GitButler and Anvil Raise as Autonomous ML and Agent Infrastructure Accelerate

VC Tech Radar

2026-04-09

GitButler and Anvil Raise as Autonomous ML and Agent Infrastructure Accelerate

By VC Tech Radar • April 9, 2026

Capital moved into agentic developer tooling and physical-AI infrastructure, while Disarray emerged as a notable autonomous ML team. The broader pattern is rising enterprise adoption, growing pressure around managed-agent lock-in, and practical infrastructure advances in compression, cyber capability, and safe model packaging.

1) Funding & Deals

- **GitButler — $17M Series A led by a16z.** GitButler is building version control designed for agentic coding, including an agent-oriented CLI, parallel branches for multi-agent workflows, and a rethink of code review, PRs, and commit messages. Scott Chacon's return to version control is explicitly tied to Git's largely unchanged UI since 2005 and the need for new tooling as developer communication becomes more valuable. a16z investing note [1, 2, 3]
- **Anvil Robotics — $6.5M led by Matter Venture Partners.** Founder @0x796F said he has spent eight years shipping hardware. Leo Polovets describes Anvil as shared infrastructure for Physical AI teams—hardware integrations and teleop software—so teams can focus on their differentiated solution and intelligence stack. Additional participants include @humbavc, @vsodera, @spacecadet, and @Position_VC [4, 5]

2) Emerging Teams

- **Disarray posted unusually strong early results for an autonomous MLE agent.** It won 28 Kaggle medals across vision, NLP, and tabular tasks, placed top 10 in nine competitions, and beat all

human teams in one competition, all within 24 hours on a single GPU. The system starts from a high-level task, plans and refines ML workflows on its own, and augments data from public sources; the founding team is two PhDs with backgrounds across Databricks, Google, LinkedIn, Microsoft, NASA, and IBM, and its backers include the Kaggle founder, the former U.S. Chief Data Scientist, and the co-founder of Databricks and Perplexity [6]

- **Instapi.co is an early example of an agent-first product.** The founder's premise is that agents are blocked by human-only UX, CAPTCHAs, and 2FA, so Instapi lets agents sign up via `curl` and pull live Instagram data without opening a browser, with automatic image and video parsing plus metadata on each request [7]
- **Userlens is tackling customer success with early churn prediction.** YC says the product predicts churn months before it happens so CSMs can intervene proactively rather than reactively. Founders are Hai Ta and Ankur D. Launch page [8]
- **ClearSpec is trying to formalize the spec layer for coding agents.** The product turns meeting notes, chat, or guided inputs into structured specs with user stories, edge cases, security gaps, and acceptance criteria, then exports to GitHub, Linear, Jira, Cursor rules, Claude Code, Markdown, and Notion. The founder frames it as a response to "garbage in, garbage out" when feeding vague requirements into AI coding tools. Early access [9]

3) AI & Tech Breakthroughs

- **TurboQuant Pro packages vector compression into a practical OSS toolkit.** The MIT-licensed toolkit compresses embeddings and KV cache by 5-42x while maintaining 0.95+ cosine similarity, with benchmarks showing 0.97+ recall@10 on 2.4M real embeddings. The authors' practical recommendation is that Matryoshka truncation plus scalar int8 often beats more complex approaches for RAG, and the new autotune CLI can find a viable compression setting in about 10 seconds; one recommended configuration reached 20.9x compression at 96% recall@10 [10, 11]
- **Small open cyber models reproduced Anthropic showcase analyses.** In tests on the Mythos showcase vulnerabilities, 8/8 open models recovered the flagship FreeBSD exploit, including a 3.6B active-parameter model costing $0.11 per million tokens; a 5.1B-active model also recovered the core chain of a 27-year-old OpenBSD bug. The broader lesson from the write-up is that the AI cybersecurity frontier is "jagged": rankings reshuffle by task rather than forming a stable leaderboard [12, 13]
- **One striking bioinformatics demo compressed a full raw-genome**

workflow into $5 and 8 hours. In a workflow shared by Garry Tan, an AI agent autonomously retrieved 67GB of raw DNA, aligned 21 million long reads with 99.83% mapped, called 5.8 million variants, phased them, annotated them against ClinVar, PharmGKB, and gnomAD, and produced condition, medication, and nutrient reports. Garry's takeaway was that intelligence alone is not enough; the system still required clever orchestration [14, 15]

- **Safetensors is moving deeper into core ML infrastructure.** Hugging Face says the format has become the most popular way to share models safely, and the project is joining the PyTorch Foundation to scale further, including torch core integration [16]

4) Market Signals

- **Large-enterprise adoption is already material.** a16z says 29% of the Fortune 500 and roughly 19% of the Global 2000 are live, paying customers of a leading AI startup [17]
- **The strongest platform thesis remains "build software for agents, not just humans" — but enterprises still lack a budgeting model.** Aaron Levie argues agent-native software will increasingly prioritize APIs, CLI, and MCP-style interfaces, especially if companies end up supporting far more agents than people. The same a16z discussion flags a real adoption gap between startups and enterprises, with CFO/CIO resistance around integrations and broad uncertainty over whether token spend should be 1% or 100% of engineering budgets [18, 19, 20]

  "If you have a hundred or a thousand times more agents than people, then your software has to be built for agents…" [19]

- **Usability and usage-based pricing are starting to matter more than standalone benchmark wins.** Nathan Benaich highlighted an FT-linked report that Perplexity's revenue surged after launching Computer and usage-based billing, with the product reaching 100 million MAUs. His framing is simple: if models are not doing actual work for users, benchmark wins are no longer enough [21, 22]
- **Managed agent stacks are improving quickly, and developers are openly contesting where memory should live.** Anthropic launched Claude Managed Agents in public beta as a bundled harness plus production infrastructure, and Jerry Liu argues builders should avoid overcommitting to custom stacks because frontier labs are shipping wrappers fast enough to obsolete them. Letta and Harrison Chase argue stateful agent APIs are becoming the norm, but memory should stay outside model providers to avoid switching-cost lock-in; early efforts like WebMCP are already trying to expose product actions as AI-callable tools instead of

forcing UI automation [23, 24, 25, 26, 27]

5) Worth Your Time

- **AI adoption by the numbers** — the cleanest short read in this set on enterprise AI penetration inside large companies [17]
- **Plug and Play clip on GDPVal** — useful benchmark framing for economically valuable knowledge-work tasks across 44 occupations and nine major industries [28]

AI Investing in 2026: Bubble, Breakthrough, or Both? (15:47)

- **AI cybersecurity after Mythos: the jagged frontier** — worth reading for concrete evidence that small open models can reproduce high-value vulnerability analysis [12]
- **TurboQuant Pro** — practical open-source work on vector and KV-cache compression, pgvector integration, and autotuning for real data [10, 11]
- **LlamaIndex on VLM OCR failure modes** — useful diligence material for document-AI companies; it catalogs two production failure modes frontier VLM users will eventually hit: repetition loops and recitation errors [29, 30]

Sources

1. X post by @gitbutler

2. X post by @a16z
3. X post by @a16z
4. X post by @0x796F
5. X post by @lpolovets
6. X post by @alliekmiller
7. r/SideProject post by u/mmoustafa
8. X post by @ycombinator
9. r/SideProject post by u/xer2
10. r/MachineLearning post by u/ahbond
11. r/MachineLearning post by u/ahbond
12. X post by @ClementDelangue
13. X post by @stanislavfort
14. X post by @sowmay_jain
15. X post by @garrytan
16. X post by @ClementDelangue
17. X post by @a16z
18. X post by @a16z
19. X post by @a16z
20. X post by @a16z
21. X post by @sarthakgh
22. X post by @nathanbenaich
23. X post by @claudeai
24. X post by @jerryjliu0
25. X post by @sarahwooders
26. X post by @hwchase17
27. r/SideProject post by u/Sunnyfaldu
28. AI Investing in 2026: Bubble, Breakthrough, or Both?
29. X post by @llama_index
30. X post by @jerryjliu0