

Raj Koli

India

koliraj911@gmail.com | +91 9619564351
github.com/Rajkoli145 | [LinkedIn](#) | [Portfolio](#)

Professional Summary

Software engineer focused on backend systems, AI agent infrastructure, and developer tooling. Experienced in building full-stack applications, REST APIs, Linux automation, CLI tools, and production-ready workflows. Active open-source contributor in AI agent evaluation and terminal benchmarking. Currently authoring *The Agent Systems Handbook*, a research-driven guide covering modern AI agent architectures, memory, planning, evaluation, and autonomous software systems.

Technical Skills

AI & Agent Systems	OpenAI API, AI Agent Systems, Prompt Engineering, Agent Evaluation, Terminal Bench, T-Bench, Harbor, Workflow Design
Backend	Node.js, Express, REST APIs, WebSockets, JWT Authentication, Middleware
Systems / CLI	Linux Automation, Shell Scripting, CLI Design, Docker, File I/O, systemd, launchd
Frontend	React, Next.js, HTML5, CSS3, Tailwind CSS, JavaScript, TypeScript
Databases	MongoDB, Firebase
Languages	JavaScript, TypeScript, Python, Bash/Shell, SQL
Testing	pytest, Postman, Swagger, API Validation, Error Handling
Tools	Git, GitHub, GitHub Actions, VS Code, Vercel, Render
Practices	Debugging, Code Review, Technical Writing, REST API Design, CI/CD, Agile

Experience

AI Evaluation Specialist — Handshake AI (Project Dynamo) Jul 2026 – Present
Remote

- Evaluate AI models on terminal-based software engineering tasks involving debugging, code generation, and reasoning.
- Assess model performance across real-world software engineering workflows and provide structured evaluation feedback to improve AI system performance.
- Work with GitHub repositories, terminal environments, and software engineering best practices to evaluate and improve AI systems.

AI Benchmark Contributor (Trial) — Parsewave Jun 2026
Remote

- Designed and completed Terminal-Bench style benchmark tasks for AI agents in a structured evaluation environment.
- Developed deterministic task environments, automated validation workflows, and reproducible evaluation artifacts for model assessment.
- Successfully completed a paid benchmark contribution that passed technical review and quality evaluation.

Frontend Developer Intern — Unified Mentor Sep 2024 – Dec 2024
Remote

- Refactored UI into modular, reusable React components adopted across the codebase.
- Reviewed and merged pull requests while maintaining code quality and coding standards.
- Delivered product features throughout a four-month internship using Agile development practices.

Engineering Projects

FreelancerFlow — Freelance Management Platform [GitHub](#)
React, Node.js, MongoDB, Tailwind CSS, JWT

- Built a full-stack platform for project tracking, invoicing, payments, and client management.
- Designed a structured export pipeline that streamlined invoice generation.
- Secured APIs using JWT authentication and middleware validation.

- Implemented automated endpoint testing for critical backend routes.

EduStory — AI-Powered Narrative Platform

Next.js, MongoDB, OpenAI API, NextAuth

[GitHub](#)

- Developed an AI storytelling platform using structured prompt pipelines and OpenAI APIs.
- Improved AI response consistency through prompt validation and output verification.
- Implemented secure authentication with NextAuth and role-based access control.

Research Projects

The Agent Systems Handbook — Ongoing Research

- Authoring a comprehensive research handbook on AI agent systems, covering agent architectures, planning, memory, reasoning, tool use, multi-agent collaboration, evaluation frameworks, and production deployment.
- Synthesizing insights from academic papers, open-source frameworks, and production AI systems into a practical engineering reference for building autonomous software agents.
- Researching modern agent ecosystems including orchestration frameworks, retrieval-augmented generation (RAG), long-term memory, agent communication protocols, observability, and evaluation methodologies.

Startup Research Handbook

[GitHub](#)

- Developing a structured handbook covering startup validation, competitor research, product strategy, MVP planning, and execution frameworks.
- Creating reusable methodologies for evaluating software businesses through technical, product, and market analysis.

Open Source & Engineering Contributions

RustChain Installer Contributor

2026

Remote

- Built a cross-platform CLI installer automating blockchain node setup, wallet configuration, and service registration.
- Implemented SHA-256 verification and systemd/launchd automation; merged and awarded a project bounty.

SDK Developer — La Tanda

2025

Remote

- Built a production TypeScript SDK with modular HTTP client architecture and centralized error handling.
- Validated API implementations against Swagger documentation before release.

Hiero SDK Python Contributor

2026

Remote

- Contributed to the official v0.2.2 release by improving CI dependency resolution.
- Collaborated with 20+ open-source contributors using GitHub Actions and branch-based workflows.

Research Interests

AI Agent Systems • Multi-Agent Systems • Autonomous Software Engineering • AI Evaluation • Developer Infrastructure • Distributed Systems

Education

BTech in Computer Science — ITM Skills University

Aug 2024 – 2028

12th Grade (PCMB) — Score: 80%

2024