

Sylvain Delahaies

Data Scientist | Data Engineer | Research Scientist

En bref

Avec un parcours combinant plus de dix ans de recherche académique en sciences des données et une expertise en développement de solutions IA, j'apporte une solide expérience des modèles de machine learning, deep learning et IA générative. Mon approche se distingue par l'innovation, la rigueur et une forte aptitude à collaborer dans des environnements pluridisciplinaires.

Contact & Info

✉ sdelahaies@gmail.com

🌐 Web site: sdelahaies.fr

🐙 Github: sdelahaies

🌐 LinkedIn: sylvaindelahaies

📄 Research Gate

📍 Condat-sur-Vézère, Dordogne, FR

Langues

🇫🇷 Français - langue maternelle

🇬🇧 Anglais - maîtrise professionnelle

Formation

• **Data Engineering Jan-Apr 2023**
DataScientest, Paris.

• **PhD 2005-2008**
Mathematics & Fluid Dynamics
University of Surrey

• **MSc 2004-2005**
Hydro-Informatics
University of Surrey

• **MPhil 1999-2004**
Mathématiques et Applications
Université d'Angers

Intérêts & Activités

- Conseiller municipal
- Vice-Président *Commanderie de Condat Confluence Hospitalière*
- Technologies, computing, Linux, AI
- Piano, Guitare, Lutherie
- Course à pied, randonnée

COMPETENCES

- **Modélisation et apprentissage automatique** : expérience des modèles de machine learning et deep learning (CV, NLP, LLM, VLM, MOE).
- **Pipelines de données** : développement de workflows pour l'acquisition, la transformation et l'exploitation des données.
- **Programmation et développement logiciel** : solide expérience en Python et autres langages.
- **Déploiement d'IA** : mise en production de modèles avec orchestration, containerisation et CI/CD.
- **Bases de données et Big Data** : gestion de bases relationnelles, non relationnelles, orientées graphes, et solutions Big Data.
- **Cloud computing** : expérience avec les solutions cloud AWS (EC2, S3, ECS, Sagemaker, Terraform,...).
- **Développement d'applications web** : création d'applications interactives pour la visualisation et l'analyse de données.
- **Résolution de problèmes complexes** : approche analytique et application de l'état de l'art des sciences et techniques.
- **Collaboration pluridisciplinaire**
- **Veille scientifique et technologique**
- **Communication scientifique et vulgarisation**

EXPERIENCE

Data Scientist

4.71 - à distance

Août 2023-Nov 2024

Clermont Ferrand, FR

- Développement d'une méthodologie de calcul de l'empreinte carbone des entreprises sur la base de leurs données comptables.
- Développement d'outils d'IA pour l'analyse de documents, l'extraction d'informations structurées.
- Développement d'un orchestrateur pour la mise en production de pipelines de traitement de données.
- Création de jeux de données synthétiques et entraînement de modèles de CV, LLM, VLM et MOE.
- Développement d'un environnement RAG pour la comptabilité carbone.
- Développement d'applications web pour l'exploitation des données en lien avec la comptabilité carbone.
- Montage de dossiers de demande de subvention de recherche.

Chercheur Post-Doctorant & associé

University of Surrey - à distance

2009-2023

Guildford, UK

Assimilation de données et Processus Stochastiques Ponctuels

- Filtres bayésiens pour les processus stochastiques ponctuels.
- Apprentissage automatique pour les risques assurantiels liés à des phénomènes de subsidence.
- Analyse des données d'appels d'urgence aux services de police de Providence (US).
- UI pour la modélisation et le suivi de l'épidémie de Covid19.

Assimilation de données et cycle du Carbone

- Développement de méthodes variationnelles, filtres d'ensembles, méthodes de Monte-Carlo.
- Etude sur l'analyse de la dégradation forestière à l'aide de techniques d'assimilation de données satellites très haute résolution.
- Etude du contenu en information de flux d'observations et quantification de l'incertitude des stocks de carbone.

Enseignement: Numerical solutions for PDEs

2011-2012