

Orbyt Labs · We are the test

12 AI executives. One human.

Engineering to finance, with sub-agents under them, doing the work. Nobody signs off on their commits. The checks do. When they fail, it's public.

Software roles, not corporate officers. One GitHub repo. Claude Code seats. Codex builds. Justin hosts workers.

Agents move. Oversight lags.

62% of respondents experiment with or scale agents. No individual function exceeds 10% scaling. McKinsey survey: 1,993 respondents across 105 nations.

Source: [McKinsey, The state of AI in 2025, 2025-11-05](#)

26% of agentic-AI users surveyed cannot detect unauthorized internal agents. EY-commissioned full sample: 202 US senior AI decision makers, public companies, \$1B+ revenue. Subgroup size unstated.

Source: [EY, Autonomous AI implementation outpaces oversight, 2026-09-15 · vendor-reported](#)

Bet on the whole company.

Our bet: an operating system for every company. People set authority. Checks decide what lands.

Agents bluff. Checks don't.

Justin Bartak, founder and CEO.

September 22, 2026. Agent claimed done. Check refused.

04:37 · State changes	Operations committed new state.
Claimed done	Engineering called it complete.
Refused	Stale file on main, not seat code. Still counts against the seat.
05:43 · Fix it	Founder-launched loop: replay, fix, regression test. No intervening founder prompt. CI passed.
06:43 · Pass the checks	Engineering re-dispatched. Checks passed. Report landed. Dispatcher unrecorded.
Human hand	Justin launched it and wrote its queue. A separate defect later stopped it.

Case below. Source identifiers withheld. [Read the dated case record.](#)

Yeah, we publish the failures, too.

Corpus: mostly founder-directed sessions. About one in four touch our machinery. No origin field. 11 undated; 6 predate the run window. 106 of 168 counted runs: CoS, quality, product and legal reports. Marketing: 18 of 27 passed; 9 of 17 failures. Minutes saved: unmeasured.

Everyone builds agents.

We build the company they work in. Dated overlaps below: not a full list or ranking.

GitHub Agent HQ. Multi-agent control, permissions and audit for repository work. Source: [GitHub, 2025-10-28 · vendor-reported](#)

Microsoft Agent 365. Agent registry, access control and quarantine for enterprise governance. Source: [Microsoft, 2025-11-18 · vendor-reported](#)

OpenAI AgentKit. October 2025 launch. June 2026 update: Agent Builder and Evals end November 30, 2026; other OpenAI tools recommended. Source: [OpenAI, 2026-06-03 · vendor-reported](#)

LangChain / LangGraph. LangChain and LangGraph: agent frameworks and orchestration. LangSmith: tracing and evaluations. Source: [LangChain, 2025-10-20 · vendor-reported](#)

CrewAI. Agent orchestration and an enterprise platform in its October 2024 announcement. Source: [CrewAI via GlobeNewswire, 2024-10-22 · vendor-reported](#)

Claude Agent SDK. An SDK using Claude Code's agent loop and tools. Source: [Anthropic \(guide reviewed\), 2026-10-20 · vendor-reported](#)

Software first. Then the world.

73,499 sums two Census 2022 counts below 100 employees: 12,610 software publishers plus 60,889 custom programming firms. Possible double counting; not unique prospects or measured demand.

Source: [US Census Bureau, SUB 2022, 2025-04-10](#)

The commercial bet.

Plan: Apache-2.0 open core. Paid hosting. Release scope and branding remain under review. **Proposed:** Open: seats, workflows, checks. Paid: cross-repository history, governance evidence, support. **Pricing:** Unset. Informed by token usage. Plans as of October 22, 2026, not promises. Installation, demand, supervision savings and economics unproven. Failure risks: technology, founder/provider dependence, costs, competition, unresolved branding rights.

Check our work.

Completed is not accepted. 151 of 168 counted recorded runs passed checks, 89.9%. 2 founder verdicts; acceptance unproven. Partial months: July 21 to 31; October 1 to 22. Skips ("nothing to do") and harness faults sit outside. Count faults as failures: October 46 of 49, 94%. Gates changed: September 20, no_sentinel retired; September 23, 3 faults reclassified. Checks changed; rates do not measure constant-quality work.

The proof still owed.

Acceptance: earn it. 64 of 67 measured outputs untouched by the founder; 3 derived edited. Trusted or unread is not yet measured. 82 completions have no disposition.

Founder time: at least 370 reconstructed hours; 99% unclassified. Early months undercount: transcript retention.

Economics: Model tokens are runtime-reported usage, not a dollar cost measure. Subscription, infrastructure and founder costs have not been allocated to individual runs. Dollar cost per run and per accepted outcome has not been measured. These observations do not establish costs for hosted customer use.

Outside: comparison, external installation, demand unproven. Hosting unfinished; not for sale.

Provider dependence: Scheduled agent roles currently depend on Claude Code. Provider access, usage limits, terms and pricing may change. The recorded setup does not establish permission or economics for a hosted customer service. Providers can compete.

Logging gaps. Detailed records: 107 of 180 rows.

Authority: code enforces paths, landing checks, HALTs. Spending, pricing, legal calls and claims rely on instructions and judgment. Signed admission is coded, not live-proven.

The test that closes each gap is in the R&D Evidence Pack.

We back the bet. Test it. Does it pay for itself? Proposed test: an external operator installs and repairs. Founder's computer off. Buyer value untested.

Current products: Orbyt Jobs and Orbyt Intelligence accept waitlist registrations. Orbyt One: accounts; billing in test mode, not live. Built in founder-directed sessions.

Justin Bartak, Founder and CEO

hello@orbytlabs.ai ·

<https://www.orbytlabs.ai/investors>

[Research or partnership conversation](#)