

Pau Alcázar Perdomo

DATA & AI ENGINEER | PYTHON, PYSPARK, AWS & AZURE

Barcelona, Spain · Open to relocation: Geneva / Vaud, Switzerland
github.com/drpaau · LinkedIn

PROFESSIONAL SUMMARY

Data & AI Engineer and consultant with 3+ years of experience delivering end-to-end data, AI and cloud solutions for public-sector and healthcare organisations. Strong hands-on background in Python, PySpark, SQL, FastAPI, AWS and Azure, spanning ETL/data integration, REST APIs, LLM and RAG workflows, document intelligence, BI and production deployment. Experienced in translating client requirements into secure, observable and maintainable solutions and collaborating across engineering, architecture, functional and stakeholder teams.

CORE TECHNICAL SKILLS

Languages & data engineering Python, PySpark, SQL, SAS, JavaScript/TypeScript, Bash, R, ETL/ELT, Data modelling, Data quality, Batch processing.

AWS Glue, Athena, S3, Lambda, API Gateway, CloudFront, Bedrock, QuickSight.

Azure & platforms Azure OpenAI, AI Search, Document Intelligence, Blob Storage, Cosmos DB, Key Vault, ARO/OpenShift, PostgreSQL, PostGIS, Delta Lake, Redis/RQ.

AI, backend & DevOps LLM/RAG, Prompt evaluation, Text-to-SQL, Document processing, PII redaction, FastAPI, REST APIs, Asynchronous workers, Docker, Kubernetes/OpenShift, Jenkins CI/CD, Git, Linux.

Web & analytics Angular, React/Vite, D3/TopoJSON, Power BI, QuickSight, Excel.

PROFESSIONAL EXPERIENCE

Accenture

Jun 2023 – Present

Data & AI Analyst (AI Decision Science) · Barcelona, Spain

Deliver end-to-end data, AI and cloud solutions for public-sector and healthcare clients, from discovery and architecture through implementation, testing, deployment and production troubleshooting.

- Design and implement generative AI applications and reference architectures for public-service use cases with Python/FastAPI, Angular/React, Azure services, Docker and OpenShift.
- Build and industrialise data pipelines and APIs across AWS and Azure, including PySpark/Glue ETL, Athena-based analytics, REST integrations, asynchronous processing and serverless delivery patterns.
- Apply data governance and engineering controls covering data quality, traceability, consistency, PII handling, testing, observability and operational ownership.
- Collaborate directly with clients and multidisciplinary teams to turn functional needs into technical requirements, implementation plans and maintainable production solutions.

Superprof

Sep 2021 – Present

Technology Subjects Tutor · Barcelona, Spain

Teaching programming, algorithms, operating systems, mathematics and physics.

- Adapt technical explanations and practical exercises to different student profiles.

EDUCATION

Bachelor of Science in Computer Engineering

2020 – 2024

Universitat Politècnica de Catalunya (UPC) – EPSEVG

Software engineering, algorithms, databases, operating systems, networks, data science, cybersecurity and parallel programming.

LANGUAGES

Catalan · Spanish · English

SELECTED PROJECT PORTFOLIO

AI-assisted document ingestion

Designed an asynchronous workflow that extracts and validates PDFs, generates metadata with GPT-4o, supports human review, maps metadata to a controlled vocabulary and publishes records through REST APIs. Implemented queue-based and no-Redis processing alternatives for enterprise deployment constraints.

Technology: Python, FastAPI, Angular, GPT-4o, DSpace REST APIs, controlled vocabularies, Redis/RQ, Azure Blob Storage, Key Vault, OIDC, OpenShift.

Production cloud architecture & deployment

Designed a containerised application platform with a single-page web app, API backend and Python worker for a managed OpenShift environment. Defined environment-specific integrations, API exposure, identity, secrets, observability, CI/CD and asynchronous processing alternatives for production readiness.

Technology: Angular, FastAPI, Python, Docker, ARO/OpenShift, Jenkins, OIDC, API gateway, Key Vault, Blob Storage, Application Insights, Log Analytics.

Healthcare data engineering & GenAI

Developed PySpark and AWS Glue ETL processes across Oracle, S3/Delta Lake and Athena; supported healthcare reporting, clinical form completion, data preparation and batch LLM workflows with Amazon Bedrock.

Technology: Python, PySpark, AWS Glue, S3, Delta Lake, Athena, Amazon Bedrock, Oracle, QuickSight.

Multi-agent mobility assistant

Contributed to a mobility assistant combining natural-language responses, map output and deterministic data queries. Worked across text-to-SQL, PostGIS/GeoJSON, ETL and serving layers, prompt evaluation and production diagnosis of environment-specific data-refresh failures.

Technology: Python, FastAPI, PostgreSQL/PostGIS, GeoJSON, SQL templates, ETL, prompt evaluation, OpenShift.

PII anonymisation platform

Worked on an asynchronous document-anonymisation service using hexagonal architecture and explicit workflow states, with secure storage, secrets management, PII redaction and automated tests designed for enterprise deployment.

Technology: Python, FastAPI, Azure Cosmos DB, Blob Storage, Key Vault, Document Intelligence, Docker, automated testing.

Serverless analytics delivery

Built an Athena–Lambda–API Gateway–CloudFront–S3 delivery layer for a custom BI visual, implementing paginated retrieval beyond default row limits and serving more than 32,000 records to the client application.

Technology: Python, Amazon Athena, Lambda, API Gateway, CloudFront, S3, QuickSight, JavaScript.

Road-safety analytics

Produced official quarterly, campaign, thematic and provincial analyses using SAS/SQL, reusable reporting macros, quality controls for provisional and consolidated records, geographic and population enrichment, and validated office-document outputs.

Technology: SAS, SQL, Excel, Word, PowerPoint, RTF, data validation, reporting automation.

Camí Curt — independent web product

Built and deployed an educational route-planning game for Catalonia using Dijkstra/A* pathfinding. The application covers 41 comarques and 1,550 route constraints, with daily, standard, timed and exploration modes.

Technology: React 18, Vite, JavaScript, D3, TopoJSON, Web Workers, Vercel, pathfinding algorithms.

camicurt.cat