fineSample

Wald

Autonomous approvals and exception handling for critical workloads.

Autonomy that is earned, and keeps being earned, on your own data, in your environment.

A brief from fineSample, November 2026

The test history quoted in this brief is synthetic and was written by fineSample.

Most approvals and exceptions never needed a person

Every team that makes decisions at volume has a queue: alerts a detector raised, payments on hold, account unlock requests, refunds and claims flagged as exceptions, approvals waiting for a sign-off. Most of these items turn out fine, and a person still opens each one to find that out.

Clearing them automatically is easy to switch on and hard to defend. A confidence score says how sure a tool is, not whether it was right. And most tools are scored by how often they agree with your people, which shows whether they copy your people, not what happened next.

	Automation today	Wald
It learns from	The items people looked at	Those, and a blind sample of its own clears
An unfamiliar item	Gets a score anyway	Goes to a person, and the verdict becomes evidence
When results turn bad	A monitor raises an alarm	Clearing stops on its own, and the queue carries on
What you show an auditor	A narrative	A replay that gives the same answer on any computer

One answer per item: needs a person, or safe to clear

For each item, Wald checks your own records and answers one question: does this need a person? If it does, the item goes to your queue with the reasons, the most urgent first. If it doesn't, Wald clears it in your own systems: releases the payment, unlocks the account, approves the request or closes the case. Every answer shows the facts behind it and replays exactly later.

Wald isn't tied to one domain. Each kind of work gets a pack that says which facts matter and what clearing means there. Your own experts write packs, or have an AI assistant draft one from your data, and two people approve it before it runs.

Autonomy is a learning loop on your data, in your environment

Wald doesn’t arrive with authority, and it doesn’t get it from its settings or from time. It earns it from outcomes on your own work that it can’t see in advance, and everything below runs inside your environment.

1 **Calibrate.** Wald scores its answers against your own outcomes, so a probability means what it says on your work.

2 **Release.** For kinds of work that have earned the right, Wald clears items inside your limits, in seconds.

3 **Inspect.** Your people check a blind sample of what Wald would clear, without seeing its answer. Later results from your systems, such as a chargeback or a reopened case, count too.

4 **Learn.** Wald retrains on those checks and outcomes. A new version has to beat the current one on cases neither has seen, two people approve it, and it starts earning from scratch.

You set the miss rate you’ll accept, one in 200 by default. Until a kind of work earns the right, every item Wald would clear is checked, which typically takes 150 to 250 checks. After that, one in twenty still is.

The right is never permanent. A miss can take it away. A watch on recent results stops clearing early when they start to turn. A missing fact sends the item to a person. Anyone with the authority can stop all clearing at once, and restarting takes two people. When clearing stops, the queue carries on as it did before Wald.

Every cleared item passes three checks

Earned authority isn’t enough on its own. Wald clears an item only when all three hold:

1. **No limit objects.** Fixed checks run on every item: an amount over the ceiling, an open hold or case, a source system that asked for a person, too many clears for one account today. Your settings decide their values, but none can be switched off.
2. **The answer clears your bar,** such as “at least 95 percent likely to be safe to clear”, calibrated on your own outcomes.
3. **This kind of work has earned the right.**

Anything that fails sends the item to a person. No failure ends with an item cleared.

Built for model risk, security and audit review

- **A model your validation team already knows.** The default judge is a points-based scorecard, and every answer shows how many points each fact contributed.
- **Exact facts, as of the moment.** Facts are worked out from your own events as they stood at the time, and say so when data is missing. The same facts serve live answers, history and training, and you can use them on their own for your own models.
- **Same records, same answer.** Export the evidence for any period and verify it on a computer that isn't connected to Wald.
- **Two people for every change,** and the person who proposes a change can't approve it.
- **Personal data stays in your network.** A relay you run replaces identifiers with tokens before anything reaches Wald.
- **Fast and cheap.** A few milliseconds per decision on ordinary CPUs.

How it compares with LLMs, hosted AI and in-house models

A frontier model, a vendor’s hosted service, a local model or your own statistics can each supply a judge. None supplies what clearing needs around it: calibration on your outcomes, limits that always run, authority earned per kind of work, and a stop that triggers on its own. In Wald they are candidates for the judge rather than rivals. Whichever judge answers, it has to earn the right to clear the same way.

AI agents fit the same way. An agent can answer Wald’s questions, send its own verdict with an item, read Wald’s facts or draft a pack. It can do the judging. It can never grant itself the right to act.

Other approaches ask how accurate a model was when it was tested. Wald asks whether this kind of work has been safe to clear on your own work lately, and keeps asking.

Two numbers that matter

- **Person rate:** how much of the queue still needs a person. The workload you get back.
- **Miss rate:** of the items Wald would clear, how many needed a person after all. The risk you take on.

On a published test history of 600 made-up alerts, Wald would have cleared 526, and 6 of those needed a person after all (1.1 percent). That is above the default bar, so Wald would not yet have earned the right to clear. We wrote that history ourselves, so it shows how the report reads, not what you would get. Only your own work counts.

Work with us

We’re looking for teams with a high-volume approval or exception queue to work with us on Wald. It starts with a year of your own history, run through a free evaluator on your own computer, which shows how much of your queue never needed a person and what Wald would have missed. From there, Wald runs beside your live queue while your people do the blind checks, and clears nothing until the results say it may.

Talk to us about Wald: hello@finesample.com