

# LOUIS DESCLOUS

AI Engineer
AI Agents · RAG · Data

## CONTACT


- louisdesclouspro@gmail.com
- +33 7 82 67 88 75
- louisdesclous.com
- linkedin.com/in/louisdesclous
- github.com/louisdesc
- Paris · remote or hybrid


## SKILLS

**Programming**
Python · TypeScript · SQL

**AI & LLM**
Agents · Tool calling · RAG · Embeddings · Fine-tuning · LLM evaluation

**Data & ML**
Dataiku · Data pipelines · Machine Learning · PostgreSQL

**Production**
AWS · Terraform · Docker · CI/CD · FastAPI · Sentry

**AI tools**
Claude Code · Codex · Hermes

## LANGUAGES

**French** — native
**English** — TOEIC 980/990

## INTERESTS

Long-distance **running & trail**
**Travel** — 10+ countries visited solo

## PROFILE

AI Engineer with an engineering degree from EFREI Paris (Big Data & AI), coding since 2018. I design and ship AI products end to end: agents with tool calling, RAG, APIs and data pipelines. I measure their quality, latency and cost before I say they work. I founded HireLayer, an AI recruiting API: **2M+ API calls**.

## EXPERIENCE

### Founder & AI Engineer — HireLayer ↗
Oct 2025 – Sep 2026

AI recruiting API: candidate parsing, scoring and voice pre-screening.

**AI voice pre-screening agent**

- Real-time agent that runs the screening interview (questions drawn from the job ad, follow-ups), then writes a per-criterion report with timestamped evidence
- Tool calling with LiveKit Agents, Deepgram, Claude Haiku and ElevenLabs; Python prototype ported to TypeScript
- QA console: interview grids generated from the job ad, simulated candidates, LLM judge, blind interviews
- **5 stacks compared**: 3.95/5, **1.8 s** median latency, **$0.80** per interview; OpenAI Realtime ruled out on measurements
- Tested guardrails: AI and recording disclosure, no illegal questions, manipulation attempts ignored

**API, RAG and MCP**

- **Parsing & matching**: CVs and job ads into structured data (OCR + LLM); per-criterion matching and reranking, NDCG@5, 10+ models compared
- **Hybrid RAG** over 11,000 skills (aliases, Cohere embeddings, Claude validation), evaluated on 1,550 labeled cases: from 10% to 67% correct answers
- **MCP server** (TypeScript): 6 tools for Claude, ChatGPT, Cursor and Codex
- Docker, GitHub Actions CI, Sentry, data isolated per customer

### Freelance AI Engineer — Wink (Station F)
Jul 2024 – Sep 2025

AI CV parsing API built alone for the client, from scoping to production.

- OCR + LLM pipeline → structured JSON; fine-tuning, prompt engineering
- AWS serverless, Terraform; **€50k/year saved** for the client

### AI Engineer / Data Scientist — AlloBrain (Station F)
Sep 2024 – Aug 2025

LLM analysis of customer calls to steer service quality.

- Python, FastAPI, LiteLLM: key-stage detection, API for the dashboards
- Structured JSON outputs, retries, fallback to another model on failure

### Data Engineer / Analyst — Bpifrance
Jul 2023 – Aug 2024

French public investment bank: data, reporting and ML in Dataiku.

- Multi-source Dataiku flows, visual recipes (Prepare, Join), Python and SQL
- Visual ML: classification and prediction models, anomaly detection
- Scheduled scenarios: automated processing and reporting, quality checks
- Rebuilt a Python quality-control tool (GUI, PowerPoint reports)

### Freelance Web & AI Developer — WebCyme
2018 – 2025

- Custom web and AI applications for clients (Next.js, TypeScript)

### Network Director (volunteer) — Cop1
Jun 2024 – Jun 2025

- Student aid non-profit: from 20 to 27 cities, from 3,000 to 5,000 volunteers

### Self-employed — Luxury goods trading
2020 – 2024

- B2C and B2B buying and reselling, **€100k+ revenue**; Python invoicing tools

## EDUCATION & CERTIFICATIONS

**EFREI Paris — Master's in Engineering, Big Data & AI**
Taught entirely in English

**Dataiku — 6 certifications** 2023 – 2026
including Generative AI Practitioner and Agent Builder

**Azure Data Scientist Associate · AWS Cloud Practitioner** 2024 – 2025