Javier Garcia Garcia
Data Engineer

Murcia, Spain · linkedin.com/in/javi-gg · github.com/drtey · javiggv99@gmail.com

---

I am a proactive, self-taught professional with a strong background in data engineering and software development, currently aiming to specialize in Medical Image Computing. My technical expertise includes Python, Scala, SQL, Spark, Git and AWS. I have hands-on experience in developing ETL, building data warehouses and data lakes on cloud platforms.

Main stack: Python, Scala, SQL, PySpark, Spark, Git, Jenkins, AWS, Snowflake, Pytorch, Scikit-Learn, Numpy, OpenCV, CUDA C/C++.

PROFESSIONAL EXPERIENCE

Logicalis   **Madrid, Spain**
Data Engineer   *Mar 2024 – Present*

External Data Engineer at Atresmedia
- Perform new development, evolutionary, and corrective maintenance of the ETL platform in AWS, utilizing Scala and Spark for data processing and partitioning to optimize performance and scalability.
- Design and implement Batch processes using Lambda to ingest data efficiently, ensuring scalability and reliability in data pipelines.
- Develop and deploy Lambda functions to automate the login and data retrieval processes from various social media platforms, including TikTok Ads Business, Meta Ads Business, and Google Ads.
- Ingest and parse requests and responses from a machine learning recommendation model hosted on AWS, capturing model scores via logs.
- Optimize data pipelines and ensure the seamless flow of data between different AWS services and other products like Snowflake.

Stack: Python, Scala, SQL, PySpark, Spark, Jupyter Notebooks, Jenkins, Git, Snowflake.
AWS: S3, Redshift, PostgreSQL, DynamoDB, Athena, Kinesis Firehose, Elasticache, Databrew, Glue, Step Function, CodeCommit, CloudFormation, SageMaker, SNS, SQS, Eventbridge, EMR, Fargate, EC2.

Nubika Cloud Solutions   **Murcia, Spain**
Data Engineer   *Nov 2022–Mar 2024*

- Development of datasets, and dataflows on Salesforce Einstein Analytics (SAQL, SQL).
- Building data warehouse and data lake using MuleSoft on Amazon S3 and Amazon Redshift services.

Sisytec Networks   **Murcia, Spain**
Software Developer   *Sep 2021 - Nov 2022*

- Development of API, JWT authentications and user interface for web environments (PHP, Node.js, Git, JS, Angular, JQuery, CSS, SQL).

Chest CT Scan Analysis - **Personal Project**
- Built and implemented a convolutional neural network model with Python and PyTorch to automatically classify CT scan images, reducing the diagnostic workload for healthcare workers.

- Used an optimized ResNet-50 architecture, pretrained on ImageNet, to enhance the accuracy of detecting pulmonary conditions.
- Applied preprocessing techniques such as resizing, normalization and data augmentation to prepare CT images.
- Employed techniques like dropout and early stopping to prevent overfitting and used evaluation metrics such as accuracy, recall, F1-score and confusion matrix to validate model performance.

EDUCATION

UNIVERSITAT OBERTA de CATALUNYA   **Barcelona, Spain**
Data Science Degree   *Feb 2022 - Sep 2026*

IES JOSE PLANES   **Murcia, Spain**
Web Applications Development Associate Degree (CFGS)   *Sep 2019 - Dec 2021*

SKILLS

- Certifications: [Microsoft Azure AI 900](#), [Databricks Generative AI Fundamentals](#) and [Databricks Lakehouse Fundamentals](#).
- Native Spanish and B2 English level.