



BEN REJAB SAHAR

Étudiante ingénierie en Data Science Recherche appliquée | Stage RD Data Exploration
(février–mars 2026)

📍 France 📩 sahar.benrejab@centrale.centralelille.fr

📞 07 44 24 51 01

LinkedIn

GitHub

Profil professionnel

Ingénierie en devenir, passionnée par la recherche en Data Science appliquée à la compréhension du comportement humain et à la santé visuelle. Je combine rigueur scientifique, exploration de données multimodales et curiosité analytique pour transformer les données complexes en connaissances exploitable. Intéressée par les projets RD liant vision, perception et innovation biomédicale.

Projets IA

Data Exploration – Human Behavior Analysis

2025

Projet personnel — Architecture multi-agents avec LangChain

- Analyse exploratoire de données multimodales décrivant des comportements humains et visuels.
- Création d'indicateurs descriptifs et de visualisations synthétiques pour appuyer une modélisation perceptuelle..
- Technologies : Python, Pandas, Matplotlib, Seaborn, Scikit-learn..

RAG LangChain Assistant – Recherche augmentée par génération

2025

Projet personnel — Assistant conversationnel contextuel

- Développement et optimisation : conception d'un assistant conversationnel exploitant FAISS et RecursiveCharacterTextSplitter pour la recherche sémantique et la génération contextuelle.
- Résultat : amélioration notable de la pertinence des réponses par rapport aux approches classiques.
- Technologies : LangChain, FAISS, OpenAI API, Python.

Expériences professionnelles

Ingénierie IA – Projet industriel (CDD)

sept.–nov. 2025

McBride

- Pipeline prédictif et exploration moléculaire : conception d'un pipeline ML pour optimiser la performance de formules détergentes, intégrant des embeddings moléculaires et des Graph Neural Networks en appliquant une démarche exploratoire de corrélation entre propriétés physico-chimiques et performances expérimentales.
- Extraction de données PDF : mise en place d'un pipeline Docing pour extraire tableaux/textes de rapports PDF (tests & formulations), normaliser les schémas et produire des jeux de données CSV/Parquet prêts pour l'analyse et l'entraînement.
- Technologies : Python, scikit-learn, DVC, Azure.

Stagiaire Data Science / IA

mars–août 2025

Regmant (Thermopatch)

- Prospection et automatisation : développement d'un moteur multi-source (LinkedIn, Google Maps, Booking) connecté au CRM Odoo avec un module de scoring IA pour hiérarchiser les leads.
- Technologies : Python, Pandas, Selenium, Odoo API.

Stagiaire Développement Full-Stack

juin–août 2024

Archimomi

- Développement web : création d'une plateforme d'éco-tourisme responsive avec chatbot intégré pour assistance utilisateur.
- technologies : Node.js, MongoDB, HTML/CSS/JS.

Formation

Master 2 — Intelligence Artificielle & Machine Learning

2025–2026 | Lille, France

École Centrale Lille

Cours clés : Deep Learning, Reinforcement Learning, Computer Vision, NLP, Multi-Agents, Generative AI, MLOps.

Diplôme d'Ingénieur — Data Science

2023–2026 | Casablanca, Maroc

CentraleSupélec — Campus Casablanca

Cours clés : Machine Learning, Data Engineering, Optimization, Graph Theory, Predictive Modeling.

Classes Préparatoires — Maths & Physique

2021–2023 | Nabeul, Tunisie

Institut Préparatoire IPEIN

Compétences techniques et linguistiques

Data Science & Exploration : Python (Pandas, NumPy), Scikit-learn, Matplotlib, Seaborn, Data Visualization, Statistical Analysis, Metadata Structuring, Data Cleaning , Python, SQL, C++, JavaScript.

IA & Dév : PyTorch, TensorFlow, Hugging Face, LangChain, FAISS, MLflow.

Outils : Power BI, Excel (avancé), Odoo CRM, PhantomBuster, Docing (extraction PDF).

Langues : Français (C1 DALF) | Anglais (C1 TOEIC, Linguaskill) | Arabe (natif).