COGN1TUM

FOR IMMEDIATE RELEASE

Cognitum One Unveils MetaHarness, the Open-Source Factory for Branded, Governed AI Agent Harnesses, in Conjunction with GitHub Universe 2026

With over 100 Million downloads and over 196,000 GitHub stars, MetaHarness arrives as the next open-source project of the prolific creator, Reuven "rUv" Cohen.

SAN FRANCISCO, Calif., October 27, 2026. Every team building AI agents ends up rebuilding the same plumbing: the command-line tool, the tool connections, memory, safety rules and release checks that sit around a model. Cognitum One™, Inc. (cognitum.one), the sovereign agentic control plane company, today announced MetaHarness™, in conjunction with GitHub Universe 2026, which takes place October 28 and 29 at Fort Mason Center in San Francisco. MetaHarness is an open-source factory for agent frameworks. Starting from a GitHub repository or a blank page, it generates a branded agent harness that a team owns and can publish under its own name. Each harness comes with its own npx command-line tool, Model Context Protocol (MCP) server, scoped memory, governance policy and witness-signed releases. The model can be swapped out. The harness is what the team builds and keeps.

MetaHarness doesn't stop at the first attempt. It tests different ways for an AI agent to work, builds on changes that perform better and pass safety checks, and uses those results to guide the next round—all without changing the AI model itself. These capabilities support Cognitum One's vision for AI that businesses can control: the company sets the rules, the system checks tool actions against those rules, and records them for review.

MetaHarness comes from the developer behind some of the most widely used open-source AI infrastructure on GitHub. Cohen's ruvnet projects, including RuView™, Ruflo™ and RuVector™, have passed 100 million downloads across npm and crates.io and earned more than 196,000 GitHub stars from a community of more than 12,000 GitHub followers. MetaHarness is also ranked #2 on the OpenEnv Arena leaderboard, a new benchmark in which AI agents design the reinforcement-learning environments used to train and test a model.

"Most teams treat the model as the product and the harness as an afterthought. We think it is the other way around," said Reuven "rUv" Cohen, Founder and Chief AI Officer of Cognitum One and creator of MetaHarness. "The model changes every few months. The harness holds your tools, your memory, your safety rules and your release history. MetaHarness lets a team mint that harness in minutes and put its own name on it. Then it builds many versions, races them against each other on real work, and keeps the one that wins, without touching the model. Shipping it alongside GitHub Universe makes sense to us, because this community made these projects what they are."

BY THE NUMBERS: rUv AND RUVNET ON GITHUB

100M+
downloads across ruvnet npm and crates.io packages

196,000+
GitHub stars across ruvnet's open-source repositories

#2
on the OpenEnv Arena leaderboard, at 22.5% across 40 held-out tasks

15M+
npm downloads in September 2026 alone, up from about 300,000 a year earlier

878
published packages: 398 on npm and 480 on crates.io

2,254
passing tests across 246 files in the MetaHarness repository

- **Flagship projects.** RuView has passed 100,000 GitHub stars, Ruflo more than 74,000 and RuVector more than 4,400. MetaHarness has passed 700 stars and 88 forks.
- **Growth.** ruvnet packages have recorded more than 13 million npm downloads every month since May 2026.

Explore MetaHarness: https://github.com/ruvnet/metaharness
Get started: `npx metaharness --wizard`

What MetaHarness Does

- **Generates a harness from your repository.** `harness analyze-repo` reads a repository's structure and recommends a harness before anything is built. Analysis is read-only and never runs repository code. Inferred build and test commands stay disabled until someone explicitly trusts them.
- **Runs on the hosts teams already use.** Generated harnesses target ten hosts: Claude Code, Codex, GitHub Copilot, OpenCode, pi.dev, Hermes, OpenClaw, RVM, Prime Agent and GitHub Actions. Nineteen quick-start templates cover verticals including coding, DevOps, research, trading, legal, support, sales and education.
- **One shared Rust kernel.** Every harness uses `@metaharness/kernel`, a Rust core shipped as WebAssembly (WASM) and native bindings. It provides governance, memory, routing and witness signing. Host-specific behavior stays in thin adapters.
- **Tools with a policy boundary.** MCP tool access is denied by default. Teams then choose one of three modes: MCP turned off entirely, local connections only, or authenticated remote connections. `harness mcp-scan` checks tool configurations for risky grants and missing audit or timeout settings without running the tools.
- **Routes against a quality bar.** `@metaharness/router` predicts each candidate model's quality on a team's labeled examples and picks the cheapest one expected to clear the team's threshold.
- **Builds many versions and keeps the best.** Darwin Mode generates competing variants of a harness, runs each one in a sandbox against the same tasks and keeps only those that measurably beat the current version and pass its safety gates. The underlying model stays frozen, and the defaults need no network and no API keys.
- **The flywheel: run, measure, mutate, verify, promote.** `@metaharness/flywheel` turns that search into a compounding loop. Each generation starts from the last proven winner, and a candidate is promoted only if it clears a frozen gate, including held-out tasks it was never tuned on.

Every promotion is signed with Ed25519, so an outside reviewer can replay the full lineage and confirm each gain was real.

- **Reviewable releases.** Ed25519 witness signing, SPDX-2.3 software bills of materials and release gates record what was generated and what shipped. Publishing stays an explicit maintainer action.
- **Red-team and blue-team testing.** `@metaharness/redblue` attacks an agent a team owns with adversarial models, covering prompt injection, tool misuse, data leakage, jailbreaks and runaway costs. It then patches the vulnerable areas with declarative rules, retests and produces a board-readable report. Real credentials, shell access and live external targets are blocked in code.

Proven on Public Benchmarks

MetaHarness's published benchmarks show that the right harness can let lower-cost models match or beat frontier models on specific workloads, at a fraction of the cost:

- **Research (DRACO).** Claude Haiku 4.5 scored higher than Claude Opus 4, 0.7566 versus 0.7143, at about one-tenth the calculated token cost, delivering 10.6 times more quality per dollar.
- **Function calling (BFCL).** MetaHarness's DeepSeek-V4-Pro configuration reached 96% accuracy versus 83.3% for GPT-5.2 across the same 150 tasks, at about 54% lower cost.
- **Research and assistant questions (FRAMES).** DeepSeek-V4-Pro matched GPT-5.2's 42.7% observed accuracy on the same 150 questions while costing about 79.5% less per question.
- **Software engineering (SWE-bench Verified).** A GLM-to-Opus cascade resolved 278 of 500 issues, or 55.6%, at an estimated $0.15 per issue.

Leading on OpenEnv Arena

Most AI benchmarks measure how well a model answers questions that people wrote. OpenEnv Arena flips that. AI agents design and code the reinforcement-learning environments, then the arena trains an open model on those environments and scores the result on held-out tasks across eight domains: cybersecurity, finance and economics, industrial and physical systems, math and formal reasoning, media and content production, natural science, office and white-collar work, and software engineering. It is designed to be hard. Every task is held out, and the agent never sees the test it is training a model for. As of October 26, 2026, no entrant had solved more than a quarter of the 40 tasks.

As of October 26, 2026, MetaHarness ranks #2 on the leaderboard with a 22.5% average, solving 9 of 40 held-out tasks. Its strongest domain is office and white-collar work, where it solved 3 of 5 tasks, followed by industrial and physical systems at 2 of 5. The leaderboard keeps each entrant's best result in each domain across submissions, which shows where an agent's environment designs are strong and where the next experiments should focus. Earlier this month, MetaHarness was tied for the top score at 15%.

Why It Matters

AI agents are capable, but regulated enterprises cannot deploy what they cannot govern. Auditors want to know what an agent did, where it ran and whether that can be proven. Evaluation has become a market of its own: on October 8, 2026, Arena, formerly LMArena, announced a $200 million Series B at a $3.1 billion valuation. MetaHarness brings that discipline into the agent itself. Every harness it generates ships with policy boundaries, adversarial testing, quality-gated model routing and a signed record of each release.

Enterprises do not want another framework to babysit. They want an agent they own, with their name on it, that they can prove is safe before it reaches production. MetaHarness delivers that in a single command. Teams can start with one harness on one repository and grow into governed agents across the business, running under Cognitum One's control plane alongside Ruflo and ruOS on infrastructure they control.

Getting Started

- **Guided setup:** run `npx metaharness --wizard` and pick a template and host.
- **Claude Code:** run `npx metaharness my-bot --template vertical:coding --host claude-code`. For Codex, use `--host codex`.
- **From an existing repository:** run `harness analyze-repo . --scaffold my-bot`, then validate it and publish under your own package name.
- **Security testing:** run `npm i @metaharness/redblue` and point it at a local copy of the agent you own.

Cognitum One recommends starting with one harness on one repository, reviewing its generated policy, and adding Darwin Mode, routing and red-team testing as each one earns its place.

Availability

MetaHarness is available today as a v0.1.x beta. It is open source under the MIT license at github.com/ruvnet/metaharness, with the `metaharness` command-line tool and the `@metaharness/*` packages on npm. Model provider subscription or usage costs still apply. Experimental modules, including AVO governed variation, field memory, the ARC-AGI-3 harness and Weight-EFT, are clearly labeled and carry their own requirements. Cognitum One offers commercial licensing, enterprise support and deployment with Ruflo and ruOS™ on infrastructure the customer controls. To learn more, visit cognitum.one/metaharness.

Resources

- MetaHarness on GitHub: https://github.com/ruvnet/metaharness
- Cognitum One: https://cognitum.one/metaharness
- Visual walkthrough: https://github.com/ruvnet/metaharness/blob/main/docs/visual-walkthrough.md
- MetaHarness explainer: https://metaharness-explainer.vercel.app/
- redblue on npm: https://www.npmjs.com/package/@metaharness/redblue
- MoE-Foundry: https://github.com/ruvnet/MoE-Foundry and https://expert-foundry.ruv.chatgpt.site/
- OpenEnv Arena leaderboard: https://openenvarena-arena.hf.space/api/leaderboard
- ruvnet registry statistics: github.com/ruvnet/ruvnet/.../data/registry-stats.json

About Cognitum One

Cognitum One is an ambient intelligence company building the sovereign agentic control plane for regulated enterprises. The control plane is the governance layer between enterprise workflows and the foundation models they use. The enterprise sets the rules, the platform enforces them, and every agent action is recorded with a cryptographic witness that can be exported to an auditor. Policy is centralized and execution stays local, on premises, inside trusted infrastructure such as PostgreSQL or on sovereign edge appliances, so data and intellectual property never leave the customer's boundary. The Cognitum stack perceives through RuView camera-free sensing. It reasons with ruvLLM private inference and RuVector self-learning memory, acts through Ruflo multi-agent orchestration with native MCP integration, and adapts through SONA on-device continuous learning, all on the CPUs, GPUs, NPUs and edge devices an enterprise already owns. Cognitum One serves organizations in defense and government, financial services, healthcare and life sciences, and cybersecurity. The company is led by Founder and Chief AI Officer Reuven Cohen, creator of the widely used open-source AI projects RuView, Ruflo, RuVector, and MetaHarness, and Founding CEO D. Christopher Keil. To learn more, visit cognitum.one.

GitHub and GitHub Universe are trademarks of GitHub, Inc. Claude, Claude Code, Claude Haiku and Claude Opus are trademarks of Anthropic, PBC. OpenAI, GPT and Codex are trademarks of OpenAI. All other trademarks are the property of their respective owners. Cognitum One is not affiliated with or endorsed by GitHub, Anthropic, OpenAI, DeepSeek, Hugging Face or Arena. *Benchmark results are Cognitum One's own published runs in the ruvnet/metaharness repository and apply to the workloads tested: DRACO costs use blended per-token price estimates; BFCL measures single-turn function calling; FRAMES used lightweight retrieval, so absolute scores are not comparable to external FRAMES leaderboards; the SWE-bench Verified cost is an estimate. The OpenEnv Arena ranking and scores are from the public leaderboard as of October 26, 2026, and combine the best per-domain results from four MetaHarness submissions made between October 9 and October 18, 2026; the leaderboard is live and rankings may change. The download figure combines 95.7 million npm downloads across ruvnet packages in the 12 months through October 10, 2026, 1.8 million cumulative crates.io downloads, and npm downloads since October 10 at October's daily rate, per the public ruvnet/ruvnet registry statistics. Star and follower figures for the ruvnet GitHub profile and RuView are as of October 27, 2026; Ruflo and RuVector figures are as of October 12 and October 19, 2026. MetaHarness star, fork and test figures are from the public ruvnet/metaharness repository as of October 26, 2026. Arena funding figures are from Arena's announcement of October 8, 2026.*

Media Contact

Nick Ruest
Chief of Staff, Cognitum One, Inc.
nick@cognitum.one
(416) 477-3336

###