

Adrian Ng, MSc.

Junior Metadata Engineer at IAG Cargo

Email: contact@adrian.ng **Portfolio:** adrian.ng **Location:** Harrow

PROFILE

*NB. I am **not** job-hunting. May this document simply serve to give you an idea of what I'm about professionally.*

I am a *Junior Metadata Engineer* at IAG Cargo, an airline based at Heathrow. I have just started my job here but I will be building data pipelines and working with big-data technologies.

Content to keep this section rather sparse for now...

- Adrian

EDUCATION

Royal Holloway 2017
Data Science & Analytics

- Master of Science
- Pass with Distinction
- Department of Computer Science

King's Collge London 2010
Mechanical Engineering

- Bachelor of Engineering
- Upper Second Class with Honours
- School of Engineering

LANGUAGES

- Java 8
- T-SQL

SOFTWARE

- IntelliJ IDEA
- MS SQL Server
- VS Code
- Git
- Jira
- Maven
- WSL

IAG CARGO

Junior Metadata Engineer

Data Platform Team

June 2019 - Present

Oracle Ingestion Automation Created an automated data ingestion process written in **Powershell 5.1** capable of picking up files from LAN directory or *Sharepoint* (via REST API) for inserting into *Oracle* database. This automation was used on multiple projects and made use of *Sharepoint Patterns & Practices*, *SQL*Loader*, and **sqlcl**.

Sales Reference db

MANCHESTER CITY FOOTBALL CLUB

Data Analyst

Fan Relationship Management

Jan. - July 2018

NYCFC Data Integration Project Owner Integrated New York City FC data into our analytical warehouse. Six-month project involving data discovery, analysis, engineering. Multiple data sources were involved: • NYCFC • Ticketmaster • Salesforce • Major League Soccer

Data Pipeline • built pipeline ingesting data from multiple databases, replacing *Informatica* solution
• achieved speed improvements using efficient DML & DDL (**OPENQUERY**, **MERGE**)

Data Cubes Storing analytical datasets in *Data Cubes* achieved
• up-stream computation of all drill-down/roll-up levels and **GROUP BY** permutations
• reduction in size of dataset, minimising bandwidth across distributed servers
• improved user-experience in *Tableau* front-end

Mentoring Dedicating time to mentoring junior colleagues remotely in Manchester/New York
• organised weekly workshops teaching basic DML and advanced DDL
• developed additional material on my website to supplement these workshops
• aimed towards self-sufficiency in writing database queries/stored procedures

GDPR Pipeline Technical Lead • integrated new GDPR schema into existing datastores (*SQL*, *Salesforce*)
• provide schema specification to **SQL** developers, advocating for indexable data types
• built efficient **MERGE** process featuring relational database design
• implemented a process to wipe personalised data belonging to any non-consenting individual stored in our data-warehouse

Customer Churn Model Modelling MCFC/NYCFC customers' future propensity to churn via *logistic regression*.
• contributed to feature selection via: – data extraction – imputation – normalisation – **R** modelling
• researched alternate models (e.g. *Beta-Geometric/Beta-Bernoulli*), academic papers, **R** APIs

I developed a number of **SQL** processes to transform customer data into CRM segments. On occasion, I took responsibility for resourcing and managing the team's workload in *Jira*.

Virgin Media Segmentation Built a flexible segmentation process able to accommodate the numerous VM mailings and myriad ad-hoc configurations.

adrian.ng/SQL/recursion

adrian.ng/openquery-xml

- wrote a flexible import process to efficiently ingest millions of tuples distributed across multiple flat-files, gaining time-savings over the built-in import wizard
- achieved efficient joining of local and remote tables via use of **OPENQUERY**, **XML**, dynamic **SQL**
- implemented efficient regex parsing via recursion, producing a one-to-many tuple mapping

Volkswagen Onboarding Worked with .NET developers and project managers to on-board a new client.

- built and tested a new process for segmenting email *and* SMS from scratch
- provided schema specification to developers for data warehousing

TUI Redesign Collaborated with TUI to integrate a new, responsive design for *Thomson* and *First Choice* large deployment broadcasts (5M+ recipients)

- wrote **TCL** scripts for dynamic **HTML** merges and gained efficiencies by moving expensive operations upstream
- provided testing; gave feedback; managed expectations on technical feasibilities
- gained recognition with client and was awarded at the end of this three-month project