

FEDERICO FIORIO

Data Engineer & ML Engineer

+39 342 596 4621 | fioriofederico99@gmail.com

[GitHub](#) | [LinkedIn](#) | [Portfolio](#)

Scalable Systems Engineer specializing in distributed computing, cloud-native infrastructure (AWS/GCP), and AI integration. Proven experience designing **reusable software frameworks** and internal tooling that automate complex workflows. Expertise in building resilient, high-throughput architectures that support data-driven decision-making.

EXPERIENCE

Data & Machine Learning Engineer | Data Reply, Milan, Italy

May 2024 – Present

- **LLM Fine-Tuning: Fine-tuned Llama 3** for fashion intent classification using **LoRA (Low-Rank Adaptation)** and **quantization**, achieving high accuracy while optimizing GPU compute resources on Databricks.
- **RAG-powered Tooling:** Developed a UI-driven engine to generate **AWS CloudFormation** templates from natural language. Optimized prompts to guarantee well-written YAML output and lower **Time To First Token (TTFT)**.
- **Data Pipelines:** Designed high-volume ingestion pipelines using **Databricks and Medallion Architecture**. Optimized historical tracking with **SCD Type 2**, reducing query time by 80% and more.
- **Sanctioned Shop Blocker:** Optimized the customer onboarding sanctions pipeline by refactoring matching logic from **permutation-based ($O(n!)$)** to **token-sorting ($O(n \log n)$)**, reducing data storage by **75%** and significantly decreasing execution time
- **Orchestration:** Migrated **Airflow** DAGs from Composer 1.x to **GCP Composer 2.x**, increasing scheduling efficiency by 10x
- **CI/CD:** Engineered **Azure DevOps** pipelines for Databricks, leveraging **Terraform** for Unity Catalog, **DABs** for deployment and integration testing with **DQx** for data quality, and **Pytest** for unit testing.

Data Engineer | Management Solutions, Milan, Italy

October 2023 - April 2024

- **IBM DataStage refactoring:** Optimized IBM DataStage ETL workflows and **SQL** for banking data warehouses.
- **Data Governance & Compliance:** Enforced **data governance** policies in cross-functional teams

Software Engineer Intern | Dilium, Milan, Italy

February 2021 - May 2021

- **GANs Research:** Conducted research on **GANs** for deepfake generation, emphasizing ethical considerations

EDUCATION

- **MSc in Computer Science | University of Milan, Erasmus+ (University of Copenhagen)** Grade: 108 / 110
[Thesis:](#) Text-to-image diffusion models attacks in scarce resource environments and privacy issues
- **BSc in Computer Science | University of Milan**

CERTIFICATIONS

- **All certifications:** Databricks Certified Data Engineer Professional, Databricks Certified Associate Developer for Apache Spark, Databricks Certified Data Engineer Associate, Neo4j Certified Professional, AWS Certified Machine Learning – Specialty, Professional Machine Learning Engineer Certification

PERSONAL PROJECTS

- **NASA Space Apps Challenge 2025 Winner (Zurich):** Led a team to develop an AI-driven traffic optimization platform using **Reinforcement Learning** and **SUMO**, designed to reduce urban congestion and emissions through dynamic traffic light coordination. | [Project](#)
- **Real-Time Edge Detection System:** Along with a friend, engineered and deployed a computer vision safety system currently controlling operations at **2 commercial car wash facilities**, used tflite for optimized edge performance.
- **Personalized Board Games RAG:** Engineered a **Retrieval-Augmented Generation (RAG)** application allowing users to query rulebooks via natural language. Built using **OpenAI embeddings** and **Qdrant** for vector search to retrieve precise game mechanics, with **MongoDB** for chat history persistence and **Docker** to make it easily available. | [Repo](#)
- **Patient journey Analysis:** Built an **LLM-powered pipeline** to extract clinical insights from unstructured patient symptom descriptions, transforming raw narratives into structured data to accelerate diagnostic accuracy and gain insights for **sales team** | [Repo](#)

SKILLS

- **AI & ML:** Prompt Authoring, Fine-tuning (PEFT/LoRA, DPO), Langchain, PyTorch, Vector Databases (Qdrant), Scikit-learn, MLflow
- **Data Engineering:** Databricks, Apache Spark, SQL, NoSQL (MongoDB), Airflow, GCP Composer, BigQuery.
- **Languages & Core Tech:** Python, Java, SQL, PySpark, Docker, Terraform, K8s, ArgoCD, Github Actions, Azure devOps
- **Foundations:** Distributed Architecture (CRDTs, CAP theorem, HLD, LLD), Data Structures & Algorithms ([Top 10% LeetCode](#)).
- **Agents Automation:** [n8n](#)