

LOUIS DESCLOUS

AI Engineer
Agents IA · RAG · Data

CONTACT

- louisdesclouspro@gmail.com
- +33 7 82 67 88 75
- louisdesclous.com
- linkedin.com/in/louisdesclous
- github.com/louisdesc
- Paris · télétravail ou hybride

COMPÉTENCES

Langages
Python · TypeScript · SQL

IA & LLM
Agents · Tool calling · RAG · Embeddings · Fine-tuning · Évaluation LLM

Data & ML
Dataiku · Pipelines de données · Machine Learning · PostgreSQL

Industrialisation
AWS · Terraform · Docker · CI/CD · FastAPI · Sentry

Outils IA
Claude Code · Codex · Hermes

LANGUES

Français — langue maternelle
Anglais — TOEIC 980/990

CENTRES D'INTÉRÊT

Course & trail longue distance
Voyages — +10 pays visités seul

PROFIL

AI Engineer, ingénieur EFREI (Big Data & IA), développeur depuis 2018. Je conçois et mets en production des produits IA de bout en bout : agents avec tool calling, RAG, API et pipelines de données. Je mesure leur qualité, leur latence et leur coût avant de dire qu'ils marchent. J'ai fondé HireLayer, une API IA de recrutement : **2M+ appels API**.

EXPÉRIENCES

Fondateur & AI Engineer — HireLayer ↗
oct. 2025 – sept. 2026

API IA de recrutement : parsing, scoring et préqualification vocale de candidats.

Agent vocal IA de préqualification

- Agent temps réel qui mène l'entretien de présélection (questions tirées de l'offre, relances), puis rédige une fiche par critère avec preuves horodatées
- Tool calling avec LiveKit Agents, Deepgram, Claude Haiku et ElevenLabs ; prototype Python porté en TypeScript
- Console QA : grilles d'entretien générées depuis l'offre, candidats simulés, juge LLM, entretiens à l'aveugle
- **5 stacks comparées** : 3,95/5, **1,8 s** de latence médiane, **0,80 $** par entretien ; OpenAI realtime écarté sur mesures
- Garde-fous vérifiés : annonce IA et enregistrement, aucune question illégale, tentative de manipulation ignorée

API, RAG et MCP

- **Parsing et matching** : CV et offres (OCR + LLM) en données structurées ; matching par critère et reranking mesurés en NDCG@5, 10+ modèles comparés
- **RAG hybride** sur 11 000 compétences (alias, embeddings Cohere, validation par Claude), évalué sur 1 550 cas annotés : de 10 % à 67 % de bonnes réponses
- **Serveur MCP** (TypeScript) : 6 outils pour Claude, ChatGPT, Cursor et Codex
- Docker, CI GitHub Actions, Sentry, données cloisonnées par client

Ingénieur IA freelance — Wink (Station F)
juil. 2024 – sept. 2025

API IA de parsing de CV livrée seul au client, du cadrage à la production.

- Pipeline OCR + LLM → JSON structuré ; fine-tuning, prompt engineering
- AWS serverless, Terraform ; **50 k€/an économisés** pour le client

Ingénieur IA / Data Scientist — AlloBrain (Station F)
sept. 2024 – août 2025

Analyse des appels clients par LLM pour piloter la qualité de service.

- Python, FastAPI, LiteLLM : détection des phases clés, API pour les dashboards
- Sorties JSON structurées, retries, bascule entre modèles en cas d'échec

Data Engineer / Analyst — Bpifrance
juil. 2023 – août 2024

Banque publique d'investissement : données, reporting et ML sous Dataiku.

- Flows Dataiku multi-sources, recettes visuelles (Prepare, Join), Python et SQL
- Visual ML : modèles de classification et de prédiction, détection d'anomalies
- Scénarios planifiés : traitements et reporting automatisés, contrôles qualité
- Refonte d'un outil Python de contrôle qualité (interface, rapports PowerPoint)

Développeur freelance Web & IA — WebCyme
2018 – 2025

- Applications web et IA sur mesure pour des clients (Next.js, TypeScript)

Directeur du réseau (bénévole) — Cop1
juin 2024 – juin 2025

- Association d'aide aux étudiants : de 20 à 27 villes et de 3 000 à 5 000 bénévoles

Auto-entrepreneur — Commerce de produits de luxe
2020 – 2024

- Achat-revente B2C et B2B, **100 k€+ de CA** ; outils Python de facturation

FORMATION & CERTIFICATIONS

EFREI Paris — Diplôme d'ingénieur (Bac+5), Big Data & IA
Cursus intégralement en anglais

Dataiku — 6 certifications — 2023 – 2026
dont Generative AI Practitioner et Agent Builder

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