



Bologna, Italia | lorenzo.ricci@gmail.com | +39 346 134 5678 | linkedin.com/in/lorenzoricci-data | github.com/lorenzoricci

Lorenzo Ricci

JUNIOR DATA ENGINEER

PROFILO

Laureato Magistrale in Informatica all'Università di Bologna con **109/110**. Stage di 6 mesi in Enel Digital con sviluppo di **5 pipeline ETL** Apache Spark in produzione che processano **2,8 milioni di record al giorno** da contatori smart. Competenze in Python, Snowflake, dbt e Apache Airflow. Certificazione Databricks Spark 3.0.

ISTRUZIONE

Titolo e specializzazione

Università di Bologna - Alma Mater Studiorum, Bologna
Set 2022 - Apr 2025

Titolo e specializzazione

Università di Bologna, Bologna
Set 2019 - Lug 2022

COMPETENZE

- Python (PySpark, Pandas)
- Apache Spark 3.x
- Apache Airflow
- Snowflake
- dbt (data build tool)
- SQL Avanzato
- Azure Data Lake
- Apache Kafka (base)
- Docker / Kubernetes
- Great Expectations
- Git / GitHub Actions
- Databricks

LINGUE

- Italiano (madrelingua)
- Inglese (C1)

REFERENZE

Ing. Simone Barone

Lead Data Engineer, Enel Digital S.p.A.
s.barone@enel.com, +39 06 8305 0000

ESPERIENZA

Tirocinante Data Engineer, Enel Digital S.p.A., Roma

Set 2024 - Feb 2025

Stage di 6 mesi nel team Data Platform di Enel Digital, responsabile dell'infrastruttura dati per i sistemi di smart metering e grid management.

- Sviluppato **5 pipeline ETL** con Apache Spark 3.4 / PySpark per l'ingestione e trasformazione di **2,8 milioni di record al giorno** da contatori smart distribuiti su 18 regioni italiane
- Orchestrato i workflow con **Apache Airflow 2.7** su Kubernetes: 5 DAG in produzione con monitoring, alerting e retry automatico su failure
- Progettato e implementato **3 layer di data mart** in Snowflake (staging, intermediate, final) con modello dimensionale per analisi consumi
- Introdotto dbt per la gestione delle trasformazioni SQL con **42 modelli**, test di qualità e documentazione automatica del data lineage
- Ridotto il tempo di esecuzione delle pipeline più critiche del **34%** tramite ottimizzazione delle partizioni Spark e caching strategico

Ricercatore (Tesi Triennale), Università di Bologna - Laboratorio AIA, Bologna

Mar 2023 - Set 2023

Ricerca applicata su sistemi di data quality per pipeline di machine learning.

- Costruito un dataset di benchmark con **500.000 record** simulati per testare framework di data quality (Great Expectations, Soda Core)
- Pubblicato i risultati in una short paper presentata al workshop ITADATA 2023 a Pisa

CERTIFICAZIONI

Databricks Certified Associate Developer for Apache Spark 3.0, Databricks

Gen 2025 - Feb 2025

PROGETTI

5 pipeline ETL PySpark in produzione: **2,8 milioni record/giorno** da 18 regioni. dbt con 42 modelli in Snowflake. Riduzione tempo esecuzione del **34%**. Monitoring con Airflow e alerting su Slack.

3 pull request merged nel repository dbt-utils con nuove funzioni di manipolazione date per Snowflake. **340 download** mensili del package aggiornato dalla community dbt.

ATTIVITÀ EXTRACURRICOLARE

Co-fondatore - DataEngineering Bologna Meetup

Giu 2023 - Apr 2025

Comunità locale di data engineering e data platform a Bologna. Organizzato **8 meetup** con relatori da Enel, Yoox Net-a-Porter e Cerved. Media di **45 partecipanti** per incontro. Canale YouTube con 320 iscritti.