Maksim Khovrov

AI Automation Engineer

mkhovrov01@gmail.com · +381 65 810 6509 · t.me/khovre · linkedin.com/in/maksim-khovrov-113633293 · github.com/mkhovrov01 · khovrov.dev

Summary

AI automation engineer with 4.5 years in development and automation, the last 2+ focused on production LLM systems: RAG pipelines, conversational AI agents, and self-hosted AI infrastructure. Shipped 30+ voice AI solutions to production and built a governed retrieval system that now runs five self-hosted personal projects end to end.

Experience

MCN Telecom — Moscow / Remote
Product Manager / Automation Engineer — Mar 2023 – Nov 2025
Promoted from QA Engineer, Jul 2021 – Mar 2023

- Launched 30+ voice AI solutions into production across medical, retail, recruitment, financial, and municipal-infrastructure clients.
- Designed and shipped 3 RAG pipelines (Qdrant, Zep) for client knowledge bases, raising response accuracy from 70% to 98%.
- Migrated the voice pipeline from a step-based to a real-time streaming architecture.
- Led 15+ technical audits and supported commercial proposals that closed $400K+ in deals.
- Scaled the automation team from 3 to 10+ people; onboarded and mentored developers, ran code review, maintained the internal knowledge base.
- Designed integration architectures with Bitrix24, amoCRM, 1C, and IP telephony; built alerting and error-handling for high-uptime external API integrations.

Personal Projects

Self-hosted and self-built. Portfolio and live demos: khovrov.dev

- **Mnemosyne —** Access-controlled knowledge retrieval broker — every AI query is scoped and re-validated against a canonical registry before it's served, so the access boundary is structural, not a prompt instruction. Python, PostgreSQL, Qdrant, FastAPI, MCP.
- **Personal AI Assistant —** Self-hosted chat assistant (Open WebUI) grounded through Mnemosyne, answering with citations instead of improvising. FastAPI, Qdrant, OpenRouter.
- **KB Auditor and RAG Builder —** Agent-orchestrated pipeline that crawls a public knowledge base, audits it, builds a RAG index, and evaluates retrieval quality end to end; n8n runs the agentic stages. Next.js, PostgreSQL, Prisma, n8n.
- **Conlang Audio Recognition —** Speech-ML project that designs a constructed-language vocabulary and its recognizer together, screening for phonetic distinctiveness before any audio is recorded; benchmarked against three baselines. Python, PanPhon, WavLM.
- **Document Editor —** Turns long source documents into structured, cited drafts; isolates OCR/PDF reconstruction on a separate untrusted-compute worker behind a versioned schema contract. FastAPI, React, TypeScript, SQLite. Live demo: editor.khovrov.dev.

Skills

- **Languages:** Python, JavaScript, TypeScript, SQL
- **AI & Retrieval:** RAG, LangChain, OpenAI API, Qdrant, Zep, Prompt Engineering, MCP
- **Backend & Infra:** FastAPI, Next.js, PostgreSQL, Docker, Caddy, Linux, Git
- **Automation & CRM:** n8n, REST APIs, Webhooks, Bitrix24, amoCRM, 1C, Swagger

Education

Software Engineering, MIREA — Russian Technological University, Moscow (incomplete, 2023)

Languages

Russian — Native · English — C1 (https://cert.efset.org/en/7MbYem)