



Antoine THÉOLOGIE

Data Scientist & MLOps Engineer · Recherche CDI

theologia@gmail.com · Disponible immédiatement

linkedin.com/in/antoine-theologien · github.com/antothl

Profil : Data Scientist & MLOps Engineer

Diplômé du Master MIND (Machine Learning, IA et Données, Sorbonne Université, mention Bien), je combine deux compétences : la **data science**, pour comprendre un besoin, explorer les données et construire des modèles, et le **MLOps**, pour les mettre en production, les surveiller et garantir leur fiabilité dans la durée. Mon objectif : transformer des prototypes en produits réellement utilisés.

Expérience Professionnelle

ML Engineer (Stage de fin d'études)

Février 2026 – Août 2026

SNCF Connect & Tech

Saint-Denis

- Exploration des données de production d'un modèle de **recommandation de trajets** avec PySpark et premiers tests de détection de dérive.
- Conception et développement d'un outil de **détection de dérive des données et des modèles** (Evidently) : comparaison statistique des distributions d'entraînement et de production (distance de Wasserstein, test de Kolmogorov-Smirnov).
- **Industrialisation** : exécution automatisée orchestrée par **Airflow** et alerting via **Datadog** en cas de dérive ; outil depuis **réutilisé sur plusieurs modèles** en production.
- Rédaction de l'ensemble de la **documentation technique** et présentation du projet à l'équipe.
- Maintenance de l'environnement : mises à jour d'images **Docker** et tests de nouvelles versions de librairies.

Développeur Web Full-Stack (2 stages de 8 semaines)

2023 – 2024

Groupe Europe Handling

Roissy-en-France

- Deux applications internes conçues avec les équipes métier (PHP Symfony, JavaScript, SQL).

Projets Data & IA

Assistant de questions-réponses sur 20 000 tableaux de données (LLM)

2025

Sorbonne Université

Paris

- Outil qui répond en langage naturel à des questions sur des tableaux Wikipédia : nettoyage automatique de 20 000 tables hétérogènes par LLM (Mistral-7B, CodeLlama-13B), chargement en base SQL, génération des requêtes (Text-to-SQL) et évaluation automatisée des réponses.

Réimplémentation d'un article de recherche en deep learning

2025

Sorbonne Université

Paris

- Implémentation en **PyTorch**, à partir de l'article, d'un modèle de séquence à **complexité linéaire** (Mixture-of-Memories) : contrairement aux Transformers, dont le coût explose avec la longueur du texte, il traite de longues séquences à coût maîtrisé.
- Plusieurs mémoires en parallèle pour retenir plus d'informations ; **résultats encourageants**.

Formation

Master Informatique MIND (ML, IA et Données) – Mention Bien

Septembre 2024 – 2026

Sorbonne Université

Paris

Licence Informatique (Cursus Master Ingénierie)

Septembre 2021 – Juin 2024

Université de Reims-Champagne-Ardenne

Reims

Compétences & Informations

- **Langages** : Python, SQL, Java
- **Data Science & IA** : PyTorch, Scikit-learn, statistiques, LLM, RAG, fine-tuning (LoRA), NLP
- **Données** : PostgreSQL, Oracle, Spark (PySpark), Airflow
- **MLOps & Cloud** : Evidently, Datadog, MLflow, Docker, Kubernetes, AWS, GitLab CI/CD ; **Outils** : Git, Linux, Bash, assistants de code IA (skill personnalisé)
- **Langues** : Anglais (B2 certifié – First Certificate of Cambridge), Espagnol (notions)
- **Centres d'intérêt** : tennis (pratiqué 5 ans), football, cinéma, musique · **Permis B**