



Gueulette Arnaud

Cloud Data Engineer | Data Platform & Governance Practitioner

Location: Brussels, Belgium | **Phone:** +32 488 69 25 31 | **Email:** arnaud_gueulette@outlook.fr



Professional Summary

Data Consultant with 3 years of experience in designing and delivering data systems, platforms, pipelines, and ML models. Specialized in Infrastructure as Code (IaC) using Terraform across AWS, with hands-on expertise in Python, SQL, and Docker. Proven track record in automating cloud infrastructure, building end-to-end data solutions, and translating business needs into scalable technical architectures.

Professional Experience

Data Consultant, Datashift NV

Feb 2025 – Present

- Scania CV – Data Governance Initiative

Feb 2025 – Present

- Leading cloud data engineering initiatives to support the integration of data governance practices at Scania, combining functional ownership of use case delivery with hands-on technical implementation in both Collibra and AWS (DEV, PROD). Responsibilities include designing and deploying end-to-end solutions leveraging Collibra Data Quality, metadata ingestion frameworks, and custom workflows, as well as building scalable AWS-based services to automate governance processes, enable lineage tracking, and enhance data quality monitoring.

- Datashift – Internal Projects

Feb 2025 – Present

- Developed an automated **Soda-Agent deployment framework** in AWS using **Terraform** and **Terragrunt**, enabling full setup within minutes through reusable infrastructure-as-code modules.
- Designed and implemented a complete **Data Quality & Governance demo environment** integrating **Soda**, **dbt**, **Airflow**, **Superset**, and **Python** for automated data validation and analytics orchestration.
- Integrated **Collibra** as the governance backbone, linking data quality metrics and lineage to catalog assets for improved visibility and accountability across the data lifecycle.
- Demonstrated **end-to-end data observability** and compliance readiness through practical data pipelines, promoting reusable best practices for client engagements.

Data Consultant, Sopra Steria Benelux

Sept 2022 – Jan 2025

- eu-LISA – Schengen Information System (SIS)

June 2024 – Dec 2024

- Enhanced and maintained eu-LISA's border security systems by developing, testing, debugging, and deploying improvements (PREPROD). Ensured system performance and compliance with EU standards. Collaborated across teams to integrate enhancements and documented processes to streamline future maintenance.

- EUROCONTROL – ISOBAR

March 2023 – Oct 2023

- Designed and deployed a multi-step flight forecasting model (via online endpoint) to support air traffic operators in optimizing arrival/departure flow based on weather conditions. The model significantly reduced manual workload and enabled data-driven operational decisions.

- Sopra Steria Benelux – Internal Projects

Sept 2022 – Jan 2025

- Researched and documented AI-powered development tools (e.g. Copilot, prompt engineering) to support SDLC tasks, highlighting productivity benefits and privacy risks.
- Evaluated and prototyped Elasticsearch integration into existing data systems to improve search performance and query expressiveness.
- Built an end-to-end pipeline to extract and process PDF-based energy consumption invoices into the SAS IBM Envizi ESG Suite, improving ESG reporting and internal analysis.
- Delivered a POC data system that extracted KPIs from energy invoice data, supporting sustainability and cost-efficiency decisions.

Education

- Master's Degree in Data Science

Option: Statistics / GPA: 3.0

Université Catholique de Louvain, 2020 – 2023

- Bachelor's Degree in Business Engineering

Technical Skills

Programming & Scripting

Python, SQL, Linux/Bash, Java

Data Engineering & Pipelines

Airflow, pandas/pySpark, dbt, DVC, MLflow

Data Systems

PostgreSQL, Microsoft SQL Server, MySQL, MongoDB, Elasticsearch, Snowflake, Dremio

Libraries & Frameworks

Scikit-learn, PyTorch, TensorFlow, SQLAlchemy, Pydantic, PyTest, Flask, FastAPI, ...

Cloud Platforms

AWS: S3, EC2, RDS, Glue, Athena, Lambda, Lake Formation, EventBridge, EKS, CloudWatch...

Azure: Data Lake, Data Factory, ML Studio

Dev, Infrastructure & Ops

Terraform, Terragrunt, Docker, Rancher, GitLab CI, Github Actions, Helm

Data Governance, Quality & BI

Collibra: Catalog, Lineage, Workflow, Integration, Collibra DQ, Soda, Great Expectations, Power BI, Grafana, Superset

Other Tools

Kafka, Debezium, Power Automate, ...

Soft Skills

Critical Thinking

Problem Solving

Continuous Learning

Agile & Iterative Delivery

Teamwork & Collaboration

Business-Driven Mindset

Languages: French (C2), English (C1), Dutch (B1)

Certifications

- [Collibra Solution Architect Certification](#)
- [AWS Certified Developer – Associate \(DVA-C02\)](#)
- [AWS Certified Solutions Architect – Associate \(SAA-C03\)](#)
- [AWS Certified CloudOps Engineer – Associate \(SOA-C03\)](#)
- [AWS Certified Data Engineer – Associate \(DEA-C01\)](#)

Other Achievements

- [EU Defence Innovation Scheme \(EUDIS\) – Hackathon 2024](#)
Team EUnify – 1st (Belgium) – 3rd (Europe)