sponsor.</i> |
| <b>2 Embedded in routing</b>          | The product can actually block a sign-off until rationale is captured. That needs write access into the approval workflow.                 | <i>Live demo of the escalation path. Integration partner confirmation.</i>                    |
| <b>3 Zero-friction capture</b>        | Reviewers record their reasoning in seconds, e.g. reason codes and smart defaults, not just a blank text box.                              | <i>Product walkthrough. Reviewer time-on-task telemetry.</i>                                  |
| <b>4 Clear purchasing trigger</b>     | Customers can point to a specific event that made them buy: a regulatory finding, an audit deficiency, a malpractice exposure.             | <i>Customer reference calls. Reasons cited in churn data.</i>                                 |
| <b>5 Within-org compounding first</b> | The product gets smarter inside one customer's data before any cross-customer benchmarking; value can be proven.                           | <i>Roadmap review. Data architecture diligence.</i>   |
| <b>6 Design partner credibility</b>   | At least one brand-name regulated firm, is actively using the product. Lets me verify real adoption, not just a logo on a website.         | <i>Reference calls with the named partner. Pilot scope and KPIs.</i>                          |

## KILL CRITERIA

*Sidecar dashboard, no routing write-access / Pricing tied to seats, not risk reduction / Cross-org data play before within-org value proven / Free-text rationale capture / No named executive buyer*

# Companies are emerging in this space, but the winner hasn't been built yet

| CRITERIA                       | Flank* | truthsystems | WITNESS AI | ConductorAI |
|--------------------------------|--------|--------------|------------|-------------|
| 1 Buyer with downside          | ✓      | ✓            | ✓          | ✓           |
| 2 Embedded in routing          | ✓      | ✓            | ◐          | ✓           |
| 3 Zero-friction capture        | ◐      | ✗            | ✗          | ✗           |
| 4 Clear purchasing trigger     | ✓      | ✓            | ✓          | ✓           |
| 5 Within-org compounding first | ◐      | ✗            | ✗          | ◐           |
| 6 Design partner credibility   | ◐      | ◐            | ✓          | ✓           |

No company currently meets these criteria. The winner hasn't been built yet.

# Appendix

# The product in depth

**LIVES:** Inside the existing approval workflow

*Replaces or wraps the sign-off screen in audit, legal review, or model validation tools via API (no rip-and-replace required)*

**FORM:** Review console · Routing engine · API

*Three components, one integration point. Console for the reviewer, engine for routing logic, API for the write-back*

**FORCES:** Structured rationale + mandatory escalation

*Reason code dropdown with AI-suggested options. Escalation paths enforced before approval is permitted*

**WRITES:** Back into workflow and blocks approval if incomplete

*Write-path into the approval system. Sign-off is gated until rationale is captured.*

**EXPORTS:** To GRC platforms and audit logs

*Structured JSON to AuditBoard, ServiceNow, Workday, or any downstream compliance system via webhook*

## WHAT GETS CAPTURED

Reviewer identity · AI conclusion seen · Decision made · Rationale (reason code + note) · Time taken · Escalation path

## WHAT GETS MEASURED OVER TIME

Reviewer quality drift · Override rate by team · AI accuracy vs. human judgment · Routing threshold performance

## HOW IT INTEGRATES

API into existing approval workflow · Write-path to routing engine · Read from AI copilot output · Export to GRC / audit log

# Key risks I would watch out for

## 01

### Execution-layer AI tools absorb the approval layer

If agentic platforms embed approval, escalation, and audit directly into their workflows, the decision moment gets captured in-place. The standalone judgment layer becomes redundant as a feature rather than a system.

#### WATCH FOR

*Execution tools (e.g. Flank, Harvey, Copilot) shipping native approval, escalation, and audit capabilities that remove the need for a separate layer.*

## 02

### AINS eats further up than expected

If AINS portfolio companies take PCAOB-regulated audit or AmLaw 50 transactional, the addressable carve-out shrinks. The structural ceiling I'm betting on becomes leaky.

#### WATCH FOR

*Big 4 acquisition of an AINS audit play, or AmLaw 50 partnership with an autopilot.*

## 03

### Foundation models close the judgment gap

If next-gen models produce defensible conclusions with embedded reasoning traces that satisfy regulators, the human sign-off becomes vestigial and the layer collapses.

#### WATCH FOR

*Anthropic / OpenAI shipping vertical-specific reasoning products with regulatory traction.*

## 04

### GRC incumbents acquire and integrate fast

AuditBoard or ServiceNow could buy the leading judgment-layer startup early and integrate before the data moat compounds. Acquisition becomes the only viable exit, capping returns.

#### WATCH FOR

*Active M&A signaling in the GRC space, particularly AuditBoard.*