mario.brajkovski.14@gmail.com
superbavario.dev
github.com/mariobrajkovski
linkedin.com/in/mario-brajkovski
Munich · Remote

Mario Brajkovski

Research Engineer · RL environments, verifiers & agent evaluation

I build RL environments for AI agents end to end: realistic work environments, hard tasks, plausible data, tool implementations and verifiers, then I run agents in them and fix whichever side is wrong. At **HUD** (YC W25, employee #7) I built 16 RL and eval environments and shipped 1,000+ tasks to two of the top three frontier labs, as one of the three people behind the company's revenue-generating deliveries, while running quality across 5+ external vendor teams. Co-author of **ZeroDayBench** (ICLR 2026 Workshop). 5+ years of production SRE underneath.

RL ENVIRONMENTS & EVALUATION

Environments across industries and capabilities

16 environments · 1,000+ tasks shipped

Healthcare: environments on two open-source health-record systems for tool-use and computer-use agents, with verifiers that check GUI actions against the real database state and 250+ task extractors and graders. **Web research:** an archive-replay browser environment (mitmproxy + pywb) for reproducible tasks on real sites. **Software & SRE:** bug-fix and incident-response tasks with hidden tests on real codebases. Also professional knowledge-work, enterprise-tool, game and robotics environments.

Verifiable task design

tasks, data, verifiers

Tasks, data, tools and verifiers designed together, so a high score means the work was actually done. Every grader is run against deliberately broken solutions and cheating agents, prompt and grader are checked for agreement, and difficulty is calibrated across model tiers before anything ships.

Trajectory review, annotation & vendor quality

hundreds of hours in agent traces

Built automated QA that separates verifier bugs from genuine agent failures (e.g. rubric checks the prompt never asked for), and an LLM-judge pipeline classifying failed trajectories into task-specific failure modes. Ran the vendor side from zero: onboarding, annotation and QA workflows and quality enforcement across 5+ external vendor teams, plus ~95% of customer technical communication, turning lab feedback into task specs and reward-design changes.

Zero-to-one coding evals · ZeroDayBench

coding · ICLR 2026 Workshop

Turning a tested module into a blank-workspace coding environment with hidden tests protected against cheating and prompt-verifier alignment checks. **ZeroDayBench** (co-author): can agents find and patch 22 novel critical zero-days in real code? Measured GPT-5.2, Claude Sonnet 4.5, Grok 4.1.

EXPERIENCE

Member of Technical Staff / Research Engineer · HUD

Remote · Jul 2025 – Jul 2026

RL environments and agentic evals for frontier AI labs. Employee #7 (YC W25).

- **Built the environment and evaluation stack** above; designed eval infrastructure for an unreleased frontier model from scratch.
- **Agent-native daily workflow:** run fleets of coding agents (Claude Code) for building, reviewing and auditing; built and run TwoPeas, a paid real-time voice AI product, almost entirely through agents.

Site Reliability & DevOps Engineer · Founder

2019 – present

Joyn (2019–2023): streaming-platform SRE, Kafka, observability, incident response. **Cloudsurf** (2024–2025, for HanseMerkur): led an on-prem to AWS EKS migration. **TwoPeas** (2025–): voice-first AI companion with paying subscribers, sub-second WebRTC voice. **Tutor.mk**: AI tutor with 320+ active students.

SKILLS

RL ENVIRONMENTS	Environment & task design · verifier and reward design · mutation-tested graders · reward-hacking analysis · pass@k calibration · trajectory review · LLM-as-judge · synthetic task and data pipelines · computer-use, browser and tool-use agents · MCP
ENGINEERING	Python · TypeScript · Bash · SQL · Docker · Kubernetes · AWS · Linux · Kafka · CI/CD
LANGUAGES	Macedonian (native) · English (C2) · Serbo-Croatian (B2) · German (B1)