



Théo Joly

DIPLÔMÉ DU MASTER DATASCALES –
RECHERCHE UN PREMIER POSTE EN CDI
EN ÎLE-DE-FRANCE.

FORMATION

2026

MASTER EN GESTION DES DONNÉES ET EXTRACTION
DE CONNAISSANCES À LARGE ÉCHELLE

Université Paris-Saclay, Versailles

2024

LICENCE EN INFORMATIQUE

Université de Versailles Saint-Quentin-en-Yvelines, Versailles

2021

BACCALAURÉAT GÉNÉRAL - MATHÉMATIQUES,
NUMÉRIQUE ET SCIENCES INFORMATIQUES

Lycée Jules Vernes, Limours

PROFIL



theo.joly@theojoly.fr



Limours



Permis B - Véhiculé

OBJECTIF

Ingénieur Data & IA combinant architectures de données (Web Sémantique) et intégration de modèles de langage (RAG, MCP) pour concevoir des solutions d'analyse à large échelle.

COMPÉTENCES

Data Science & IA : Python (Expert), R, Machine Learning, LLMs (GenAI), RAG, MCP, Graphes, Recherche Opérationnelle.

Big Data & Data Engineering : SQL Avancé (PL/SQL), NoSQL, Hadoop Ecosystem, Elastic Stack (ELK), Web Sémantique (RDF/SPARQL), ETL (Talend).

Développement & Outils : Rust, Java, Docker, Git/GitHub, Linux (Bash), FastAPI, FastMCP.

Langues : Français (Langue maternelle), Anglais (Technique).

CENTRES D'INTÉRÊT

- Culture Japonaise** : Apprentissage de la langue, mangas et projet de voyage.
- Aménagement Paysager** : Rénovation d'espaces verts et plantation d'arbres.
- Cuisine** : Gastronomie.

EXPÉRIENCE PROFESSIONNELLE

STAGE DE FIN D'ÉTUDES – DATA & IA

Ecolab - CGDD, Ministère de la Transition Écologique | avr. - sept. 2026

Développement de **HarmoniA**, un assistant IA de modélisation de données pour aligner les standards sémantiques. Conception d'une architecture RAG avec LLMs locaux (Qwen, Gemma) et base vectorielle (Qdrant), couplée à un serveur Model Context Protocol (MCP). Définition des scénarios et conduite des tests utilisateurs. (Technos : Python, FastMCP, Qdrant, OWL).

MARAÎCHER

Ferme Lafouasse | juil. 2025

PROJETS ACADÉMIQUES & RÉALISATIONS

Exploration Hybride de Data Lakes (M2 DataScale) : Conception d'un moteur de recherche sémantique innovant combinant LLMs (GenAI) et Graphes de Connaissances. L'outil permet d'identifier automatiquement les tables pertinentes (Union/Jointure) au sein de lacs de données massifs pour enrichir les analyses.

Data Engineering & Architecture SQL : Double compétence en gestion de la donnée : Création d'un **pipeline ETL** complet sous Talend (audit qualité, nettoyage, consolidation) et architecture d'une base de données complexe en **PL/SQL** (Triggers, Procédures stockées) pour un réseau social.

Algorithmique sur Graphes & Recherche Opérationnelle (TER) : Étude avancée sur la structure des réseaux : Comparaison d'heuristiques (**Tabu Search**, **Recuit Simulé**) pour l'optimisation de l'isolation virale dans des graphes orientés. Réimplémentation critique de l'algorithme **PageRank** sur données massives.