

Benny LAY

Data Scientist / Ingénieur IA

☎ 06.16.74.19.86 | ✉ benny.lay.contact@gmail.com | 📍 Ile-de-France, France |  [linkedin.com/in/benny-lay](https://www.linkedin.com/in/benny-lay) |  github.com/BennyBobby |  bennybobby.github.io

EXPÉRIENCE

Aubay

Fév. 2026 – Présent

Stagiaire R&D IA

Boulogne-Billancourt, France

- **Détection automatisée du blanchissement corallien** par Deep Learning et Vision par Ordinateur, en exploitant des modèles tels que **SAM** et **DINOv2** pour la segmentation et la représentation visuelle
- Construction d'un dataset annoté sur mesure et développement d'algorithmes de quantification du blanchissement des coraux via **similarité colorimétrique (distance euclidienne RGB)** et modèle par régression
- Gestion du cycle de vie du projet en **méthodologie Agile**

Ulster University

Juil. 2024 – Nov. 2024

Stagiaire Analyste de Données

Londonderry, Royaume-Uni

- Analyse de données complexes pour soutenir une recherche académique axée sur **l'amélioration des performances étudiantes**
- Développement d'une interface utilisateur pour faciliter l'accès aux données environnementales du **projet national Land Use for Net Zero (LUNZ Hub)**
- Conception et **automatisation d'un système d'affectation de projets étudiants**, réduisant significativement la charge administrative manuelle

PROJETS

Application Deep Learning de Prédiction d'Âge | *PyTorch, FastAPI, React, Docker, Optuna*

- Entraînement d'un modèle de régression **ConvNeXt-Tiny** sur le dataset UTKFace atteignant une **MAE de 4,49 ans** ; optimisation des hyperparamètres avec **Optuna** et conteneurisation avec support GPU via **Docker**
- Développement d'un backend **FastAPI** avec détection de visage OpenCV et frontend **React**
- Fiabilité du code assurée par 17 cas de test **pytest** et pipeline **CI/CD** automatisé via GitHub Actions

SotaRAG — Assistant de Recherche RAG | *Ollama, Qdrant, FastAPI, Redis, Streamlit*

- Développement d'un système **RAG** pour interroger des articles ArXiv en langage naturel, avec **Qdrant** pour la recherche sémantique et **Ollama** pour l'inférence LLM locale
- Architecture d'un pipeline d'ingestion asynchrone (**TaskIQ + Redis**) pour le traitement PDF non-bloquant et orchestration des microservices avec **Docker Compose**

Chess Analytics Dashboard | *Python, Streamlit, SQLite, Docker, GitHub Actions*

- Développement d'une application web d'analyse de parties d'échecs : pipeline **ETL** intégrant l'API Chess.com avec ingestion incrémentale et déduplication vers **SQLite**
- Dashboard **Streamlit** interactif (évolution du classement, analyse des ouvertures, taux de victoire par mode de jeu) ; conteneurisé avec **Docker Compose** et CI/CD via GitHub Actions

FORMATION

ESEO

2021 – 2026

Diplôme d'Ingénieur

Vélizy-Villacoublay, France

- **Spécialisation** : Data Science et Intelligence Artificielle

COMPÉTENCES

Programmation : Python, SQL

IA & Machine Learning : PyTorch, TensorFlow, Scikit-Learn, Pandas

Backend & DevOps : FastAPI, Docker, Git, GitHub Actions (CI/CD)

Visualisation de Données : Power BI, Streamlit

Langues : Français (Natif), Anglais (B2)