

Andrés Menor



Data Engineer · HPC & Cloud Infrastructure

andresmr997@gmail.com · 633 60 78 03 · andresmenor.com
linkedin.com/in/andres-menor · Catalonia / Hybrid

PROFILE

Data Engineer with experience in cloud-native pipelines, data workflow orchestration, and distributed computing in HPC environments. Has designed and implemented production RAG systems on GCP, executed large-scale workloads on HPC clusters using SLURM, and built end-to-end ETL pipelines. Strong mathematical foundation (GPA 9/10, honours in Probability, Linear Algebra, and ML). Seeking a role with real technical ownership in a data team.

EXPERIENCE

Data Engineer · **Findspo (startup)** 04/2026 – Present

- Designed and implemented a full ingestion pipeline from PubMed (Python) for a longevity research platform grounded in scientific evidence.
- Built a RAG architecture: semantic search with vector embeddings + LLM reranking for a clinical evidence-based chatbot.
- Managed all GCP infrastructure via Terraform: Cloud Storage, BigQuery, Cloud Run, and Vertex AI.
- Sole data engineer responsible for pipeline reliability, data quality monitoring, and iterative improvement of the retrieval system.

Data Scientist Intern · **CSIC – IBBM (Institut de Biologia Molecular de Barcelona)** 09/2025 – 12/2025

- Processed large-scale scRNA-seq datasets using Python and scanpy for biological age prediction.
- Trained and evaluated regression and classification models (Elastic Net, XGBoost, Random Forest, SVM); best result with Logistic Elastic Net.
- Executed all computational workloads on the IBBM HPC cluster using SLURM.
- Built reproducible ML experimentation workflows with automated preprocessing, feature selection, and model evaluation.

Junior Data & BI Analyst · **Grup Boadella** 10/2025 – 03/2026

- Automated SQL reporting workflows and Python preprocessing pipelines with data validation and quality control.
- Designed data transformation processes to support analytical models and business KPIs.

TECHNICAL SKILLS

Languages	Python, SQL (PostgreSQL), R, Java
Data Eng.	ETL/ELT, Airflow, dbt, data validation, RAG, semantic search, reranking
Cloud & Infra	GCP (BigQuery, Cloud Storage, Cloud Run, Vertex AI), Terraform, Docker, Linux, Git
HPC	SLURM, distributed workload execution on HPC clusters
ML & Analytics	scikit-learn, XGBoost, Elastic Net, Random Forest, Pandas, feature engineering, scRNA-seq

EDUCATION

BSc Data Science — Universitat Oberta de Catalunya (UOC)

2023 – 2027

GPA: 9.0 · Honours: Probability & Linear Algebra · Distinction: Machine Learning, Bayesian Modelling

LANGUAGES

Spanish – Native · Catalan – Native · English – Fluent (professional)