

Nikolas de Hor

AI Engineer · MCP & Agent Systems

São Paulo, Brazil (remote) | nikolasdehor79@gmail.com | dehor.dev
linkedin.com/in/nikolasdehor | github.com/nikolasdehor

SUMMARY

AI Engineer focused on production agent systems: tool calling, MCP servers, RAG and memory, evaluation with golden sets, and per-turn cost and latency control, including multimodal flows (voice, images, documents). Author of MCP Fiscal Brasil (45 tools, published on PyPI). Works end to end, from architecture to operations, with Python, TypeScript, and AWS.

EXPERIENCE

AI Engineer, CAST (remote) Jun 2025 - Present

- Lead architect of LUME, an enterprise hub for AI agents: multi-agent orchestration, RAG, per-project memory, model governance, and cost, latency, and token tracking.
- Agent lifecycle (create, curate, version, run) with secret hygiene and usage feedback.
- Python and TypeScript/Node.js services on AWS, with Docker and CI/CD.
- AI across the engineering lifecycle, including internal technical-documentation quality tools for a client in the regulated banking sector.

Co-founder & COO, FOR6 Solutions (remote) Nov 2024 - Present

- Technology boutique delivering full-stack products and AI and automation systems for clients; leads product direction, architecture, and delivery.

Founder & AI Engineer, DeHor Labs (part-time) Feb 2022 - Present

- Products built on AI agents, MCP servers, and WhatsApp automation: AgroVoz, PlantonIA, TranscreveAI, and MCP Fiscal, Agro, and Jurídico Brasil, with end-to-end ownership.

Rubi S.A. (Rubiataba-GO, Brazil) Feb 2022 - May 2025

Systems Analyst (Nov 2023 - May 2025)

- Selenium and Telegram automation: -50% manual intervention, ~20 h/month saved; automated validations: -40% ERP rework.
- CI/CD and cloud monitoring on AWS/GCP: -60% deploy time, -35% downtime incidents.
- Automation and monitoring for sugar-bag labeling: 100% label availability, -70% emergency tickets.

Technical Support Analyst (Nov 2022 - Nov 2023): automated diagnostic scripts, -30% troubleshooting time; 95% SLA on internal tickets.

IT Assistant (Feb 2022 - Nov 2022): automated reports with Python and APIs, -70% generation time.

PROJECTS

MCP Fiscal Brasil - open-source (MIT) MCP server with 45 tools on FastMCP, published on PyPI and in the official MCP registry; offline NF-e parsing and CNPJ lookup, including the 2026 alphanumeric format.

AgroVoz - WhatsApp agent that turns voice (Whisper), text, and photos into farm records: intent classification, tool calling, and per-farm data in Postgres; 17 modules, in closed beta.

MCP Jurídico Brasil - MCP server for case tracking over the CNJ's public DataJud API, with 92 court codes mapped.

TranscreveAI - multimodal pipeline that turns video and audio into cited dossiers (transcription, OCR, embeddings, RAG), also exposed as an MCP server.

GerminaVision - seedling germination analysis from photos with YOLO and OpenCV; academic project at IF Goiano (Jan - Jun 2026).

Portfolio with a live agent and detailed case studies at **dehor.dev**.

TECHNICAL SKILLS

AI & agents: LLM APIs (Anthropic, OpenAI, Gemini), MCP, tool calling, RAG, pgvector, embeddings, LLMOps and evaluation (golden sets, cost and latency), Whisper, OCR, YOLO.

Development: Python, TypeScript, FastAPI, React/Next.js.

Data & platform: PostgreSQL/Supabase, Redis, Docker, AWS, CI/CD.

EDUCATION

Information Systems - IF Goiano, Ceres Campus | 2025

AWS re/Start - Campinho Digital + AWS | 2025