

Manel LABIDI

AI Engineer & Researcher

+33 6 95 82 27 99

manel_labidi@hotmail.fr

linkedin.com/in/manel-labidi

orcid.org/0009-0000-9386-506X

Annecy 74000, France

Nationalité française

Mobile sur la Suisse romande / sur site

Profil

Ingénieure IA et chercheuse avec plus de 5 ans d'expérience dans la conception et le développement de solutions d'intelligence artificielle dans plusieurs secteurs d'activité. Expertise en IA générative, systèmes multi-agents, RAG, NLP, vision par ordinateur et machine learning. Intervention sur les phases d'avant-vente, chiffrage, architecture technique et développement de solutions IA utilisées en production. Auteure de publications scientifiques et formatrice sur les technologies IA.

Compétences

Analyse et modélisation de données

Transformation de données métier hétérogènes en solutions IA exploitables, du cadrage du problème à la livraison de bout en bout (Python, pandas, NumPy, SQL).

IA générative

Développement d'applications exploitant des LLM propriétaires et open source, avec sorties structurées et adaptation de modèles (OpenAI, Claude, Gemini, Mistral, Llama, Ollama, Pydantic, LoRA, QLoRA).

IA agentique

Conception de systèmes mono-agent et multi-agents, automatisation de workflows métier et intégration d'outils externes type ERP (LangGraph, LangChain, CrewAI, OpenAI Agents SDK, MCP).

RAG & recherche d'information

Développement de moteurs de recherche documentaire et de systèmes RAG produisant des réponses sourcées et vérifiables (recherche sémantique, hybride, reranking, Text-to-SQL ; Elasticsearch, Qdrant, FAISS, Cohere Rerank, Sentence Transformers).

Vision par ordinateur

Détection, segmentation, reconnaissance visuelle et OCR, avec modèles vision-langage pour structurer l'information (YOLOv9, Mask R-CNN, Cascade Mask R-CNN, Mask2Former, U-Net, Detectron2, MMDetection, Tesseract OCR).

Machine learning & deep learning

Développement de modèles de classification, régression et prévision de séries temporelles, optimisation d'hyperparamètres et évaluation comparative de modèles (scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, LSTM, Transformers, Optuna).

Traitement documentaire

Développement de pipelines d'extraction, de structuration et d'enrichissement de données documentaires afin d'alimenter bases de connaissances, moteurs de recherche et systèmes IA (Docling, Tesseract OCR, pypdf, pdfminer, pdf2image, Textract, openpyxl).

Développement & industrialisation

Développement de services IA en APIs REST et microservices, conteneurisés et intégrés aux plateformes clients (cloud et on-premise), avec test du déploiement en local (FastAPI, Flask, Docker, Kubernetes / Minikube, GitLab, Pytest).

Veille technologique & workshops

Veille technologique et intégration des innovations IA dans les projets ; conception et animation de formations techniques sur les LLM, les systèmes RAG et les agents IA ; accompagnement des équipes dans l'adoption des technologies d'intelligence artificielle dans leurs métiers.

Expérience

Ingénieure IA & Consultante delivery — Leviatan

2021 – Présent

- Conception et développement de solutions IA (GenAI, RAG, agents IA, vision par ordinateur) pour les secteurs industrie, immobilier et BTP.
- Participation aux avant-ventes, chiffrages, choix d'architecture et développement backend.
- Développement de services IA sous forme d'APIs REST et de microservices intégrés aux plateformes clients.
- Animation de formations et accompagnement des équipes sur les technologies LLM, RAG et agents IA.

Recherche quantitative & modélisation de données — Universités Lyon 3, Nice Sophia Antipolis & Grenoble

2007 – 2021

- Recherche en statistiques, économétrie et modélisation de séries temporelles appliquée aux données financières.
- Expertise transférée vers l'intelligence artificielle et prolongée par des travaux de recherche publiés en IA.

Publications scientifiques

Predicting Stock Performance: A Comparative Study of Econometric and AI-Based Models.

2025

M. Labidi, Y. Zhang, M. Petit Guillaume, A. Krauth. CNIA@PFIA 2025. <https://hal.science/hal-05168124>

CryptoGPT: A 7B Model Rivaling GPT-4 for Financial News Analysis and Classification.

2024

Y. Zhang, M. Petit Guillaume, A. Krauth, M. Labidi. JNFT 2024. <https://hal.science/hal-04617152v1>

Projets clients significatifs

- **Système RAG multi-agents (industrie, BTP, immobilier)** : développement d'un système permettant à des conseillers métier d'interroger plus de 1 500 documents issus du corpus client (PDF natifs et scannés). Recherche vectorielle, reranking et LLM hébergés en Europe pour répondre à des questions client. Solution déployée en production chez 4 clients.
- **Analyse intelligente de rapports PDF (sécurité des bâtiments)** : automatisation de l'extraction et de l'intégration en base de données d'informations issues de centaines de rapports d'analyse technique de sécurité des bâtiments et de rapports de commission (PDF natifs et scannés : observations, dates, rapporteurs, tableaux, etc.) avec une qualité d'extraction des informations d'environ 92 %. Réduction du temps de traitement de plusieurs heures à moins de 60 secondes par document.
- **Module d'extraction documentaire multi-formats** : développement d'un pipeline d'extraction et de normalisation de données à partir de PDF, Word et PowerPoint afin d'alimenter automatiquement des bases de connaissances et moteurs de recherche documentaires.
- **Comparaison automatisée de fichiers Excel volumineux** : développement d'une solution automatisant la comparaison de fichiers contenant plusieurs milliers de lignes et générant des écarts directement exploitables par les utilisateurs métier ; mise en production.
- **Parsing automatique de CV (immobilier)** : développement d'un système d'extraction automatique de données (nom, prénom, adresse, téléphone, e-mail, âge) à partir de CV PDF, Word ou image, utilisé en production pour le préremplissage automatique des dossiers candidats, avec un taux d'extraction supérieur à 96 %.
- **Matching candidats-offres (recrutement immobilier)** : développement d'un moteur de recommandation identifiant automatiquement les profils les plus pertinents en quelques secondes (contre plusieurs minutes de recherche manuelle auparavant) et les classant selon leur adéquation et leur proximité géographique. Approche initiale par similarité sémantique entre CV et historique de recrutement, puis remplacée par une classification des candidats une fois les données suffisantes disponibles. Solution utilisée en production.
- **Chatbot métier (immobilier)** : développement d'un assistant conversationnel, utilisé en production par plusieurs milliers d'utilisateurs externes pour interroger une base documentaire et accéder rapidement aux informations métier pertinentes.
- **Chatbot multilingue (FR, EN, ES, NL) — parentalité** : développement d'un assistant conversationnel alimenté par recherche documentaire permettant de répondre à des questions liées à la grossesse et à la parentalité à partir de contenus validés par des experts. Solution utilisée en production par une base de plusieurs centaines de milliers d'utilisateurs.
- **Vision — segmentation d'instances sur plans architecturaux 2D** : développement d'un modèle de segmentation d'instances entraîné sur plus de 4 000 plans annotés (IoU supérieur à 80 % après post-traitement), automatisant la détection des polygones de murs auparavant tracés à l'aide d'un logiciel et intégré aux processus opérationnels du client.
- **Extraction d'informations de plaques signalétiques industrielles (Vision + LLM)** : développement d'une solution de lecture et de structuration des données issues de plaques produits (numéro de série, marque, gamme, produit), restituées via API REST et intégrées aux outils métier du client. Traitement complet en moins de 20 secondes, réduisant fortement les opérations de saisie manuelle et le temps de traitement des dossiers.
- **Vision — reconnaissance de produits par l'image** : développement d'un système combinant détection open-vocabulary et recherche par similarité visuelle sur un catalogue d'environ 20 000 références produits multi-marques. Solution en cours de déploiement en production, avec un taux de reconnaissance d'environ 90 %.
- **Génération automatisée d'attributs de prénoms** : enrichissement automatique de données à grande échelle à partir de règles métier, données internes et LLM afin d'alimenter un service d'aide au choix de prénoms.

- **Prédiction de tendances crypto** : développement et comparaison de modèles de machine learning et de deep learning pour la prévision de séries temporelles financières, avec automatisation des traitements, exposition via API REST et validation par backtesting.

Formation

Certification professionnelle en Data Science , Campus Numérique des Alpes	2020–2021
Data Science Professional Certificate , HarvardX (edX)	2019–2020
Doctorat en Finance de marché , Université Grenoble Alpes	2011–2015

Langues

- **Français** — langue maternelle
- **Anglais** — B2
- **Italien** — débutant

Qualités

Leadership technique · Rigueur scientifique · Pédagogie · Communication · Résolution de problèmes

Centres d'intérêt

Ski · Randonnée · Natation