# Persistent Agents, Security SaaS, and a Higher Bar for AI Outcomes

VC Tech Radar

2026-05-03

## Persistent Agents, Security SaaS, and a Higher Bar for AI Outcomes

*By VC Tech Radar • May 3, 2026*

Early traction this cycle clustered around AI security and workflow automation, while the strongest technical signals came from persistent-agent architectures and personal-agent stacks. The broader investor read-through: open-source model momentum remains credible, security diligence is becoming more urgent, and the bar for venture-scale SaaS outcomes keeps rising.

### Funding & Deals

- **Meta/Manus is the clearest deal signal in the set.** Harry Stebbings flagged that China blocked Meta's $2B Manus deal, and his accompanying interpretation is that an attempted unwind would pressure the acquirer and future cross-border deal behavior more than already-distributed VC proceeds [1]. For investors, it is a reminder that AI exits can pick up meaningful geopolitical risk [1].
- **Autonomous incident response is already a funded category.** A founder building a code-level production-crash agent points to Resolve AI's $125M raise as evidence that autonomous incident handling has serious capital behind it, while arguing the open gap is code-level reproduction and fixing rather than infra-only workflows [2].
- **YC remains a strong early filter for agentic commerce tooling.** LocusFounder says it joined YC this year and is opening 100 free beta spots for an AI system that builds a website, conversion copy, ads, and back-office operations around a user's project idea, with users keeping any revenue generated [3].

## Emerging Teams

- **CheckVibe: security for AI-built apps is getting paid, quickly.** The two-person bootstrapped team built a scanner for apps shipped rapidly with AI tools; they report $3.4k in gross revenue, 100+ paying customers, and 2.5k signups within six weeks, with a public Stripe dashboard linked in the post [4]. The team also says security-critical scanner logic was architected manually rather than vibe-coded [4].
- **Zeriflow: repo-level analysis looks like the sticky wedge in security SaaS.** Eight months after v1 and 12,400 scans later, the founder reports that about 70% of paying users connect GitHub repos, monitoring with score-drop alerts drives repeat usage, and v2 adds a PR-blocking GitHub Action, live README badge, and REST API [5]. The biggest open product issue is false positives, which the founder says are the top cause of churn [5].
- **Transita: narrow workflow, fast UX, and MCP-native distribution.** The visa-eligibility product says it shipped as an MCP server inside Claude Desktop, Cursor, and Cline, uses an anonymous quiz plus token-based paid unlocks, and returns a deterministic top-six country match before slower AI enrichment streams in [6]. After about six months, the founder reports 41 quiz completions, 22% email capture, and 5% paid conversion at $9 [6].
- **A white-box compliance engine is attacking a high-trust niche.** A solo founder says the product shows the exact logic path used to verify a document against a compliance rule, aiming to replace opaque probabilistic outputs with inspectable reasoning; the compliance workflow has just launched and is seeking beta feedback [7].
- **Code-level crash automation is a credible new agent wedge.** Another founder says their CLI converts a Sentry crash URL into a failing pytest on the current branch and verifies whether a fix worked, targeting the 30-40 minutes engineers often spend reconstructing state before debugging even starts [2]. The open design question is how much autonomy developers will trust in production, especially for billing or payments code [2].

## AI & Tech Breakthroughs

- **Persistent agents are moving beyond disposable task runners.** AIPass distinguishes disposable sub-agents from persistent "citizens" that keep identity, memory, tests, and domain-specific behavior inside a layered orchestrator -> citizen -> sub-agent architecture [8]. The project gives concrete examples: a mail citizen with 696 tests built through failures and a routing citizen shaped across 80+ sessions of bugs and fixes [8]. Its memory model uses `passport.json`, `local.json`, and `observations.json`, injected each session so the citizen does not start cold [9]. The project says the repo is CLI-based on Claude Code, Linux-focused, and currently

at 85 stars, 400+ PRs, and 6,500+ tests [8, 9].

- **Garry Tan is open-sourcing a personal-agent stack, not just a demo.** Garry Tan describes GBrain as his OpenClaw/Hermes-based personal agent setup with custom retrieval, graph DB, schema, and skillpacks, with 100+ skills planned [10]. The shipped "book-mirror" skillpack maps an author's ideas to the user's own life and projects, while newer skills cover article structuring with verbatim quotes, strategic reading, concept synthesis, web research against knowledge gaps, and archive mining gated by an explicit allow-list [11, 12, 13]. Tan also says the project is experimental and still in a "Homebrew Computer Club stage" [14].
- **A new RoPE paper offers a concrete explanation for compositional reasoning gains.** The paper argues Rotary Positional Embeddings let transformers solve compositional reasoning tasks where additive positional layers fail, proving a toroidal structure on finite groups and validating the claim with Qwen2.5-0.5B on modular arithmetic and sequential composition tasks [15].
- **The training-tooling layer is getting stricter.** Parallelogram is positioned as a linter for LLM fine-tuning datasets that catches broken data before a GPU run starts [16].

## Market Signals

- **AI-assisted software creation is outrunning basic security controls.** In audits of eight vibe-coded SaaS apps, five had row-level security turned off on at least one user-data table, three exposed Supabase service-role keys to the browser, two trusted `user_id` from form bodies without session checks, and none had rate limits, including on auth [17]. The auditor's broader claim is that codegen tools assume senior engineers review the diff, and increasingly there isn't one [17]. One founder was reportedly pitching ARR projections while a service-role key was exposed in bundled client JS [17].
- **The hurdle for venture-scale SaaS outcomes continues to rise.** Harry Stebbings amplified the view that $400M ARR growing 30% is no longer enough; companies now need a path to $1B+ growing 40% to produce strong outcomes, while everything below that risks a weak public-market result [18]. He separately highlighted a venture market increasingly driven by a small number of massive winners [19].
- **Open-source model momentum has elite support, but benchmark optics are messy.** Marc Andreessen endorsed a post arguing that Kimi k2.6 and DeepSeek v4 show open-source scaling is continuing, and that the market cap of companies built on top already exceeds OpenAI plus Anthropic combined [20, 21]. In a separate exchange, he amplified criticism of IRT ELO charts: as benchmarks approach saturation, moving from 97% to 99% accuracy can show up as a 200-point ELO gain, which can exaggerate apparent model gaps [22, 23].
- **Engineer demand is not collapsing.** A Citadel Securities analysis

cited on X says software-engineer job postings are up 18% from the May inflection point last year, and Andreessen endorsed that readout [24, 25]. Another Andreessen-amplified post argues "we need more engineers, not less," while Garry Tan-adjacent discussion around personal agents points to emerging roles like "personal agent designer," "second brain engineer," and "context editor" [20, 26, 27].

- **Capability still appears to matter more than cheaper inference.**

  "i keep thinking i want the models to be cheaper/faster more than i want them to be smarter but it seems that just being smarter is still the most important thing" [28]

Sam Altman framed intelligence as the higher priority, while Naval's shorter thesis is that AIs replace UIs and APIs [28, 29].

## Worth Your Time

- **Harry Stebbings on AI capex, agent wars, and Google's infrastructure position** — compact read on compute intensity, agent-driven vendor selection, and why Google may benefit regardless of Gemini vs. Anthropic outcomes [1].
- **Doug's IRT ELO thread, amplified by Marc Andreessen** — useful before taking model leaderboard charts at face value [23, 22].
- **GBrain repository** — the clearest concrete artifact in the set for personal-agent architecture and skill design [10].
- **AIPass repository** — useful if you are tracking persistent-agent systems instead of one-shot copilots [8].
- **Vibe-coding security checklist** — practical diligence aid for founders and investors reviewing fast-shipped AI products [17].

---

**Sources**


1. X post by @HarryStebbings
2. r/SaaS post by u/sszz01
3. r/SideProject post by u/IAmDreTheKid
4. r/SaaS post by u/funfunfunzig
5. r/SaaS post by u/famelebg29
6. r/SaaS post by u/KanekiAyato
7. r/SaaS post by u/Sweaty-Ad5953
8. r/artificial post by u/Input-X
9. r/artificial comment by u/Input-X
10. X post by @garrytan
11. X post by @garrytan
12. X post by @garrytan
13. X post by @garrytan
14. X post by @garrytan

15. r/deeplearning post by u/Dan23RR
16. r/deeplearning post by u/Quiet-Nerd-5786
17. r/SaaS post by u/damn_brotha
18. X post by @HarryStebbings
19. X post by @HarryStebbings
20. X post by @casper_hansen_
21. X post by @pmarca
22. X post by @pmarca
23. X post by @0xdoug
24. X post by @Konstantine
25. X post by @pmarca
26. X post by @nxt3d
27. X post by @garrytan
28. X post by @sama
29. X post by @naval