

HUSEYIN KAPLAN

Applied AI Engineer | RAG, Agent Systems, LLM Evaluation

Istanbul, Turkiye | Remote B2B / EOR | +90 533 689 59 04 | hello@huseyinkaplan.dev | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

Applied AI and full-stack engineer with 4.5 years building B2B SaaS and LLM-enabled products. Delivers retrieval, citations, abstention, tool calling, human handoff, evaluation gates, APIs, and product interfaces in TypeScript/Node.js and Python. Public work proves tested RAG, deterministic evaluation, and agent-runtime trade-off analysis.

CORE SKILLS

Applied AI: RAG, embeddings, hybrid retrieval, RRF, reranking, citations, abstention, tool calling, agent workflows, evaluation datasets | Engineering: TypeScript, Node.js, Python, FastAPI, Django, React, Next.js, PostgreSQL, Supabase | Delivery: Docker, GitHub Actions, Vercel, Playwright, Jest, Vitest

PROFESSIONAL EXPERIENCE

Full Stack Engineer - Revotas

Istanbul | Apr 2023 - Present

- Built a private document assistant over company files with parsing, embeddings, retrieval, provider integrations, and streamed answers.
- Architected configurable recommendation-delivery and search-merchandising interfaces used across multiple brand deployments; owned product and integration layers, not recommendation-model training.
- Delivered React/Next.js product surfaces backed by Node.js, Django, and PostgreSQL services; handled production debugging across Firebase, RabbitMQ, APIs, and legacy systems.
- Reduced Largest Contentful Paint by 40% using code splitting, lazy loading, and SSR/ISR patterns.

Freelance Full-Stack Engineer - Upwork

Remote | Dec 2022 - Apr 2023

- Delivered dashboards, landing pages, and data-driven product interfaces for international clients from requirements through deployment and iteration.

Front End Developer - Kobisi

Remote | Jan 2022 - Dec 2022

- Built responsive e-commerce interfaces with reusable React/Vue components and backend-integrated user flows.

SELECTED WORK

- Agent Runtime Bench - Implemented one refund workflow as a native tool loop and LangGraph graph. Three deterministic scenarios produce identical outcomes; native path is 50 source lines versus 97 for the graph. Live server-run demo, tests, benchmark, and CI.
- LLM Reliability Lab - Deterministic CI gate for citation precision/recall, expected-fact coverage, abstention policy, latency SLOs, and estimated cost. Typed API plus JSON/HTML reports; no LLM judge or API key.
- Production RAG and Evals Reference - Runnable FastAPI reference with overlapping chunks, lexical/vector ranking, RRF, citations, abstention, Docker, tests, and CI. Deterministic fixture reports Hit@3 1.0 and MRR 1.0; presented as fixture validation, not real-world quality.
- Tedavio Clinic AI - Private demo-first MVP with scoped retrieval, tool-driven appointment intent, deterministic human handoff, admin workflows, and Supabase row-level security.

EDUCATION, LANGUAGES, AVAILABILITY

B.Sc. Software Engineering - Istanbul Okan University, in progress | Associate Degree, Computer Programming (English) - Istanbul Aydin University, 2022

Harvard CS50 | CS50 Web Programming | Turkish: Native | English: Professional working proficiency | Available for remote B2B contractor or EOR roles across EMEA, Europe, Turkiye, or Worldwide.