

# Omar Essam El Sherbiny

Head of Data • Data Engineering, Cloud Platform & Machine Learning

Zahraa El Maadi, Cairo, Egypt  omaessam.web.app

 +20 1004819774

 eng.omar.e@gmail.com

 HuggingFace

 Credly

 LinkedIn

 GitHub

 dev.to

 Skillshop

 Coursera

 Kaggle

 MS Learn

 GC Skills

## PROFESSIONAL SUMMARY

Head of Data with 8+ years of technical experience across data engineering, cloud platform ownership and applied machine learning — work that began inside a civil-engineering career, where automating project controls and reporting became the route into data full time. Currently own an enterprise data platform end to end on Google Cloud: the warehouse, pipelines and production databases, the BI portal on top of them, and the CI/CD that ships it all. Equally at home tuning a production database, cutting cloud spend structurally, putting validated machine learning in front of the teams that act on it, and reporting the results to the CTO.

## CORE COMPETENCIES

### Cloud & Platform Engineering

- GCP (BigQuery, Cloud Run, Cloud SQL, Cloud Build, Cloud Functions, Scheduler, Cloud Storage, Vertex AI, Compute Engine, Load Balancing, VPC, IAM, IAP)
- AWS (EC2, S3, Lambda, SageMaker, Redshift), Azure (ML, Data Factory, Synapse)
- Docker, Kubernetes, Terraform, systemd, Nginx, Cloudflare
- CI/CD: GitHub Actions, Cloud Build, SonarCloud, CodeQL

### Machine Learning & AI

- PyTorch, TensorFlow, Scikit-learn, XGBoost, SHAP
- LLMs, RAG, Vector Databases, Vertex AI, LangChain
- MLOps, probability calibration, walk-forward validation, deployment
- Agentic AI engineering (Claude Code, Cursor, Windsurf)

### Data Engineering & Analytics

- ETL/ELT: Apache Airflow, dbt, Spark, Kafka; Databricks, Snowflake
- Warehousing: BigQuery, dimensional modelling, data quality SLAs
- PostgreSQL, MySQL, Oracle, Redis, MongoDB: indexing, partitioning, tuning
- Full-stack BI: Python/FastAPI, React/TypeScript, SQL; Power BI, Tableau

### Leadership & Project Management

- Agile/Scrum/Kanban; Jira, Confluence, GitHub
- Team leadership, mentoring, stakeholder management
- CTO-level reporting, RFC & architecture review
- Data governance, risk assessment, strategic planning

## PROFESSIONAL EXPERIENCE

### DATA ENGINEERING, PLATFORM & MACHINE LEARNING LEADERSHIP

#### Head of Data | WalaPlus

Jul 2026 – Present

*End-to-end ownership of data, databases and BI on GCP, and of the application teams' backend and cloud decisions (VP of Data / Chief Data Officer scope)*

- Own the enterprise data platform on Google Cloud — BigQuery warehouse, Cloud SQL (PostgreSQL/MySQL), Redis and MongoDB Atlas — and the BI portal serving Finance, FinTech, Sales, Marketplace, Marketing and Customer Success; lead six data engineers and analysts, own backend, database and GCP decisions for the application teams, and report monthly to the CTO.
- Led the company-wide migration of the warehouse from a legacy cloud vendor onto BigQuery to completion: every dashboard and service re-pointed, the history archived and verified byte for byte, the legacy code removed and the vendor retired.
- Run platform DevOps: moved source control to GitHub and deploys to Cloud Build over an IAP tunnel with fast, gated releases; own the PR gates (TypeScript, ruff, Playwright, SonarCloud, CodeQL) and cleared the dependency and code-quality alert backlog.
- Extended ownership to every production database: audited every production estate into an owned remediation programme, verify proposed index drops against live usage counters before they ship, partitioned and clustered the hot BigQuery tables at load time to cut scan costs, and issue a daily database health report to the CTO and DevOps.
- Put the churn model (AUC 0.929, 21× lift, validated out of sample against what actually happened) in front of Customer Success as a live at-risk list, and run an AI-assisted engineering practice — parallel agentic sessions under written guardrails and review gates — with every change traceable in Jira and Confluence.



## Staff Data Engineer | WalaPlus

Oct 2025 – Jun 2026

*Built the BI portal, the finance data model and the migration programme that moved the company onto GCP*

- Built the BI portal from the ground up — FastAPI/Python services, React & TypeScript frontend, Redis caching with nightly preload, Google OAuth with group-to-category authorization, user management and view-as-user — replacing scattered reporting tools for Finance, FinTech, Marketplace, GRC, Customer Success and Management, with Excel/PDF export on every dashboard.
- Delivered the finance and FinTech reporting suite — ledger summaries, multi-way reconciliation, revenue, unit-economics and cost reporting and executive overviews — with every figure verified at source before release.
- Replaced dozens of fragmented bronze views with a normalized star schema as the single source of truth for all finance dashboards, refreshed nightly, and closed the reconciliation gaps it exposed.
- Built reporting assurance: every published client metric is re-derived independently before release, and partner settlements are reconciled against the partners' own records.
- Ran two platform migrations from legacy MySQL to Laravel PostgreSQL on Cloud SQL: the marketplace seller-platform rebuild (production cutover with zero errors, backed by a verification suite and a portal section tracking every issue and answer) and the core customer platform (per-client rolling cutover with balances reconciled 1:1).
- Designed the warehouse-to-GCP path: a resumable bulk-mirror engine into Cloud SQL with atomic per-table swaps, then direct nightly BigQuery loads with parity checks, legacy fallbacks and a live cutover registry; ran the CTO database workshops that approved the cost-reduction plan and the data-asset naming standard.
- Built the client churn model end to end (point-in-time panel, walk-forward XGBoost with SHAP, calibration, four model dashboards); hardened the portal's hosting; shipped a company-wide strategy KPI app behind portal SSO.

## AI Researcher (part-time) | Momita.ai (formerly Gravigol)

Jul 2025 – Present

*Part-time second role: AI Team Lead on the maternal-care nutrition chatbot (Jul 2025 – Mar 2026), then research and model improvement*

- Led maternal-care nutrition chatbot on GCP Vertex AI; implemented RAG and fine-tuning with safety, traceability, and responsible recommendations.
- Built a Firebase-hosted frontend web portal around the chatbot for browser users, integrating authentication, session telemetry, and release workflows.
- Logged interactions to Firestore, BigQuery and Cloud SQL for auditing and analytics, integrated global food-statistics APIs for calorie and vitamin calculations, and ran evaluation and feedback loops to improve recommendation quality.
- Since April 2026: rebuilt the agent so Arabic is a first-class language rather than a translation layer, added a second LLM beside Gemini with per-model ratings and a bilingual speed, token and cost comparison, and cut typical answer time by a third and model cost by half.

## Data Engineering And Machine Learning Lead | Rowad Modern Engineering

Jan 2025 – Sep 2025

*Leading ML/AI initiatives and advanced analytics solutions*

- Architected and deployed a production RAG chatbot on vector databases (Pinecone, Chroma) and LangChain, improving internal knowledge access by 60%
- Built a full-stack analytics platform linking Oracle ERP to the on-premises warehouse for real-time BI
- Built deep learning models in PyTorch and TensorFlow: CNNs for image classification, transformers for NLP
- Led the cross-functional data and AI team of 10 using Agile/Scrum, achieving a 95% sprint completion rate and reducing development cycles by 40%
- Built GPU-accelerated (NVIDIA 4090) AI agent tooling that cut manual debugging time by 60%, with MLOps on Kubernetes, MLflow and GitLab CI/Jenkins for automated model deployment and monitoring

## Data Engineering Lead | Rowad Modern Engineering

Nov 2022 – Dec 2024

*Scaling data infrastructure and leading engineering teams*

- Established and led the data team — 6 data engineers and analysts, 2 ML engineers and 2 technical-office engineers — with Scrum for project tracking and sprints
- Designed and deployed enterprise-scale data pipelines across AWS (S3, Redshift, Lambda), Azure (Data Factory, Synapse), and GCP (BigQuery, Dataflow)



- Built real-time streaming data solutions using Apache Kafka and Spark, processing 10M+ records daily with 99.9% uptime
- Architected cloud-native data warehouse solutions using Snowflake and Databricks, reducing query response times by 75%

#### Data Scientist & Scrum Master | Rowad Modern Engineering

May 2022 – Oct 2022

*Advanced analytics, predictive modeling, and agile team facilitation*

- Developed cash-flow prediction models using ensemble methods (XGBoost, Random Forest) and time-series analysis, with automated feature-engineering pipelines that lifted model accuracy by 15%
- Facilitated Scrum ceremonies and coached cross-functional team of 6 members in Agile methodologies, improving sprint velocity by 25%

#### Senior Data Engineer & DBA | Rowad Modern Engineering

Nov 2020 – Apr 2022

*Infrastructure automation, optimization, and database administration*

- Automated complex business processes using Python and SQL, saving 200+ hours monthly through intelligent workflow optimization
- Integrated data from multiple enterprise sources into centralized data warehouse including Oracle ERP, Oracle Unifier, CRM, Primavera P6 EPPM, and HR Hits systems
- Managed database administration including user management, configuration optimization, table design, and performance monitoring across production systems

#### Data Engineer | Rowad Modern Engineering

Jan 2020 – Oct 2020

*Infrastructure setup and data pipeline development*

- Established on-premises server infrastructure and configured Apache Airflow and Spark for distributed data processing
- Developed PL/SQL ETL processes loading Oracle ERP data into the warehouse, with database optimization and backup procedures for critical business data

## CONSTRUCTION & CIVIL ENGINEERING LEADERSHIP

- **Projects Control Manager** | Rowad Modern Engineering, Bridges Sector (Jan 2019 – Dec 2019)
- **Senior Project Control Engineer** | Rowad Modern Engineering, Head Office (Aug 2018 – Dec 2018)
- **Project Control Engineer** | Rowad Modern Engineering (Mar 2018 – Jul 2018)
- **Planner** | Concrete Plus (Apr 2015 – Mar 2018) - Multi-site project baseline development
- **Planner, QS Engineer** | Concrete Plus (Jan 2015 – Mar 2015) • **Technical Office, QS Engineer** (Aug 2014 – Dec 2014)
- **Site Engineer** | Concrete Plus, Porto October Project (May 2014 – Jul 2014) • **Site Engineer Training** | Orascom Construction Industries (Mar 2014 – May 2014)

## EDUCATION

**Cairo University** | Bachelor of Science in Civil Engineering | May 2013

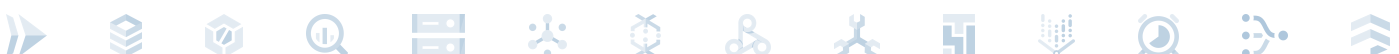
*Graduation Project:* Steel Bridges Design, graded Distinction (A)

**Orouba Language School** | High School Diploma (Thanaweya Amma, 2005) | 97.9% (401.5/410)

## ADVANCED TECHNICAL SKILLS

**Cloud & Platform:** GCP (BigQuery, Cloud Run, Cloud SQL, Cloud Build, Cloud Functions, Cloud Scheduler, Pub/Sub, Cloud Storage, Datastream, Vertex AI, Model Garden, Compute Engine, Load Balancing, VPC, IAM, IAP, Secret Manager, Cloud Logging & Monitoring, Firestore, Firebase), AWS (SageMaker, Bedrock, Lambda, S3, Redshift, RDS), Azure (ML Studio, Synapse, OpenAI), Databricks (Unity Catalog, Delta Live Tables), Kubernetes, Docker, Terraform, systemd, Nginx, CUDA

**Databases & Optimization:** PostgreSQL, MySQL, SQL Server, Oracle (PL/SQL, ERP, Unifier), BigQuery, Snowflake, Redis, MongoDB Atlas, vector databases; partitioning & clustering, index and query tuning, execution-plan



analysis, replication & WAL management, retention and archival design, storage & cost governance

**Data Engineering & Analytics:** Apache Spark, Hadoop, Airflow, dbt, Kafka, Delta Lake, NDJSON/Parquet bulk loading, change-data capture, dimensional modelling, data quality SLAs, Statistical Methods (ANOVA, t-tests, Regression), Model Evaluation (Cross-validation, KFold, walk-forward), Feature Engineering, Power BI, Tableau

**AI & Machine Learning:** PyTorch, TensorFlow, Scikit-learn, XGBoost, SHAP, probability calibration, Transformers (Hugging Face, ONNX), LangChain, LlamaIndex, PEFT, RLHF/PPO/DPO, RAG, SFTTrainer, SkyPilot, Computer Vision (CNNs, YOLO), NLP (BERT, GPT, T5), Reinforcement Learning, Recommendation Systems, agentic AI development (Claude Code, Cursor, Windsurf)

**Development & Operations:** Python, SQL, TypeScript/React, R, Java, C++, FastAPI, Vite, Tailwind, pytest, Playwright, MLflow, GitHub Actions, Cloud Build, Jenkins, GitLab CI, SonarCloud, CodeQL, Dependabot, Ansible, MPI4Py

## LEADERSHIP & PROJECT MANAGEMENT

**Methodologies:** Agile, Scrum, Lean, Kanban, SAFe      **Tools:** Jira, Confluence, GitHub, Taiga, OpenProject, MS Project, Primavera P6

**Expertise:** Team Leadership, Data Governance & Ownership Models, Executive & CTO-Level Reporting, RFC and Architecture Review, Cross-Team Stakeholder Management, Risk Assessment, Strategic & Capacity Planning, Mentoring and Technical Onboarding

## CERTIFICATIONS & PROFESSIONAL DEVELOPMENT

**Data Engineering & Architecture:** Snowflake Data Engineering Professional Certificate (Dec 2025); IBM Data Warehouse Engineer Professional Certificate (Sep 2025); IBM Relational Database Administrator with GenAI Professional Certificate (Oct 2025); Database Design & Operational Business Intelligence Specialization (UC Irvine, Sep 2025); Generative AI for Data Engineers Specialization (IBM, Aug 2025)

**AI & Machine Learning:** Google AI Professional Certificate (Oct 2026); IBM AI Engineering Professional Certificate (Sep 2025); IBM Machine Learning Professional Certificate (Oct 2026); IBM Generative AI Engineering Professional Certificate (Sep 2025); Natural Language Processing Specialization (DeepLearning.AI, Jul 2025)

**Data Science & Analytics:** IBM Data Science Professional Certificate (Aug 2025); Google Data Analytics Professional Certificate (Oct 2025)

**Leadership, Strategy & Architecture:** Agile Leadership Specialization (University of Colorado System, Oct 2026); IBM & Darden Digital Strategy Specialization (Oct 2026); Generative AI for IT Systems Analysts & Architects Specialization (IBM, Oct 2026)

**Professional & Technical:** Certified Data Scientist Professional (Epsilon AI, 2024); CCNA preparation course (Cairo University High Tech Center, 2008); Mastering Primavera P6 (YAT Learning Solutions, 2014); Steel Bridges Design (graduation project, Cairo University)

**In progress:** Google Cloud Professional Data Engineer, Professional Cloud Architect and Professional Cloud DevOps Engineer certification preparation. **Full record:** 330+ Coursera certificates and 120+ Credly badges, each publicly verifiable at [credly.com/users/eng-omar-essam](https://credly.com/users/eng-omar-essam)

## PERSONAL PROJECTS, ACHIEVEMENTS & LANGUAGES

**Selected technical projects:** a CNN that identifies Egyptian hieroglyphs by Gardiner code (96.9% top-1 on 59 signs), published on Hugging Face with a live demo; a reproducible Wikidata-to-dataset pipeline for the pharaohs of Egypt (SPARQL, pandas, pandera validation) on Kaggle and Hugging Face; an applied-ML portfolio (churn, segmentation, image classification, recommenders) at [github.com/engomaressam](https://github.com/engomaressam).

**Music Performance:** Guitarist in Iron Maiden tribute band • **Digital Content Creation:** Built YouTube channel to 100,000+ subscribers • **Community Building:** Co-founded social media group ESS • **Media & Communication:** Co-hosted 3 podcasts • **Creative Writing:** Authored two science fiction novels • **Community Service:** Active in street animal rescue operations in Maadi • **Languages:** Arabic (Native) • English (Professional Working Proficiency, Fluent)

