LEAD DATA ENGINEER & DATA ARCHITECT

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

CDI - CDD - Freelance
Normandie - Ouest parisien - Paris - International
On-site - Full or Partial Remote

SUMMARY

Lead Data Engineer & Data Architect with 14 years of experience in data engineering, distributed systems architecture, and cloud/on-premise infrastructure. Specialist in deploying Data Lakehouses (handling up to 10 billion requests/day), modernizing legacy workflows, FinOps optimization, and GDPR compliance. Proven track record in founding Data departments, cost optimization, and technical lead on high-impact business projects.

PROFESSIONAL EXPERIENCE

LEAD DATA ANALYTICS ENGINEER - Groupe BZ

Rouen, France | 04/2025 – 08/2026

Major agri-food player (grain export & trading, 2.5 MT exported, multimodal logistics).

- Strategic Framework & Data Governance: Conducted full data ecosystem audit, led Executive Committee (CODIR) meetings, and aligned technical roadmaps with strategic group goals.
- Data Department Creation & Change Management: Structured the Data department from scratch as Founding Engineer, establishing a cross-functional data-driven culture for 180+ users.
- Security & GDPR Compliance: Deployed Linux/Docker infrastructure via Azure DevOps CI/CD, ensured GDPR compliance, and implemented OpenMetadata catalog.
- Data Lakehouse FinOps Architecture: Designed and deployed a self-hosted Lakehouse (Docker, MinIO, PostgreSQL, dbt) using medallion architecture, eliminating licensing costs.
- High Performance & Orchestration: Enabled analytical processing with DuckDB and Polars, and built resilient ETL/ELT pipelines using Dagster.
- Business App Development & Digitalization: Built custom web applications (FastAPI, Streamlit), automating business processes and replacing 35+ complex legacy Excel files, removing numerous time-consuming tasks.

LEAD DATA & OSINT INVESTIGATOR - Baltic Shield Systems

Moscow, Russia | Remote | 12/2022 – 02/2023

Firm specializing in strategic intelligence, supplier risk analysis, and brand protection.

- Cybersecurity Intelligence Platform: Architected and deployed an advanced analytics platform in a highly critical and constrained environment.
- Mass Data Flow Engineering & Analysis: Audited, extracted, and normalized global data flows using deep OSINT and data mining expertise.
- Proactive AI / ML Detection: Implemented GenAI and Machine Learning models (GLiNER, XGBoost) for early identification of vulnerabilities and weak signals.
- Cross-functional technical collaboration to resolve critical vulnerabilities and strengthen defense protocols.

DATA ENGINEER CONSULTANT - Wildberries

Moscow, Russia | 03/2022 – 03/2024

E-commerce leader in Eastern Europe with 150M+ customers.

- Data Silo Unification & Scalability: Consolidated regional data pipelines (10B requests/day, €60B GMV) while ensuring local regulatory compliance and maintaining SLA targets.
- Data Engineering & NLP: Automated natural language processing and multi-language product catalog cleaning.
- Predictive Modeling & Analytics: Developed advanced targeting and high-volume customer segmentation models to optimize conversion rates.

DATA MINING ENGINEER - Independent OSINT Consultant

01/2012 - 07/2022

Expertise in massive data collection, open-source intelligence (OSINT), and international R&D.

- Clients & Partners: BBC, Le Monde, Interpol, Institute for the Study of War (ISW), Oryx, LTT, as well as government and military entities (Saudi Arabia, Australia, Estonia, Poland, France, Denmark, Netherlands).
- Tools & Data Mining: Designed custom Python pipelines for extracting and processing massive unstructured data asynchronously or in real time.
- Forensic Investigation (GEOINT / SOCMINT): Analyzed satellite imagery, social media metadata, and public records for high-frequency critical intelligence reports.
- Security & Geopolitics: Tracked illicit networks and analyzed high-stakes events supporting law enforcement and specialized agencies.

TEACHER & DATA PROGRAM ASSISTANT - LePont Learning

Paris, France | Remote | 05/2022 – 08/2024

Private vocational training institution (B2B/B2C) for data professions.

- Professional & Executive Training: Designed and delivered intensive Python and Data Visualization (Power BI) programs for 150 students (Audencia) and 80 executives (KIKO Milano).
- Instructional Design: Developed industry-aligned curricula in academic partnership with UC Berkeley and the Fischer Center.
- Mentorship & Career Transition: Provided individual technical coaching to 380 learners making career transitions to Data roles, maintaining a 95% satisfaction rate.

SENIOR SERVER ADMINISTRATOR - Keen Software House

Prague, Czechia | Remote | 08/2018 – 02/2022

Independent video game developer and publisher (Space Engineers, Medieval Engineers).

- Bare-Metal Infrastructure Administration: Managed and optimized 35 dedicated servers running Proxmox and Linux, ensuring 99.9% SLA for 24,000 concurrent players.
- Performance Optimization & Virtualization: Fine-tuned Linux kernel and Proxmox virtualization, reducing server latency by 30% in hybrid environments.
- In-Game Economic Engineering: Designed and implemented a complex in-game economic system based on player behavioral data analysis.

EDUCATION & DEGREES

- Self-taught in data engineering and infrastructure since 2012, backed by 14 years of production experience.
- Guest Lecturer in Python & Data Visualization: KEDGE, Audencia, KIKO Milano (Master's to PhD level profiles).

KEY SKILLS

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
Data & Modeling	DuckDB, Polars, Snowflake, BigQuery, ClickHouse, PostgreSQL, Redshift, Data Vault, Kimball, Metabase, OpenMetadata, Databricks, MS Fabric
Languages & Dev	Python, SQL, JavaScript, FastAPI, Streamlit, Git, Bash
AI - R&D	Pinecone, Vector DB, LLM Integration, RAG, LLM, MLFlow, Ollama, Claude, Gemini
Leadership & Produit	Technical Roadmap, Project Management, Mentorship, Data Governance, Architecture Trade-offs, Data Contracts, Lineage
Languages	French (Native/C2), English (Full Professional/C1), German (Elementary/A2), Russian (Elementary/A2)

PERSONAL PROJECTS

- **Arken** (Docker, PHP, JS, Python, LLM Ollama, n8n, Pinecone): Geopolitical intelligence terminal based on open-source intelligence (OSINT). Features real-time event mapping, weak signal detection algorithms, and knowledge graph construction.
- **Databox** (Docker, Python, LLM, RAG, Dagster): Low-cost hardware/software solution running AI services on Raspberry Pi for SMBs, providing autonomous data capability.
- **Vortex**: Real-time financial analysis terminal.
- **Russia With My Cat**: Interactive mapping and exploration platform.
- **Tresor Interactive & Hexengard**: Video game development projects.
- **Space Engineers France**: Founder and manager of the French-speaking community.