

MARK JACKSON

Bath, UK | mcjacks@pm.me

PROFILE

Senior Data Engineer with 10+ years of experience building and scaling data platforms in growing organisations. Strong track record of establishing data architecture from the ground up, modelling trusted data with dbt and enabling self-service analytics across technical and commercial teams. Combines hands-on Python, SQL and AWS expertise with pragmatic technical leadership, mentoring and a focus on reliable, maintainable delivery. Experienced in growth analytics and the practical application of generative AI, RAG and recommendation systems.

TECHNICAL EXPERTISE

- **Data engineering:** Python, SQL, dbt, Airflow, Dagster, ETL/ELT, data modelling, Apache Spark, Pandas
- **Platforms:** AWS, S3, Apache Iceberg, Athena, Redshift, PostgreSQL, MySQL, BigQuery, MongoDB, DynamoDB
- **Analytics:** QuickSight, Looker, Metabase, self-service analytics, KPI reporting, acquisition funnels, conversion and clickstream analysis
- **AI and machine learning:** LLM APIs, LangChain, RAG, vector and multimodal embeddings, recommendation systems
- **Engineering and leadership:** Docker, TypeScript, CI/CD, automated testing, data quality, technical leadership, mentoring, stakeholder collaboration

PROFESSIONAL EXPERIENCE

Senior Data Engineer | KYND

Aug 2025 – Present

- Joined the cybersecurity analytics company as its first dedicated data hire, turning a raw bronze data lake into a trusted medallion architecture for historical analysis and customer-facing data products.
- Selected and implemented a deliberately simple AWS stack using S3, Apache Iceberg, Athena and dbt Core, orchestrated by Dagster on ECS and scheduled with Step Functions.
- Unblocked a long-running embedded-dashboard initiative by championing a pragmatic design using QuickSight SPICE; delivered three production dashboards backed by around 20 dbt models within the first year.
- Replaced monthly Excel reporting for around 30 clients, representing approximately 90% of the business, saving one person-week per month while helping retain clients and support £500k-£1m in new annual recurring revenue.
- Owned all transformations and QuickSight dataset implementation, working with Operations, Engineering, Marketing and Sales to agree consistent definitions and reusable data models.
- Established dbt modelling, documentation, unit-testing, data-testing and review standards. Added source-system reconciliation, client-level row-level security and Slack alerts for pipeline failures.
- Centralised complex versioning and deduplication logic into tested current-state and historical-change models, removing the need for consumers to repeatedly reconstruct entity state.
- Mentored four Operations colleagues to develop their data skills and progress towards data analyst and analytics engineer roles. Promoted AI-assisted questionnaire extraction, saving 5-10 hours per month.

Technologies: SQL, dbt Core, Dagster, AWS, S3, Apache Iceberg, Athena, ECS, Step Functions, QuickSight, SPICE, Kafka, Git

Senior Data Engineer | Validis

Dec 2024 – Apr 2025

- Established testing, Agile and development practices for a new team while mentoring junior engineers.

Senior Data Analytics Engineer | Seccl*Dec 2022 – Dec 2024*

- Designed and built a data platform for financial transactions covering more than 100,000 clients, establishing the ingestion and warehouse architecture from the ground up.
- Modelled a new Redshift warehouse using dbt, creating a scalable foundation for operational, marketing and regulatory reporting.
- Championed data-quality testing and led improvements to deployment and testing practices.
- Partnered with teams across the business to turn varied requirements into reusable models, Looker reporting and a cohesive data roadmap.
- Promoted self-service analytics and mentored junior engineers in data technologies and engineering practices.

Technologies: Python, dbt, Airflow, SQL, Redshift, MongoDB, Docker, Looker, AWS, Fivetran**Senior Data Engineer / Data Lead | MoveGB***Mar 2018 – Dec 2022*

- Designed and built the company's data warehouse and data lake from the ground up, integrating operational databases, clickstream data and third-party platforms.
- Partnered closely with the Acquisition team to analyse user-acquisition performance, conversion funnels and clickstream behaviour, helping identify drop-off points and growth opportunities.
- Created datasets, dashboards and KPI reporting that enabled teams to independently explore performance and make data-informed decisions.
- Designed and deployed a collaborative-filtering recommendation system for a customer-facing content feed.
- Built external partner reporting and supported the CEO with M&A due diligence and investor presentations.
- Led data team members and an external data-science team, balancing stakeholder priorities, platform performance and infrastructure cost while promoting self-service data practices.

Technologies: Python, Airflow, SQL, Redshift, PostgreSQL, MySQL, TypeScript, Docker, Metabase, AWS, Apache Spark, Segment, HubSpot**Data Engineer | OVO Energy***Dec 2017 – Mar 2018*

- Built ETL pipelines for new warehouse sources and supported BI users with analytical tables, views and APIs.
- Connected data producers and consumers to promote dependable delivery and clear data ownership.

Senior Software Developer | Genius Digital / Verimatrix*Mar 2014 – Dec 2017*

- Designed scalable ETL pipelines and maintained warehouse infrastructure for TV analytics.
- Redesigning an Apache Spark analytics platform, reducing client AWS costs by 90% and cutting daily processing time from six hours to 30 minutes.
- Developed a proof of concept for real-time TV IoT processing and an automated cloud-based testing framework.
- Led cross-platform client-library development, mentored engineers and delivered training to an international team.

Software Developer | Vero Software*Oct 2013 – Mar 2014*

- Developed CAD/CAM software using C, C++ and C#.

SELECTED PROJECTS**Buzforge — AI-assisted social content platform | 2025 - Present**

- Developing a serverless social platform with AI-assisted content creation, multi-provider LLM integration and token-efficient prompting.
- Built multimodal embedding capabilities for image retrieval using TypeScript, AWS Lambda and Aurora PostgreSQL.

Cricket Conversational AI Assistant | 2023

- Built a Python and LangChain assistant combining live match data with RAG, deployed on serverless AWS infrastructure.

EDUCATION**University of Bath | 2010 - 2013**

BSc (Hons) Mathematics and Physics — 2:1