

Martin LE CORRE

Mastère spécialisé Big Data & IA — Télécom Paris, France

Ingénieur généraliste en mastère spécialisé Big Data & IA à Télécom Paris, je me concentre sur les systèmes RAG et les pipelines ELK. Projets personnels en cours ou achevés, avec démonstrateurs.

Recherche un poste d'ingénieur des données ou de scientifique des données à partir de janvier 2026, dans le monde entier.

Contact

martinlecorre@live.fr
+33 6 72 92 51 88
28 ans
Permis de conduire
international
[linkedin.com/in/martinlecorre/](https://www.linkedin.com/in/martinlecorre/)

Langues

Français : langue maternelle
Anglais : TOEIC 840/990
Espagnol : pratique occasionnelle
Portugais et chinois : notions

Compétences

Langages

Python, Java, R, SQL, PHP,
HTML/CSS

Science des données et IA

Apprentissage automatique et profond, vision par ordinateur, traitement du langage, modèles de type Transformer

Écosystème Python

Pandas, NumPy, Scikit-learn,
Seaborn, Selenium, Django

Bases de données

PostgreSQL, MySQL, Oracle,
SQL Server

Données massives

Spark, Hadoop MapReduce,
Hive, Elasticsearch

DevOps

Docker, Kubernetes, Linux

Outils et méthodes

Power BI, Git, Trello, Flutter,
Laravel, Symfony ; méthodes
agiles et gestion de projet

Ingénierie

Simulation : ANSYS, COMSOL,
MATLAB
Conception assistée par
ordinateur : Catia V5,
Solidworks

Projets personnels et principaux projets

Système RAG – Traitement de données confidentielles

2025 • Projet personnel • Démonstrateur prêt • Utilisation des données soumise au RGPD

Système avancé de traitement de données confidentielles ou locales, déployé sur site avec Ollama ou intégré à plusieurs fournisseurs de modèles de langage (Claude, GPT...). Comparaison des performances en temps réel et analyse des coûts.

- Recherche sémantique hybride : représentations vectorielles denses avec PostgreSQL/PGVector et HNSW, recherche lexicale BM25 et fusion RRF, avec reclassement par encodeur croisé.
- Ingestion automatisée multiformat : PDF, DOCX, TXT, Markdown, py, ipynb, jpg, mp3... Découpage intelligent, filtrage qualité et représentations vectorielles accélérées par GPU (BAAI/bge-m3).
- Fonctionnement entièrement local pour les données sensibles, ou hybride. Démonstrateur prêt et système déployable ; dépôt disponible sur demande.

Technologies : LangChain, Ollama et API cloud, PostgreSQL/PGVector, recherche vectorielle/BM25/RRF, traitement du langage, Sentence-Transformers, PyTorch/CUDA, HNSW, CrossEncoder, Whisper, Streamlit, Docker.

Pipeline ETL – Analyse de données massives

2025 • Projet personnel • Démonstrateur prêt • Cours personnels et ingestion multiformat

Plateforme automatisée de collecte et d'analyse : pipeline ETL, collecte web intelligente en plusieurs passes, prédiction du genre par apprentissage automatique et tableau de bord en temps réel avec Elastic et Kibana.

Infrastructure conteneurisée, déployable localement ou sur Hetzner/Oracle Cloud, avec scripts de déploiement automatisés.

Technologies : Airflow (orchestration et planification des DAG), Docker, Spark/PySpark, Parquet colonnaire et Snappy, Selenium, PostgreSQL, Flask, apprentissage automatique, Pandas, Jupyter, Redis, CI/CD, SQL avancé et optimisation des requêtes, contrôle qualité, Elastic, Kibana, cloud.

Airbus Defence and Space — Changements environnementaux

2024/2025 • Projet académique confidentiel • Images satellites

Détection automatisée de changements dans les séries temporelles d'images satellites. Comparaison de SAM2, DinoV2, Clay et Anysat pour la détection non supervisée d'anomalies et leur priorisation.

- Prétraitement et exploration du jeu de données Dynamic Earth Net (Planet Labs).
- Implémentation et comparaison des modèles fondamentaux ; segmentation sémantique et extraction de représentations vectorielles.
- Analyse temporelle, apprentissage non supervisé, regroupement LOF et détection d'anomalies avec PyTorch sur des images multispectrales.
- Évaluation par courbes ROC : les modèles spécialisés en imagerie satellite surpassent les modèles généralistes pour cette tâche.

Formation et expériences

Associations

Bureau des élèves

- 2018 : organisation d'un séjour au ski pour 600 étudiants à Vars et Risoul.
- 2017 : organisation d'événements et de soirées.
- 2017 : participation à l'organisation d'activités nautiques.

Kitesurf

2022 : membre du conseil d'administration de la FKSO, association de kitesurf dans les Landes.

Centres d'intérêt

Kitesurf, surf, bodysurf et boxe.

Chaîne YouTube de sport et de voyage : plus de 100 000 vues, partenariats et événements.

Contact

martinlecorre@live.fr
+33 6 72 92 51 88
[linkedin.com/in/martinlecorre/](https://www.linkedin.com/in/martinlecorre/)

Septembre 2024 — En cours · Télécom Paris

Paris, France • Mastère spécialisé Big Data & science des données

Statistiques, apprentissage automatique avancé, apprentissage profond, exploitation des modèles d'apprentissage automatique, systèmes distribués pour les données massives, DevOps, Internet des objets, intelligence artificielle, vision par ordinateur, projet académique Airbus, projet ETL, traitement du langage, séries temporelles, Hadoop, Docker/Kubernetes, Python, Java et SQL.

2023-2024 · Année de césure — Moniteur de kitesurf IKO

Sri Lanka, Brésil, Zanzibar et Vietnam

Enseignement international du kitesurf dans quatre pays, en anglais, en français et en espagnol.

Expériences professionnelles

2022-2023 · Sofren (CMA-CGM) — Ingénieur naval R&D

Marseille, France • Département recherche et développement des navires

Responsable des études d'intégration de systèmes de propulsion vélique pour la flotte CMA-CGM : voiles rigides, cerfs-volants de traction et rotors Flettner. Participation à l'équipe chargée d'optimiser les pilotes automatiques.

- Évaluation technique de l'intégration des systèmes de propulsion vélique et développement des pilotes automatiques.
- Proposition de nouvelles solutions d'implantation adaptées à la flotte CMA-CGM.

2021-2022 · Ailes Marines (Iberdrola)

Saint-Brieuc, France • Assistant directeur Développement et Territoires

Participation au développement, à la construction et à l'exploitation du parc éolien en mer de Saint-Brieuc (500 MW).

- Réunions avec les services de l'État : préfecture maritime et DDTM.
- Responsable de l'appel à manifestation d'intérêt pour les ports de maintenance de Saint-Malo.
- Participation aux projets portuaires de Saint-Malo, SQP et Lezardrieux.

2020-2021 · SOFRESID Engineering — Ingénieur naval

Lorient, France

Études d'intégration de systèmes de propulsion vélique sur des navires commerciaux : pétroliers, méthaniers et très grands pétroliers (VLCC). Évaluation technique et économique des systèmes selon les règles de classe, de pavillon et les normes locales et internationales.

2020 · CorPower Ocean — Stage ingénieur essais et analyse

Stockholm, Suède

Campagne d'essais à sec avant le développement d'un prototype en mer et programmation Matlab. Stage interrompu en raison de la COVID-19 et de retards de livraison du banc d'essai.

2018 · UTM — Stage élève ingénieur

Johor Bahru, Malaisie

Détermination de l'incidence optimale de convertisseurs d'énergie des vagues pour maximiser le rendement. Programmation Matlab, étude des sites d'implantation en mer et simulations numériques avec ANSYS Fluent.

2015-2020 · EPF — Diplôme d'ingénieur généraliste

Paris, France • Spécialisation Énergie et Environnement

Formation généraliste de grande école française, avec projets pratiques et expérience internationale.