

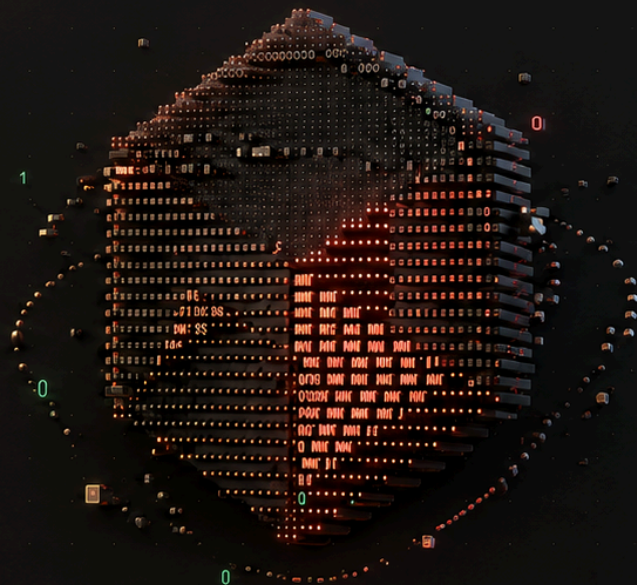


VIGOLIUM

// What is Vigolium?

SERIOUS SECURITY AUDIT, NOT JUST A PR REVIEW

Catch critical vulnerabilities that traditional scanners and AI code review tools miss, with validated proof your team can act on.



www.vigolium.com / jessie@vigolium.com

The Problem

Every company shipping software faces a painful tradeoff

- **Hiring pentesters** – \$20k-\$100k per engagement, takes weeks, and the report is stale the moment your team ships new code
- **Buying scanner software** – install it, configure it, tune it, and hire someone to run it. Most companies don't have that person
- **Bolting on an AI code reviewer** – reads the diff, not the app. Pattern-matches your PR against a training set. No runtime, no exploitation, no proof a bug is real

Meanwhile, you're shipping weekly. Your attack surface grows every sprint. Your security coverage does not.

The gap: Security testing is stuck in the consulting era – and the new wave of AI reviewers never leaves the diff. Nobody is actually *hacking* your running app.

Why AI Code Review Isn't Security Testing

The new wave of AI tools reviews your `diff`. A hacker attacks your `running app`. These are not the same job.

Scope

`AI code reviewers` see the 40 lines in your PR.

`Vigolium` ingests the entire repo – every route, every auth flow, every downstream service – plus the live running app.

You can't find a cross-endpoint auth bypass by staring at one file.

Action

`AI code reviewers` reason about code and leave suggestions.

`Vigolium` writes exploit payloads, fires them at your app, and watches the response.

Static reasoning can't prove a bug is exploitable. A real request can.

Signal

`AI code reviewers` produce prose hints – often wrong, always unvalidated.

`Vigolium` produces a finding with an HTTP request, a response, and a reproduction.

Your engineers don't need more opinions. They need proof.

What Vigolium Is

A hacker that reads your whole codebase – and actually breaks in.

- **Connect your app** – a URL, a GitHub repo, an API spec, or all three
- **Vigolium ingests the entire repo**, maps your attack surface, plans exploits, and **fires real payloads at your live app** – every time you ship
- **You get validated findings** – each with a real HTTP request, response, and reproduction. Not hints. Not diff annotations. Proof.



The Pentester That Never Sleeps

Always on. Always learning. On every deploy – without an ops team.

Never off

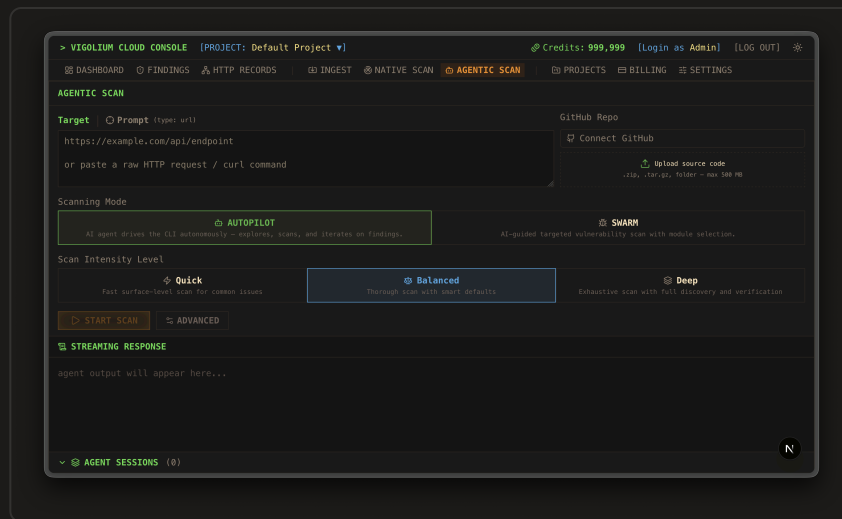
First scan in minutes – just a URL and optional GitHub connection. Not 3 weeks of scoping calls.

Never forgets

Every push scanned, every deploy watched. Not an annual snapshot you'll forget about by Q3.

Never quits

Learns from every app, every customer, every scan.
Not one expert's playbook, frozen in time.



Meet Agentic Mode (The AI Pentester)

An **AI security agent** that works the way a senior human pentester works – but at machine speed, and never gets tired.

1. **Reads your *entire* source code** – not just the diff. Maps every route, auth flow, data access path, and business-logic chain across files and services
2. **Decides what to attack** based on what it found – not a fixed checklist
3. **Writes custom exploit scripts on the fly** for logic flaws no generic scanner could catch
4. **Fires real payloads at your live app** and watches how it responds – exploitation with evidence, not suggestions on a PR
5. **Reviews every finding** and throws away false alarms before you ever see them
6. **Explains each real issue** in plain English, with a reproducible HTTP request and a suggested fix

All automatically. All in one click. A seasoned hacker – on retainer, at machine speed.

Two Modes, One Platform

Most "AI" security tools bolt a chatbot onto a traditional scanner. Agentic Mode is the opposite: the AI **drives the entire scan** – planning, probing, validating, and reporting.

	Native Scan	Agentic Scan
Speed	Very fast (seconds-minutes)	Deeper (minutes-hours)
Approach	Deterministic checklist of 200+ built-in checks	AI plans, writes custom tests, triages results
Best for	Every deploy, CI/CD gates, broad coverage	Pre-release audits, new features, sensitive apps
Finds logic flaws?	Limited	Yes – this is its strength
False-positive rate	Low	Near zero (AI triage)
Cost per scan	Low	Higher (real AI compute)

Competitive Landscape (Vigolium compares to the usual options)

What you want	Hire a pentester	Buy a scanner	Bug bounty	AI code reviewer	Vigolium
Instant setup	✗ (weeks)	✗ (days)	⚠	✓	✓ (hours)
Always on, not a snapshot	✗	⚠	✓	⚠ (per PR)	✓
Reads entire repo (not just the diff)	⚠	✗	✗	✗	✓
Runs against your live app	✓	⚠	✓	✗	✓
Produces validated PoC / evidence	✓	⚠	✓	✗	✓
Catches blind / runtime-only bugs	✓	⚠	✓	✗	✓
AI filters false alarms	✗	✗	N/A	✗	✓

Vigolium vs. Claude Code with Opus 4.7

STATIC AUDIT KICKOFF

```
Claude Code v2.1.114
Opus 4.7 (1M context) - Claude Max
~/Desktop/oss-to-run/codex

do a thoroughly static security audit on this repo for me

Skill(security-review)
└─ Successfully loaded skill

Listed 2 directories (ctrl+o to expand)

• This is the Codex repo (Rust + Node CLI + SDKs). I'll run a thorough static security audit across the codebase.

• Imagining (53s - 4.151 tokens - thought for 4s)
└─ Tip: Use /btw to ask a quick side question without interrupting Claude's current work

> |

>> bypass permissions on (shift+tab to cycle) - esc to interrupt
```

FINDING REVIEW AND FP CHECK

```
Agent(Verify PowerShell spawn finding)
└─ Initializing

Wrong! Log. (12s 11s - 4.44.8k tokens - thought for 4s)
Claude Code v2.1.114
Opus 4.7 (1M context) with xhigh effort - Claude Max
~/Desktop/oss/codex

• Let me verify the one HIGH finding with an independent false-positive check sub-task.

Agent(Verify PowerShell spawn finding)
└─ Done (25 tool uses - 73.8k tokens - 3m 19s)
(ctrl+o to expand)

• Codex Repo Static Security Audit

Scope: Full-repo audit across sandbox/exec, auth/secrets, MCP/plugin/IPC, and SDK/CLI/install surfaces. Trust model: LLM output, MCP server responses, and network content are UNTRUSTED; user CLI/env/config are TRUSTED.

Vuln 1: Sandbox escape via PowerShell-parser pre-spawn - codex-rs/shell-command/src/command_safety/powershell_parser.rs:104

- Severity: HIGH
- Category: sandbox_escape / command_injection
- Confidence: 9/10

Description

is_known_safe_command (codex-rs/shell-command/src/command_safety/is_safe_command.rs:22-24) unconditionally calls is_safe_command_windows on all platforms - the windows_safe_commands module has no MCFG(is_opsys)) gate, unlike its sibling is_dangerous_command_windows (is_dangerous_command.rs:13-3). When an LLM-supplied tool-call argv has command[0] whose basename matches powershell, powershell.exe, powershell, or pwsh.exe (windows_safe_commands.rs:107-110), the code invokes PowerShellParserProcess::spawn(executable), which executes Command::run(executable), argset["-nologo"], "-NoProfile", "-NonInteractive", "-SuccessCommand", "...")::spawn() - passing the LLM's raw path verbatim, with no PATH resolution, no absolute-path requirement, and no system-directory check.

This runs inside the orchestrator process during pre-dispatch is_muting_evaluation (codex-rs/core/src/tools/registry.rs:334 - before handle_argv), so the spawned process inherits the user's full ambient authority, outside the seatbelt/landlock/berap/Windows-sandbox that the later-child exec would have been wrapped in.

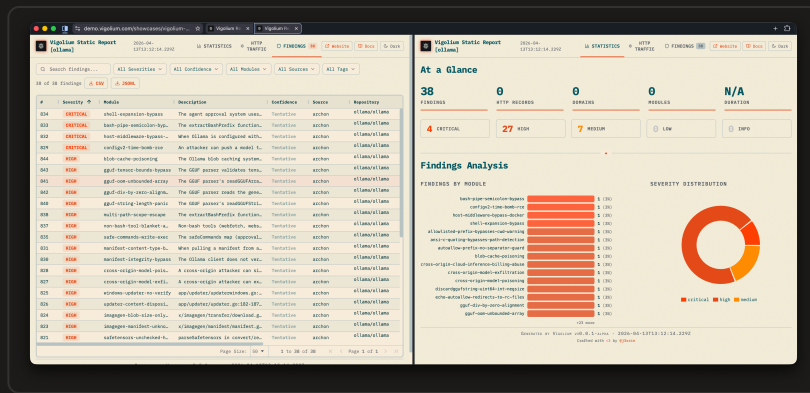
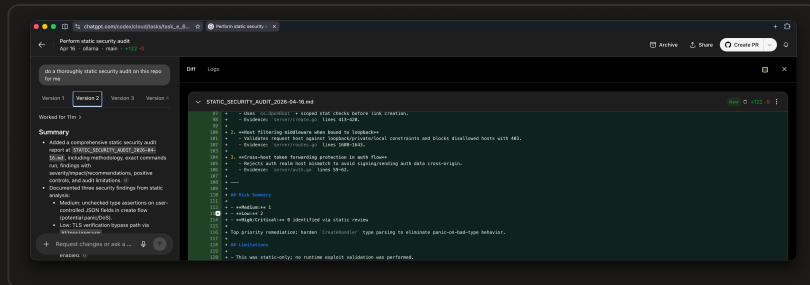
Call sites that feed untrusted LLM argv into the gate:
- codex-rs/core/src/exec_tool.rs:157
```

Takeaway: Claude Code is a strong coding agent. Vigolium is built for serious security audits. Vigolium found **37 critical and high severity vulnerabilities** compared to **2 findings in normal Claude Code**, even with skills enabled. Claude Code with Opus 4.7 can produce a solid **static security review**. Vigolium goes further: audit the **entire repo**, validate findings with **proof**, and deliver actionable results.

Vigilium vs. GPT-5.5 Cyber

Head-to-head on the exact job a security buyer is trying to do.

Dimension	GPT-5.5 Cyber	Vigilium
Findings	3 findings, low severity	38 findings, including criticals
Input scope	Diff / pasted file, misses most of the repo	Entire repo + live app
Method	Reasons about source	Writes exploits, fires them
Output	Natural-language hints	Validated finding + HTTP PoC
False-positive rate	High – not validated	Near zero – runtime-confirmed
Cross-file auth / IDOR chains	⚠️ partial	✅ whole-repo reasoning
Runtime misconfig	❌ static only	✅ observed live
Evidence a dev can act on	Prose	Request / response + repro



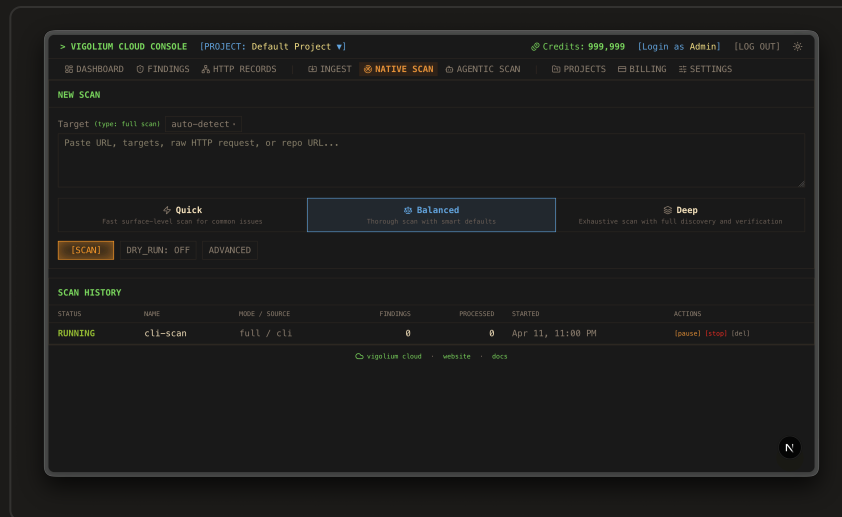
Technical Appendix

For folks familiar with the AppSec market

Capability	Veracode / SonarQube	Acunetix / Qualys WAS	HackerOne / Bugcrowd	Vigolium
Runtime DAST (tests live app)	⚠️	✓	✓ (human)	✓
SAST (reads source code)	✓	✗	✗	✓
AI-driven payload generation	✗	✗	N/A	✓
AI triage / noise filtering	⚠️	⚠️	✓ (human)	✓
Custom checks on the fly	✗	⚠️	N/A	✓
Pricing model	per-dev/mo	per-asset/mo	per-bounty	per-scan or per- app/mo

How It Works

1. Sign up at vigolium.com
 - ↓
2. Add your app
 - Paste a URL, API spec or
 - Connect GitHub, or
 - Just describe your requirement in english
 - ↓
3. Vigolium's AI reads your code, plans the attack, runs 200+ checks, and filters false alarms
 - ↓
4. Clean report in your dashboard
 - Prioritized by real risk
 - With clear fixes
 - ↓
5. Every new deploy → fresh scan, automatically.



From zero to your first real finding in under 10 minutes.

What We Catch (OWASP Top 10 and far beyond)

Everything a human pentester would look for – and a whole category of bugs that diff-scoped AI reviewers **structurally cannot reach**.

- **Data leaks** – sensitive info exposed in responses, headers, or error pages
- **Broken access controls** – users seeing or changing data they shouldn't
- **Injection attacks** – SQL, XSS, command injection, and modern variants
- **Authentication & session flaws** – weak logins, broken tokens, privilege escalation
- **Cloud misconfigurations** – leaked credentials, exposed storage buckets, risky API keys
- **Framework-specific bugs** – Next.js, Django, Rails, Spring, Laravel, FastAPI, and more
- **"Blind" vulnerabilities** – subtle bugs with no visible symptoms, caught via advanced callback techniques

What AI code reviewers can't catch and never will: Cross-endpoint auth bypass (needs whole-repo context) · IDOR chains spanning multiple files or services · Runtime-only misconfig (leaked headers, debug endpoints, env drift) · Business-logic race conditions

Why Customers Trust It

- **Benchmark-validated** against the industry-standard vulnerable apps every security vendor tests on
- **Real-world validated** through bug bounty programs finding previously-unknown bugs in production
- **Transparent AI** – every finding shows the reasoning and evidence; nothing is a black box
- **Data sovereignty options** – managed AI by default, or private deployment for enterprise

The screenshot displays the Vigolium Static Report interface for the target 'ollama'. The interface includes a search bar, filters for All Severities, All Confidence, All Modules, All Sources, and All Tags, and a '45 of 45 findings' indicator. A table lists findings with columns for #, Severity, Module, and Description. Finding #9 is highlighted, showing a severity of HIGH and a description of 'Three compounding defects'. A detailed view of finding #9 is shown on the right, including a summary, a list of three defects (Pipe-only splitting, Denylist as substring, and bash -c executor), and a cache key.

#	Severity	Module	Description
3	CRITICAL	ggvf-shape-unist4-overfl...	fs/ggf/gppl.go:505-514 cor
2	CRITICAL	puLWithtransfer-egest...	puLWithtransfer (Ollama's
1	CRITICAL	autoallow-prefix-metach...	IsAllowed (/x/agent/appr
10	HIGH	agent-approval-command-s...	Distinct primitive from p8
9	HIGH	agent-approval-shell-net...	Three compounding defects
8	HIGH	allowedhost-suffix-squat...	allowedHost() at server/ro
7	HIGH	stdc-null-derof-ineage-bl...	llama/llama.go:566-578 pas
6	HIGH	ggvf-string-unbounded-al...	Two independent GGVF parse
5	HIGH	manifest-token-oon	Two unbounded io.ReadAll(r
4	HIGH	api-pull-ssrf	POST /api/pull accepts a us
45	MEDIUM	weauth-reala-coma-pars...	The custom WWW-Authenticati
44	MEDIUM	ispriate-rc1918-host-h...	The host-header filter at:
43	MEDIUM	dns-rebinding-drive-by-l...	Chain of primitives, each:
42	MEDIUM	signin-ur-umarshal-err...	The original hypothesis (H
41	MEDIUM	whoami-public-key-disclo...	WhoamiHandler (server/rout
40	MEDIUM	editor-visual-flag-injec...	In interactive mode, the o
39	MEDIUM	web-search-fetch-umath...	Two gin-registered routes:
38	MEDIUM	readrequestbody-unbound...	readRequestBody at server/
37	MEDIUM	client2-experiment-bypas...	When OLLAMAXPERIMENT=cli
36	MEDIUM	ollama-host-nonloopback...	The allowedHostsMiddleware
35	MEDIUM	cgo-log-callback-reentra...	llama/llama.go:34-56 regis
34	MEDIUM	streaming-response-clean...	The ollama WJDSION streamin
33	MEDIUM	llama-adapter-lora-struct...	llama/llama.go:944-356 (Ap
32	MEDIUM	audio-nel-path-uncapped...	Ollama exposes two audio i
31	MEDIUM	cstring-leak-load-model...	llama/llama.go:264-318 (Lo
30	MEDIUM	lora-path-lpc-to-cgo-pas...	The llmarunner subprocess
29	MEDIUM	atxrunner-manifest-path...	/x/aiagegen/manifest/manif

FINDING #9 HIGH Tentative Draft
agent-approval-shell-metachar-bypass
Three compounding defects produce a single command-injection primitive in the agent's approval layer.

Summary
Three compounding defects produce a single command-injection primitive in the agent's approval layer:

- Pipe-only splitting:** `extractBashPrefix (/x/agent/approval.go:284-286)` only splits on `|` before extracting the prefix. Any `;`, `&&`, `|`, `{`, `{...}`, backticks, or newlines fall inside `parts[0]` – they are never examined.
- Denylist as substring:** `IsDenied (/x/agent/approval.go:94-122)`, called from `/x/cmd/run.go:378` uses `strings.Contains(cmd, pattern)` with literal patterns (`"rm -rf"`, `"curl -X POST"`, `"*/*/*/shadow"`, etc.). Any inline empty-quote (`"rm -rf"`, `"curl -X POST"`, `"*/*/*/shadow"`, etc.). Any inline bash still evaluates the intended command.
- bash -c executor:** `/x/tools/bash.go:64` invokes `exec.CommandContext(ctx, "bash", "-c", cmd)`. Bash is invoked AS a shell – metachars are interpreted by `/bin/bash`, not by Go's `exec`. This is the sole execution sink for approved commands.

The cache key produced by `extractBashPrefix` ignores everything after the first path-like argument. Once a user approves `cat /tools/run.sh` for the session, the prefix `cat/tools/` enters the allowlist, and the agent's subsequent tool-calls reach `IsAllowed=true` when their first pipe-segment's first non-flag path arg happens to start with `/tools/` – no matter what metachars follow.

Details
Three compounding defects produce a single command-injection primitive in the agent's approval layer:

- Pipe-only splitting:** `extractBashPrefix (/x/agent/approval.go:284-286)` only splits on `|` before extracting the prefix. Any `;`, `&&`, `|`, `{`, `{...}`, backticks, or newlines fall inside `parts[0]` – they are never examined.
- Denylist as substring:** `IsDenied (/x/agent/approval.go:94-122)`, called from `/x/cmd/run.go:378` uses `strings.Contains(cmd, pattern)` with literal patterns (`"rm -rf"`, `"curl -X POST"`, `"*/*/*/shadow"`, etc.). Any inline empty-quote (`"rm -rf"`, `"curl -X POST"`, `"*/*/*/shadow"`, etc.). Any inline bash still evaluates the intended command.
- bash -c executor:** `/x/tools/bash.go:64` invokes `exec.CommandContext(ctx, "bash", "-c", cmd)`. Bash is invoked AS a shell – metachars are

Proof: Real Code, Real Bugs

We pointed Vigolium at **some of the world's most popular open-source projects** – the same code running inside Fortune 500 companies – and it found real, reportable vulnerabilities.

Pricing

Security scanning that fits your stage: pay-as-you-go scans, pre-paid credits, team volume, or dedicated enterprise infrastructure.

ON-DEMAND SCAN

\$29

per 100K lines of code

Pay-as-you-go scan for vibe-coded apps, one-off audits, or benchmarking other scanners.

- ✓ One-time full agentic scan
- ✓ Pay only for what you scan
- ✓ Validated PoC for every finding
- ✓ Markdown, PDF & JSON report export
- ✓ No subscription, no commitment

STARTER PACK

From \$299

pre-paid credits

Pre-paid credits for MVPs and small teams with validated PoC on every finding.

- ✓ 5,000 scan credits included
- ✓ Native + Agentic scans
- ✓ Validated PoC for every finding
- ✓ Markdown, PDF & JSON report export
- ✓ Email support

TEAM PACK

From \$2,999

volume credits

Volume credits for production with continuous monitoring and team collaboration.

- ✓ 60,000 scan credits included
- ✓ Everything in Starter
- ✓ Scheduled & continuous scanning
- ✓ Continuous monitoring
- ✓ Cloud dashboard & scan history
- ✓ Priority support

ENTERPRISE

Custom Pricing

dedicated deployment

Dedicated infrastructure, SLA, and white-glove onboarding for large teams.

- ✓ Everything in Team
- ✓ Isolated data environment
- ✓ Custom integrations
- ✓ Custom SLA & uptime guarantees
- ✓ SSO / SAML integration
- ✓ Custom integrations & webhooks
- ✓ On-premise deployment option

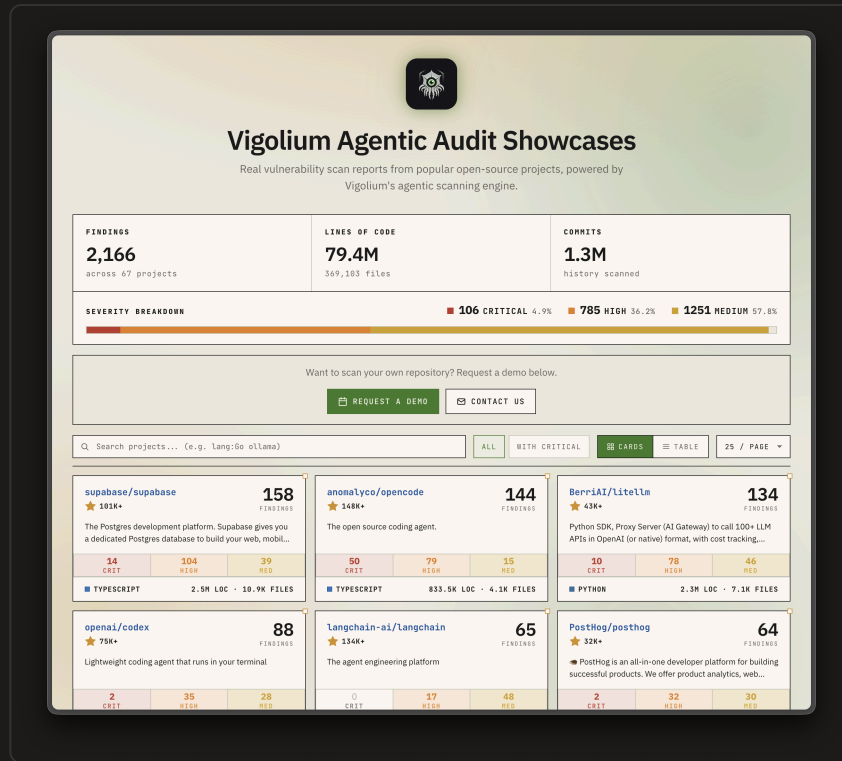
Traction & Go-to-Market

Where we are today

- **200+** built-in security checks live in production
- **Continuous scanning** with GitHub, GitLab, Bitbucket integrations
- **Dashboard & API** for team workflows
- **67+** open-source projects scanned, **2,166+** real findings surfaced

Go-to-market

- **Bottom-up:** free tier → self-serve upgrade (product-led growth)
- **Top-down:** design partners at scaleups → enterprise tier



The Ask

For investors

Security testing is a **\$10B+ market** stuck in the consulting era.

We're building a **continuous, AI-native, self-serve** AppSec platform – replacing one-off audits with always-on coverage.

Raising a seed round to scale the engine, grow the team, and land our first 20 design partners.

For design-partner customers

Free tier during beta, direct line to the founding team, and you shape the roadmap.

Get continuous AI-driven security review on your real codebase – and findings you can act on the same day.

Ideal fit: **Series A-C startups** without a dedicated AppSec team.

For partners & advisors

We're looking for **integration partners** (CI/CD, code hosts, ticketing) and **advisors** with deep AppSec, GTM, or enterprise-security experience.

Intros to **CISOs and Heads of Security** are the single highest-leverage thing you can offer us right now.

Reach out: jessie@vigolium.com · www.vigolium.com

Demo

Live, on stage, in under 5 minutes

1. Request Demo at www.vigolium.com
2. Connect a sample GitHub repo
3. Watch AI read the code and plan the attack in real time
4. Show the first real vulnerability found

See demo result and audit result: <https://demo.vigolium.com/>
or reach out to me at jessie@vigolium.com