

# Best Technology Talent For Companies Around The World



## Henry

Senior Data Engineer

### Summary

Senior Lead Data Engineer with 8+ years of experience in big data, data architecture, data engineering, and cloud-based data platforms. Strong expertise across Azure and AWS cloud services, with hands-on experience building end-to-end data pipelines, Medallion Architecture, ETL/ELT processes, data migration, and BI solutions.

Proficient in Python, PySpark, Scala, SQL, and frameworks including Databricks, Snowflake, Spark, Django, and Airflow. Experienced in leading teams and delivering large-scale data migration and transformation projects across industries including automotive, insurance, aviation, and AgriTech. Based in Colombia.

#### TECHNICAL SKILLS:

- Cloud: Azure (Databricks, Synapse, Data Factory, Data Lake Gen2, Blob Storage, Logic Apps, Function Apps, Event Grid, Event Hub, Service Bus, Key Vault), AWS (Lambda, Glue, S3, API Gateway, CloudWatch, Event Bridge)
- Languages: Python, PySpark, Scala, R, SQL, Node.js
- Data Platforms: Snowflake, Databricks, SAP HANA, Teradata, MongoDB, PostgreSQL, SQL Server, Sybase, DB2
- Frameworks & Tools: Spark, Airflow, Django, ETL/ELT, DWH, Medallion Architecture, Data Mesh
- BI & Visualisation: Power BI, QlikView, Matplotlib, Pandas
- Integration: REST APIs, OAuth2/JWT, MQTT, SharePoint, ServiceNow, Salesforce, SAP
- DevOps & Agile: GitHub, Jenkins, Azure DevOps, Jira, Scrum/Kanban, CI/CD pipelines
- Other: SSMS, Postman, MongoDB Atlas, Data Migration Service

# Professional Career

## Data Architect / Senior Data Engineer

June 2024 – Present | Colombia

### Project Info:

Large-scale data migration and processing platform on Azure, covering heterogeneous source systems and real-time event-driven integration.

### Responsibilities:

- Data processing, transformation, and advanced analytics with Databricks (PySpark, SQL, Python)
- Migrated data from heterogeneous sources — SQL Server, Sybase, AS400 DB2, SFTP, flat files into Azure
- Implemented Medallion Architecture (Bronze, Silver, Gold) for structured and governed data pipelines
- Developed ETL processes and pipeline orchestration with Azure Data Factory and Synapse Analytics
- Built Azure Logic Apps workflows and Function Apps (Node.js) for automation and system integration
- Designed event-driven solutions using Event Grid, Event Hub, and Service Bus
- Integrated REST services with OAuth2 (Private Key JWT) authentication
- Built monitoring and error dashboards in Power BI; performed SQL debugging and query optimisation.
- Automated SharePoint file transfers to external APIs and implemented real-time integration with ServicePlatform
- Maintained CI/CD pipelines and DevOps processes using GitHub and Jenkins

### Software and technologies:

Databricks, PySpark, Synapse, Data Factory, Data Lake Gen2, Blob Storage, Logic Apps, Function Apps, Event Grid, Event Hub, Service Bus, Key Vault, Snowflake, Power BI, Python, SQL, Node.js, REST APIs, OAuth2/JWT, MQTT, GitHub, Jenkins, SharePoint, Data Mesh

## Senior Lead Data Architect

January 2024 – June 2024 | Colombia

### Project Info:

Cloud migration project. Migrated terabytes of on-premise data to Azure with minimal downtime, leading a team of 3 engineers.

#### Responsibilities:

- Transformed and cleansed terabytes of on-premise data into Azure using Synapse and Databricks
- Minimised downtime by migrating on-premise databases to Azure SQL with Database Migration Service
- Utilised Azure Blob Storage, Data Lake Storage, and File Storage for scalable data staging and storage (Bronze, Silver, Gold layers)
- Led and mentored a team of 3 engineers throughout cloud data migration and orchestration

#### Software and technologies:

Databricks, SAP HANA, Synapse, Data Factory, Data Lake, Blob Storage, Database Migration Service, SQL Pools, Logic Apps, Python, PySpark, SSMS, Snowflake, SAP, REST APIs

## Senior Data Engineer

March 2023 – September 2023 | Colombia

#### Project Info:

Social media web scraping platform. Built RESTful API endpoints, web scraping modules, and serverless data pipelines on AWS.

#### Responsibilities:

- Implemented RESTful API endpoints and web scraping modules integrated with MongoDB
- Designed RESTful APIs using AWS Lambda, Glue, and API Gateway
- Developed web scraping modules with Python and BeautifulSoup targeting Facebook, TikTok, and Instagram
- Leveraged AWS Lambda for serverless execution, optimising resource usage and scalability
- Managed data storage in MongoDB Atlas and S3 for reliable retrieval and manipulation

#### Software and technologies:

AWS Lambda, Glue, S3, CloudWatch, EventBridge, API Gateway, Python, PySpark, MongoDB Atlas, Snowflake, SQL, REST APIs

## Senior BI Data Engineer

June 2022 – March 2023 | Colombia

#### Project Info:

Reaccommodation platform for flight and reservation management, integrating a DWH and real-time monitoring.

#### Responsibilities:

- Designed Django API endpoints for flight and reservation management, integrating with the DWH

- Implemented business logic for reservation rearrangement, capacity management, and passenger preference handling
- Integrated Snowflake for scalable data storage and analysis across multiple source systems
- Developed monitoring and logging for API availability, performance, and debugging support

#### **Software and technologies:**

Django, AWS Lambda, Glue, S3, CloudWatch, EventBridge, API Gateway, Python, PySpark, SQL, Synapse, Data Lake, MongoDB, Snowflake, Power BI, REST APIs

## **Senior Lead Data Migration Consultant**

**December 2021 – February 2023** | Colombia

#### **Project Info:**

SAP data migration initiative. Led strategy and execution for migrating SAP data estates to Snowflake, Azure, and AWS.

#### **Responsibilities:**

- Led end-to-end strategy for SAP data migration to Snowflake, Azure, and AWS
- Directed ETL processes to ensure secure and complete data migration across platforms
- Guided integration of SAP with Azure for seamless data synchronisation
- Promoted cross-team collaboration between SAP and Azure teams to accelerate delivery

#### **Software and technologies:**

Databricks, Synapse, Data Factory, Data Lake, Blob Storage, SQL Pools, Logic Apps, Python, PySpark, SSMS, Snowflake, SAP, REST APIs

## **BI Analyst**

**July 2019 – December 2021** | Colombia

#### **Project Info:**

Taxes Department. Built ETL pipelines to consolidate data from Teradata, SAP, and reporting tools.

#### **Responsibilities:**

- Developed ETL processes in Databricks to integrate data from Teradata with SAP and Power BI reporting

#### **Software and technologies:**

Databricks, Data Factory, Data Lake, SQL, Python, PySpark, SSMS, SAP, Power BI

## **BI Intern**

**July 2018 – January 2019** | Colombia

**Project Info:**

Mazda Motors — CX and BI Marketing Department. Data analytics solutions to enhance customer experience.

**Responsibilities:**

- Developed and implemented data analytics solutions using QlikView, Python, and SQL to support CX initiatives

**Software and technologies:**

QlikView, Python, SQL, Power BI, SAP

## Education

### 2014 – 2019

National University of Colombia, Bogotá

Bachelor's Degree in Industrial Engineering, Systems and Industrial Department

## Certifications

- Azure Databricks and Spark SQL, Python, PySpark — Udemy (Dec 2023)
- Big Data Analytics with Hadoop and Apache Spark — LinkedIn / NASBA (Dec 2023)
- Azure Data Factory — LinkedIn (Nov 2023)
- TOEFL Equivalent — English B2 International Certification — Alcaldía de Medellín (Jan 2023)
- Django Practical Course — Udemy (Nov 2022)
- Data Analysis and Visualisation with Python: Pandas and Matplotlib — Udemy (Aug 2022)
- Python — HackerRank & Sololearn (2022)
- SQL Basic, Intermediate, Advanced — HackerRank (2021–2022)
- Teradata Advanced SQL / Fundamentals — Teradata University (2021)
- Scrum: Agile with Atlassian Jira — Coursera (Mar 2021)
- SQL for Data Science — Coursera (Aug 2020)

## Languages

English: Advanced

Spanish: Native