



Sid TALHA Docteur en Intelligence Artificielle

Data Scientist Senior
NLP & IA Générative
Paris, Île-de-France, France

CONTACT INFORMATION

EMAIL: sidotalha@gmail.com LINKEDIN: [Sid Talha](#)
PHONE: +33 6 65 63 98 10 GITHUB: [Sid Talha](#)

RÉSUMÉ

Data Scientist senior avec plus de 5 ans d'expérience dans la conception et l'industrialisation de solutions d'IA from scratch, en particulier dans le Traitement Automatique du Langage (NLP) et les modèles génératifs (Transformers, LLM, embeddings). Docteur en Intelligence Artificielle, j'ai mené à bien plusieurs projets industriels sensibles, de la phase exploratoire jusqu'au déploiement en production, notamment dans les secteurs bancaire et retail. Passionné par l'innovation appliquée, j'excelle à transformer des besoins métiers complexes en solutions IA opérationnelles et fiables. Mon expérience reconnue en communication et vulgarisation technique me permet de collaborer efficacement avec des équipes pluridisciplinaires pour développer des applications stratégiques, telles que des assistants IA nouvelle génération dans des environnements exigeants et sécurisés.

EXPÉRIENCE PROFESSIONNELLE

- JANV. 2023
– AUJOURD'HUI | **Data Science Consultant – IA & NLP Génératif | SCIAM (mission Société Générale)**
Conception et déploiement de solutions d'IA from scratch pour l'automatisation de processus critiques (Successions, FTI).
- **NLP avancé** : Développement de modèles basés sur les Transformers et embeddings pour l'analyse sémantique d'emails, l'extraction d'informations clés et la détection de priorités.
 - **Automatisation intelligente** : Mise en production de systèmes d'orchestration d'emails et de workflows, réduisant significativement la charge manuelle.
- Technologies : Python, Transformers, LLM, Embeddings, Scikit-Learn, Automatisation, MLOps.
- MARS 2021
– AUJOURD'HUI | **Data Scientist Consultant | SCIAM, Paris, France**
Pilotage de projets d'IA pour des clients industriels et financiers, de la preuve de concept au déploiement.
- Conception de solutions scalables en NLP et Machine Learning, adaptées aux besoins métiers complexes et sensibles.
 - Conseil stratégique et accompagnement sur la mise en place d'architectures IA industrielles.
- Technologies : Python, NLP, Deep Learning, Stratégie IA, Cloud, Industrialisation.
- JUIL. 2021
– DÉC. 2022 | **Data Scientist – Migration & Analytics | Catalina Marketing France, Paris, France**
Refonte des processus analytiques et migration vers des solutions IA modernes.
- Développement d'un framework de migration SAS -> Python, garantissant robustesse, traçabilité et performances accrues.
 - Formation et mentorat des équipes sur Python, Machine Learning et bonnes pratiques de data engineering.
- Technologies : Python, Machine Learning, SQL, Migration de systèmes, Optimisation de workflows.
- FÉV. 2020
– MARS 2021 | **Enseignant-Chercheur | IMT Nord Europe, Lille, France**
Activités de recherche et d'enseignement en Intelligence Artificielle et Data Science.
- Encadrement de cours et projets étudiants en Machine Learning, Deep Learning et Python.
- Technologies : Python, PyTorch, TensorFlow, Scikit-Learn, MATLAB.

FORMATION

- 2016 – 2020 | **Doctorat (Ph.D.) en Intelligence Artificielle & Machine Learning | IMT Nord Europe, Lille**
Thèse : "Modélisation et reconnaissance de comportements complexes via apprentissage automatique et capteurs intelligents".
Conception de systèmes d'analyse temps réel et de modèles prédictifs appliqués à des données complexes et multivariées.
- 2014 – 2016 | **Master en Électronique et Automatique, option Data & Intelligence Artificielle | Université Paul Sabatier, Toulouse**
Spécialisation : Machine Learning, traitement de données complexes et applications en IA

COMPÉTENCES TECHNIQUES

IA GÉNÉRATIVE & NLP:	Fine-tuning et déploiement de LLMs, Hugging Face Transformers, embeddings, classification et extraction d'information, recherche sémantique, clustering multilingue.
MACHINE LEARNING:	Modèles supervisés et non supervisés, optimisation et évaluation de modèles, feature engineering, industrialisation et MLOps (CI/CD, monitoring).
DEEP LEARNING:	Réseaux de neurones (CNN, RNN, Transformers), architectures modernes pour NLP et vision.
LANGAGES & FRAMEWORKS:	Python (expert), PyTorch, TensorFlow, Scikit-Learn, Pandas, SQL.

PROJETS SIGNIFICATIFS

2022 - 2025	Clubini – Plateforme Communautaire Étudiante Conception et développement d'une application web complète (Angular + Spring Boot) permettant aux étudiants de s'inscrire, de se connecter et d'être regroupés automatiquement selon leur établissement. Mise en place d'un système d'authentification sécurisé, d'une base de données relationnelle pour la gestion des utilisateurs et d'une logique de regroupement par école afin de favoriser l'échange et la collaboration au sein des communautés étudiantes.
2025	Sidjo – Plateforme d'Actualités et d'API en Intelligence Artificielle Développement et déploiement d'une plateforme web innovante combinant Angular, Node.js et Azure Web Apps , permettant de centraliser des actualités et des services d'API liés à l'IA. Mise en place d'un backend robuste pour gérer les contenus, d'une interface utilisateur performante et sécurisée, et optimisation du déploiement cloud pour assurer la fiabilité et la scalabilité de la plateforme.

RÉCOMPENSES ET PUBLICATIONS

JUIN 2019	Prix d'Excellence en Communication <i>Collège Doctoral des Hauts de France</i> Récompense pour la clarté, l'impact et la qualité de la présentation d'un projet de recherche, démontrant une forte capacité de vulgarisation scientifique.
	PUBLICATIONS CLÉS: <ul style="list-style-type: none">Hierarchical classification scheme for real-time recognition of physical activities... (EMBC 2019)Features and Classification Schemes for View-Invariant and Real-Time... (IEEE TCDS 2018)

LANGUES

FRANÇAIS:	Langue maternelle
ANGLAIS:	Niveau professionnel