Alain Lorenz

CTO & cofondateur — IA & MLOps

Entre Genève et Lyon · al@inlore.nz · ala.inlore.nz · linkedin.com/in/alain-lorenz

Vous êtes humain ? Le CV visuel et le site, c'est par ici : CV visuel · ala.inlore.nz

PROFIL

Cofondateur et CTO, j'ai dirigé des équipes jusqu'à dix personnes sans lâcher le clavier : je suis toujours le premier contributeur de l'API que j'ai lancée en 2016. Depuis 2023, mon terrain principal, c'est l'IA en production — entraîner, déployer, servir et optimiser des modèles qui répondent à de vrais utilisateurs. Mes sujets de prédilection : les LLM, le RAG et le fine-tuning. Et la paperasse qui fait tourner la boîte — RGPD, DAC7, fiscalité, crédit d'impôt innovation — passe aussi par moi. Le tout entre Genève et Lyon.

EXPÉRIENCE

Cofondateur technique · CTO — JobyPepper

depuis 2016 · jobypepper.com

La plateforme qui met en relation des entreprises et des indépendants pour des missions.

- Web, mobile, API et infrastructure : tout l'écosystème produit.
- Deux apps mobiles publiées sur les stores iOS et Android, de la build à la fiche : versions, soumission, mises en conformité, rédactionnel.
- Dispatch financier, pas un simple bouton « payer » : le client paie une fois, la plateforme génère la facture de chaque prestataire, répartit les fonds et déclenche chaque versement. Automatiquement.
- Agent IA de sourcing et de matching de candidats, branché via MCP.

Cofondateur technique · CTO — Ozly

2023 · en pause · hello.ozly.ai

L'IA appliquée à la recherche d'emploi.

- Pipeline complet : ingestion multi-sources, classification CamemBERT, index vectoriel.
- Recherche sémantique et RAG conversationnel, dans un widget intégrable.
- Inférence GPU optimisée : coût divisé par quatre.

MLOPS

- **Collecter :** Pipelines planifiés qui agrègent, nettoient et dédupliquent des données venues de plusieurs sources.
- **Entraîner :** Construction de datasets, fine-tuning de classifieurs par secteur et par type de contrat.
- **Packager :** Modèles versionnés, images Docker CUDA, publication dans un registre de conteneurs.
- **Servir :** Inférence GPU, service d'embeddings sur Kubernetes, recherche vectorielle.
- **Exploiter :** CI/CD, observabilité sur tous les services, optimisation continue : embeddings réduits à 512 dimensions, fuite mémoire traquée et corrigée.

PILOTER

- Direction technique d'une équipe jusqu'à dix personnes.
- Décisions d'architecture, organisation de la livraison, arbitrages produit / technique.
- Montage des dossiers de financement de l'innovation.

- Transformer un besoin opérationnel en outil, et une expérimentation IA en service exploitable.

FAIRE TOURNER LA BOÎTE

- Conformité RGPD et déclarations associées.
- Export et déclarations DAC7, l'obligation fiscale des plateformes collaboratives.
- Impôts et obligations déclaratives de l'entreprise.
- Suivi des projets éligibles au crédit d'impôt innovation (CII).

PROJETS PERSONNELS

Projet personnel — Origin AI

2026

Des vidéos de vrais postes de travail transformées en données d'entraînement pour la robotique.

- Upload vidéo fragmenté, traitement ffmpeg, tagging par LLM.
- Floutage automatique des visages sur GPU, pipeline de création de datasets.

Side project — EV Comparator

2026

À partir de combien de kilomètres l'électrique devient-il moins cher ?

- Ingestion quotidienne d'open data, modèle de coût au kilomètre.
- Estimation de revente présentée comme telle, jamais comme une vérité.

COMPÉTENCES

IA & MLOps : LLM, RAG, Embeddings et recherche vectorielle, Fine-tuning (CamemBERT), Triton Inference Server, CUDA, Serveurs MCP, LangChain

Déploiement : Docker (images CUDA, Compose GPU), Kubernetes, CI/CD (CircleCI), Nginx, Sentry

Back-end : Node.js / Express, Python, RabbitMQ, MongoDB, Redis, WebSocket

Front & mobile : React, React Native / Expo, Publication iOS et Android

Gestion : Direction technique (jusqu'à 10 personnes), RGPD, DAC7, Fiscalité, CII

LANGUES

Français, Allemand, Anglais