LEAD DATA ENGINEER & ARCHITECT DATA

Paul Tuillet
Vernon, France • 07 61 92 44 09 • me@poltui.com

CDI - CDD - Freelance
Normandie - Ouest parisien - Paris - International
Présentiel - Télétravail Complet ou Partiel

INTRODUCTION

Lead Data Engineer & Architecte Data avec 14 ans d'expérience en ingénierie des données, architecture distributed-systems et infrastructure cloud/on-premise. Spécialiste du déploiement de Data Lakehouses (jusqu'à 10 Md requêtes/jour), de la modernisation de workflows legacy, du FinOps et de la gouvernance RGPD. Historique prouvé en création de départements Data (Founding Engineer), optimisation de coûts et pilotage technique de projets à fort impact business.

EXPÉRIENCE PROFESSIONNELLE

LEAD DATA ANALYTICS ENGINEER - Groupe BZ

Rouen, France | 04/2025 – 08/2026

Acteur majeur de l'agroalimentaire (exportation & commerce de céréales, 2,5 MT exportées, logistique multimodale).

- Cadrage stratégique & Gouvernance Data : Audit complet de l'écosystème data, animation des comités de direction (CODIR) et alignement des feuilles de route techniques sur les objectifs stratégiques du groupe.
- Création du Pôle Data & Conduite du changement : Structuration de A à Z du département Data en tant que Founding Engineer, instaurant une culture data-driven transverse pour 180+ utilisateurs.
- Sécurité & Conformité RGPD : Déploiement d'infrastructures sous Linux/Docker via Azure DevOps CI/CD, mise en conformité RGPD et implémentation du catalogue de données OpenMetadata.
- Architecture Data Lakehouse FinOps : Conception et déploiement d'un Lakehouse auto-hébergé (Docker, MinIO, PostgreSQL, dbt) en architecture médaillon, éliminant les coûts de licence.
- Haute Performance & Orchestration : Traitement analytique avec DuckDB et Polars, et orchestration résiliente et fiabilisée des pipelines ETL/ELT via Dagster.
- Développement d'Apps Métier & Digitalisation : Création d'applications web sur mesure (FastAPI, Streamlit), automatisant les processus métier et remplaçant 35+ fichiers Excel legacy complexes. Suppression de nombreuses tâches chronophages.

LEAD DATA & OSINT INVESTIGATOR - Baltic Shield Systems

Moscou, Russie | Distanciel | 12/2022 – 02/2023

Cabinet spécialisé en veille stratégique, analyse des risques fournisseurs et protection de la marque.

- Plateforme de Renseignement Cybersécurité : Architecture et déploiement d'une plateforme analytique avancée en environnement hautement critique et contraint.
- Ingénierie & Analyse de Flux Massifs : Audit, extraction et normalisation de flux de données mondiaux grâce à une expertise poussée en OSINT et data mining.
- Détection Proactive par IA / ML : Implémentation de modèles GenAI et de Machine Learning (GLiNER, XGBoost) pour l'identification précoce de vulnérabilités et de signaux faibles.
- Collaboration technique transverse pour résoudre les failles critiques et renforcer les protocoles de défense.

DATA ENGINEER CONSULTANT - Wildberries

Moscou, Russie | 03/2022 – 03/2024

Leader du e-commerce en Europe de l'Est avec +150M de clients.

- Unification des Silos Data & Scalabilité : Consolidation des pipelines de données régionales (10 Md requêtes/jour, 60 Md€ GMV) en garantissant la conformité réglementaire locale et un SLO cadre.
- Data Engineering & NLP : Automatisation du traitement du langage et du nettoyage des catalogues produits multi-langues.
- Modélisation Prédictive & Analytics : Développement de modèles avancés de ciblage et de segmentation client à très forte volumétrie pour optimiser les taux de conversion.

DATA MINING ENGINEER - Consultant OSINT Indépendant

01/2012 - 07/2022

Expertise en collecte de données massives, renseignement open-source (OSINT) et R&D à l'international.

- Clients & Partenaires : BBC, Le Monde, Interpol, Institute for the Study of War (ISW), Oryx, LTT, ainsi que des acteurs gouvernementaux et militaires(Arabie Saoudite, Australie, Estonie, Pologne, France, Danemark, Pays-Bas).
- Outils & Data Mining : Conception de pipelines Python sur mesure pour l'extraction et le traitement de données massives et non structurées en asynchrone ou temps réel.
- Investigation Forensique (GEOINT / SOCMINT) : Analyse d'imagerie satellite, de métadonnées de réseaux sociaux et de registres publics pour la rédaction de rapports de renseignement critiques à haute fréquence.
- Sûreté & Géopolitique : Suivi de réseaux illicites et analyse d'événements à fort enjeu en soutien aux forces de l'ordre et services spécialisés.

TEACHER & DATA PROGRAM ASSISTANT - LePont Learning

Levallois-Perret, France | Distanciel | 05/2022 – 08/2024

Établissement privé de formation professionnelle B2B/B2C aux métiers de la donnée.

- Formation Professionnelle & Executive : Conception et animation de programmes intensifs en Python et Data Visualisation (Power BI) pour 150 étudiants (Audencia) et 80 cadres dirigeants (KIKO Milano).
- Ingénierie Pédagogique : Élaboration de curriculums alignés sur les standards industriels en partenariat académique avec UC Berkeley et le Fischer Center.
- Mentorat & Reconversion : Accompagnement technique individuel de 380 apprenants vers les métiers Data, maintenant un taux de satisfaction exceptionnel de 95 %.

SENIOR SERVER ADMINISTRATOR - Keen Software House

Prague, Tchéquie | Distanciel | 08/2018 – 02/2022

Studio indépendant d'édition et développement de jeux vidéo (Space Engineers, Medieval Engineers).

- Administration Infrastructure Bare-Metal : Gestion et optimisation d'un parc de 35 serveurs dédiés sous Proxmox et Linux, garantissant un SLA de 99.9% pour 24 000 utilisateurs simultanés.
- Optimisation des Performances & Virtualisation : Reconfiguration fine du kernel Linux et de la virtualisation Proxmox, réduisant la latence serveur de 30% tout en opérant des environnements hybrides.
- Ingénierie Économique In-Game : Conception et implémentation d'un système économique complexe basé sur l'analyse de données comportementales des joueurs.

FORMATIONS & DIPLOMES

- Parcours autodidacte en ingénierie de la donnée et en infrastructure depuis 2012, validé par 14 ans de production.
- Enseignant invité en Python et Data Visualisation : KEDGE, Audencia, KIKO Milano, profils masters à doctorats.

COMPÉTENCES CLÉS

Architecture & Dataops	Medallion, Modern Data Stack, Data Lakehouse, DataOps, FinOps, Dagster, dbt, Prefect, SQLMesh, Airflow, Spark / PySpark, Kafka, Delta Lake, Iceberg, ETL / ELT, GCP, AWS, Azure, n8n
Infrastructure & CI/CD	GitLab CI, GitHub Actions, Terraform, Ansible, Docker, MinIO / S3, Linux, GCP, AWS, Kubernetes, CI/CD
Données & Modélisation	DuckDB, Polars, Snowflake, BigQuery, ClickHouse, PostgreSQL, Redshift, Data Vault, Kimball, Metabase, OpenMetadata, Databricks, MS Fabric
Langages & Dev	Python, SQL, JavaScript, FastAPI, Streamlit, Git, Bash
Ia - R&D	Pinecone, Vector DB, LLM Integration, RAG, LLM, MLflow, Ollama, Claude, Gemini
Leadership & Produit	Roadmap Technique, Direction de Projet, Mentorat, Gouvernance des Données, Arbitrages d'Architecture, Data Contracts, Lineage
Languages	Français (C0), Anglais (C1), Allemand (A2), Russe (A2)

PROJETS PERSONNELS

- **Arken** (Docker, PHP, JS, Python, LLM Ollama, n8n, Pinecone) : Terminal de renseignement géopolitique fondé sur des sources ouvertes (OSINT). Intègre une cartographie d'événements en temps réel, des algorithmes de détection de signaux faibles et la construction de graphes de connaissances.
- **Databox** (Docker, Python, LLM, RAG, Dagster) : Solution matérielle et logicielle bas coût embarquant des services d'IA sur Raspberry Pi à destination des TPE et PME, afin de mettre à disposition un service data autonome.
- **Vortex** : Terminal d'analyse financière en temps réel.
- **Russia With My Cat** : Plateforme d'exploration et de cartographie interactive.
- **Tresor Interactive & Hexengard** : Projets de développement vidéoludique.
- **Space Engineers France** : Fondateur et animateur de la communauté francophone.