

Ammar Hamdi Hammouch

Data Analyst

✉ ahlan@ahammouch.me ☎ 055 498 4459 📍 Riyadh 🔗 linkedin.com/in/ammam-hammouch 🐙 github.com/AmmarsGeeks

Professional Summary

Builds the pipelines and warehouses that reporting and dashboards actually run on. Built a dbt + DuckDB medallion warehouse (bronze/silver/gold) feeding Metabase, joining evaluation history with operational telemetry across cost, quality, and ops marts. At Thmanyah, automated performance reporting from Tally, Airtable, Forms, and Excel into a CRM, cutting the manual cycle by ~60%. At Reachware, built Snowflake pipelines for financial reporting.

Education

Bachelor in Computer Science, Jazan University

2021 – 2025

GPA: 4.86 / 5 — First Class Honours

Professional Experience

Software Engineer, Entropy

09/2025 – Present

- Built a dataset quality pipeline that validates CSV, XML, and JSON uploads — format checks, metadata validation, and corrupt-file detection — stopping bad data before it reaches downstream annotation and evaluation systems, orchestrated on BullMQ + Redis.
- Contributing to dataset management, annotation pipelines, and model-scoring infrastructure for an Arabic-language LLM evaluation platform.

Integration Specialist, Thmanyah

01/2025 – 05/2025

- Designed a multi-source aggregation engine pulling from Tally, Google Forms, Airtable, and Excel into an internal CRM. Configurable flows similar to n8n, auto-distributed performance reports, and roughly 60% less manual reporting time.
- Built an internal HR platform covering payroll, notice periods, and 3-month probation evaluations — PDF report generation via Puppeteer, job queuing on BullMQ + Redis, and heavy schema migration work on Drizzle ORM.

React Native Developer, Motmaina Medical Center

04/2024 – 12/2024

Built the admin dashboard in Next.js with React Query (hydration, interceptors, SSR, query invalidation) for real-time visibility into sessions, providers, and clinical activity.

Technology Integration Intern, Reachware

08/2023 – 04/2024

- Built data integration pipelines between Snowflake and ServeMe for financial reporting and analytics, handling schema mapping and scheduled ETL across the two systems.
- Contributed to Oracle NetSuite and Oracle Fusion ERP integration architecture, focused on finance module sync between the two systems.
- Worked on ZATCA e-Invoice Phase 2 (FATOORA) compliance integration: XML schema validation for invoice data feeding downstream reporting.

Enterprise Application Developer, COMT

08/2022 – 08/2023

- Built a business intelligence layer with cloud-synced analytics dashboards, inventory management architecture, and session handling to support operational decision-making.
- Architected the inventory and product-catalog data model underlying a full POS transaction cycle, from order entry through payment and receipt.

Projects

Balsam, LLM Evaluation Pipeline

Batch LLM evaluation platform running an 11-task Airflow DAG: dataset validation, Spark ETL for preprocessing, human-in-the-loop annotation routing, and three scoring tiers — classical metrics, semantic similarity, and LLM-as-a-Judge — producing ranked leaderboards. The pipeline ingests datasets through format validation and metadata quality checks before they reach the evaluation stage, with schema contracts enforced in CI so downstream dashboards never see broken data.

Basira, dbt + DuckDB Medallion Warehouse

Medallion-architecture warehouse (bronze/silver/gold) built on dbt and DuckDB, feeding Metabase dashboards. Joins LLM evaluation history from Balsam with operational telemetry from Mizan across three marts: cost (token usage, compute hours), quality (scoring accuracy, annotator agreement), and ops (pipeline latency, throughput). Schema contracts are enforced in CI — a breaking upstream change fails at the boundary rather than silently corrupting a downstream dashboard. The warehouse layer turns raw evaluation logs and streaming metrics into queryable business views that the platform team uses daily.

Mizan, Streaming Metrics and Anomaly Detection

Real-time observability service on Kafka and Faust, streaming evaluation metrics with EWMA and z-score anomaly detection. Anomalies surface as dashboard alerts when scoring drift or pipeline latency crosses thresholds. Sinks to TimescaleDB for time-series querying and Apache Iceberg for long-term analytical storage, making historical trend analysis possible across months of evaluation runs.

Flow Studio 📄, Webhook Aggregation and Reporting Pipeline

Webhook-driven platform that aggregates form submissions from Google Forms, Tally, and Typeform, runs each through a configurable DAG pipeline, and generates formatted output reports (React-PDF) routed to Gmail, a CRM, or a webhook. Orchestration runs entirely on Postgres — no separate scheduler or coordinator — with the next pipeline step always derived from persisted state, so a crashed worker never leaves a report half-generated or a fan-in step double-counted. The same aggregation-into-formatted-output pattern used in the Thmanyah reporting pipeline, generalized into a reusable DAG-based reporting engine.

Stock Sentiment Analysis and Ranking System 📄

MLOps pipeline for financial news sentiment analysis — ingests news data, scores sentiment through NLP models, and ranks stocks by market impact. Full pipeline from ingestion through inference to ranked output, deployed as a live demo on HuggingFace Spaces.

Skills

Languages

Python, SQL, TypeScript, JavaScript

Infrastructure

Airflow, Redis, BullMQ, Docker, Webhooks

Cloud

GCP (Cloud Run, Firebase), AWS (S3, Lambda)

Certifications

Oracle Autonomous Database Cloud Certified Professional —

Issuer: Oracle | Date: November 2025

ML Engineering and MLOps Training — Issuer: Qafza | Date:

March 2025

AWS Cloud Practitioner CLF-C02 — Issuer: AWS

Data and Analytics

dbt, DuckDB, Metabase, Snowflake, PostgreSQL, data pipeline, ETL, reporting

Frameworks

FastAPI, Next.js, React Native

Languages

Arabic — Native | English — Professional