

☎ +33 6 95 82 27 99

✉ contact@manel-labidi.com

in linkedin.com/in/manel-labidi

🔗 orcid.org/0009-0000-9386-506X

📍 Annecy 74000, France

Nationalité française

Sur site : Annecy, Haute-Savoie & bassin genevois — télétravail au-delà

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. Expertise en IA générative, systèmes multi-agents, RAG, NLP, vision par ordinateur et machine learning. Impliquée dans l'avant-vente, le cadrage, l'architecture technique et le développement de solutions IA utilisées en production. Autrice de publications scientifiques et formatrice sur les technologies d'IA.

Compétences

Analyse et modélisation de données

Transformer des données métier hétérogènes en solutions IA exploitables, du cadrage du problème jusqu'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- et multi-agents, automatisation de workflows métier et intégration d'outils externes tels que les 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 et hybride, reranking, Text-to-SQL ; Elasticsearch, Qdrant, FAISS, Cohere Rerank, Sentence Transformers).

Vision par ordinateur

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

Machine & 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, structuration et enrichissement de données documentaires pour alimenter bases de connaissances, moteurs de recherche et systèmes IA (Docling, Tesseract OCR, pypdf, pdfminer, pdf2image, Textract, openpyxl).

Développement & mise en production

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

Veille & formation

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

Expérience

Ingénieure IA & Consultante delivery — Leviatan

2021 – Présent

- Conception et développement de solutions IA (GenAI, RAG, agents IA, vision) pour l'industrie, l'immobilier et le BTP.
- Avant-ventes, cadrage, choix d'architecture et développement backend.
- Services IA exposés en APIs REST et microservices intégrés aux plateformes clients.
- 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'IA et prolongée par des travaux de recherche publiés.

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 sélectionnés

- **Système RAG multi-agents (industrie, BTP, immobilier)** : développement d'un système permettant à des conseillers métier d'interroger plus de 1 500 documents du corpus client (PDF natifs et scannés). Recherche vectorielle, reranking et LLM hébergés en Europe pour répondre aux questions clients. 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 d'informations issues de centaines de rapports techniques de sécurité et de rapports de commission (PDF natifs et scannés : constats, dates, rédacteurs, tableaux, etc.), avec une qualité d'extraction d'environ 92 %. Temps de traitement réduit 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 depuis PDF, Word et PowerPoint pour alimenter automatiquement bases de connaissances et moteurs de recherche documentaire.
- **Comparaison automatisée de fichiers Excel volumineux** : développement d'une solution automatisant la comparaison de fichiers de plusieurs milliers de lignes et générant des différences directement exploitables par les équipes métier ; déployée en production.
- **Parsing automatique de CV (immobilier)** : développement d'un système d'extraction automatique de données (nom, prénom, adresse, téléphone, email, âge) depuis des CV PDF, Word ou image, utilisé en production pour pré-remplir les fiches 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) et les classant par adéquation et proximité géographique. Approche initiale par similarité sémantique entre CV et historique de recrutement, puis remplacée par une classification une fois suffisamment de données 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 à l'information métier.
- **Chatbot multilingue (FR, EN, ES, NL) — parentalité** : développement d'un assistant conversationnel appuyé sur la recherche documentaire pour répondre aux questions de grossesse et de 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 la main et intégré aux processus opérationnels du client.
- **Extraction d'informations sur plaques industrielles (Vision + LLM)** : développement d'une solution pour lire et structurer les données de plaques signalétiques produit (numéro de série, marque, gamme, produit), renvoyées via API REST et intégrées aux outils métier du client. Traitement complet en moins de 20 secondes, réduisant fortement la 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 produit multi-marques. Solution en cours de déploiement, 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, de données internes et de LLM pour alimenter un service de suggestion de prénoms.
- **Prédiction de tendances crypto** : développement et comparaison de modèles de machine learning et deep learning pour la prévision de séries temporelles financières, avec traitement automatisé, 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 , Université Grenoble Alpes	2011–2015

Langues

- **Français** — natif
- **Anglais** — B2
- **Italien** — débutant

Atouts

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

Centres d'intérêt

Ski · Randonnée · Natation