Martin Spirovski

Skopje, North Macedonia | mspirovski9@gmail.com

Professional Summary

Results driven Data Engineer with 5+ years of experience designing and delivering scalable ETL pipelines, cloud native infrastructures, and analytics ready data platforms. Proven expertise in AWS, Azure, Databricks, PostgreSQL, Apache Airflow, and Python. Skilled in implementing Medallion, Kimball, Inmon Architecture, orchestrating workflows with Azure Data Factory, optimizing pipelines to improve data reliability, and ensuring compliance in regulated environments. Founder of a consultancy specializing in automated ingestion, transformation, and BI integration.

Experience

Data Engineer — Resume Reader (Feb 2025 – Present)
- Developing AI powered resume screening platform using RAG architecture on ChatGPT.
- Architected secure AWS infrastructure, reducing processing latency by 12%.
- Designed PostgreSQL schemas and ingestion pipelines to streamline candidate evaluation.

Data Engineer — DataLing.io (Founder, Data Engineering Consultancy) (Mar 2025 - Present)
- Founded and lead DataLing.io, delivering scalable ETL pipelines, cloud infrastructure, and analytics-ready platforms.
- Built production grade data workflows using Python, Apache Airflow, and PySpark, improving ETL reliability by 15%.
- Architected and deployed AWS infrastructures (Lambda, S3, RDS, EC2, IAM) and Azure/Databricks Medallion Architectures.
- Delivered HIPAA compliant solutions for healthcare clients with zero security incidents.
- Designed RAG-based AI systems integrating ChatGPT for document processing and enterprise search.

Data Engineer — CRM for Psychologists (Jun 2024 – Jan 2025)
- Built and maintained **end-to-end ETL pipelines**:
  - **Extraction** with Python (APIs) for batch ingestion of psychologist records.
  - **Transformation** using PySpark (Python) for data cleaning and normalization.
  - **Loading** into **AWS S3 and RDS** for secure storage and analytics.
- Orchestrated workflows and **monitored cron jobs** with Apache Airflow, ensuring reliable data delivery.
- Leveraged **AWS Glue** to automate and scale ETL processes, improving maintainability and reducing manual intervention.
- Designed and managed **PostgreSQL databases** for structured and archived physical records, ensuring secure, compliant storage.

Data Engineer — upperway.io (Nov 2024 – Feb 2025)

- Managed IAM policies and secure access, ensuring compliance with audit requirements.
- Optimized PostgreSQL instances to improve query performance.

Data Engineer — Causality (Dec 2021 – Nov 2023)
- Deployed ETL pipelines with Python, Airflow, and PySpark, increasing data throughput by 10%.
- Maintained PostgreSQL data warehouse, improving BI reporting responsiveness.
- Introduced validation scripts that increased pipeline stability.

Data Engineer and Web Scraper — Upwork (May 2020 – Nov 2021)
- Built custom scraping pipelines in Python, automating repetitive client workflows.
- Created automated extraction tools, delivering data with high reliability.

Technical Skills

- **Languages & Tools:** Python, SQL, Bash
- **Cloud:** AWS (Lambda, S3, EC2, IAM, Elastic Beanstalk, Route 53), Azure (Data Factory, Databricks, Storage Accounts)
- **Workflow & Orchestration:** Apache Airflow, dbt, PySpark, Azure Data Factory
- **Architecture:** Medallion, Kimball, Inmon Architecture, Data Warehousing, API Ingestion
- **Databases:** PostgreSQL, Redshift, RDS, MongoDB, DuckDB, DBeaver
- **DevOps & Infra:** Docker, Git, Linux
- **Other:** Web scraping (Selenium, BeautifulSoup, Scrapy)

Education

Data Science Academy, Brainster - 2021
- Specialized in data science, machine learning, and analytics.

Medical Faculty Skopje, BSc in Physical Therapy - 2011 to 2014