

Archie Dobiss

London, UK | +44 (0)7917 332 892 | archie.dobiss@gmail.com | [linkedin.com/in/adobiss](https://www.linkedin.com/in/adobiss)

RELEVANT EXPERIENCE

Career Break

April 2023 - Present

SoundCloud (Musiio), London, UK

Solutions Engineer

July 2022 - March 2023

- Managed catalogues for AI-as-a-Service enterprise clients via API using Python, ensuring successful product trials and smooth onboarding.
- Diagnosed and resolved API overuse issues by collaborating with Tech and Sales teams and designing a diagram to align API implementation with pricing, enhancing client satisfaction.
- Created comprehensive AlaaS API documentation on Confluence, contributing to a centralised knowledge base.
- Developed a tiered pricing model, laying the groundwork for automated API billing.

Musiio (Acquired by SoundCloud), Singapore (Remote)

Freelance Software Engineer

August 2021 - November 2021

- Developed a Python script to segment audio files using AI-generated timestamps, extract metadata, upload to the cloud and generate a CSV for streamlined quality assurance.
- Debugged internal API client scripts in Python, addressing errors related to song metadata formatting and column order on CSV export.

Kobalt Music, London, UK

Distribution Manager

July 2018 - July 2020

- Developed a novel SQL-based suspense royalty transaction data mapping algorithm, building on previously designed custom matching logic, identifying over £500k in previously unallocated funds.
- Centralised royalty transaction data ingestion, securing stakeholder buy-in and enabling scalable ingestion error resolution.
- Oversaw data ingestion, transforming various structured files into a single database table according to external specifications and internal requirements.
- Conducted royalty value distribution and threshold analysis, increasing efficiency and enabling the team to take on new responsibilities.
- Curated a dataset and recruited colleagues to use NLP for musical composition matching, winning the Kobalt Hackathon 2019 End User Impact award.

Assistant Processing Manager

July 2016 - June 2018

- Conducted exploratory data analysis on datasets exceeding 10 million rows using SQL, focusing on range, percentiles, distributions and aggregation to identify data issues.
- Organised ETL algorithm enhancements to unify external data on import, contributing to over 80% automation.
- Developed custom matching algorithms to enrich third-party identifiers and utilise previously unused datapoints (e.g., sound recording info), reducing data inconsistency impact by 40%.
- Introduced a scalable long-term solution to utilise existing metadata matches, eliminating workload duplication and establishing a robust foundation for an internal transaction data ingestion tool.
- Designed SQL reports to analyse outstanding workload and communicated results to team members and senior management.

Processing Assistant

January 2015 - June 2016

- Developed and conducted metadata matching training, maintaining team-wide 98% matching accuracy.
- Identified an innovative short-term solution to reuse existing metadata matches, maintaining automation despite changes in royalty transaction data structure.
- Collaborated with Tech department colleagues to understand and document the matching algorithm structure, sharing insights with the department.

PROJECTS

- **Machine Learning Spam Filter:** Developed an XGBoost-based email spam filter in Python tailored for a sparse, unbalanced dataset. Implemented weighted loss for enhanced security and L1 regularisation to minimise overfitting, achieving 92% precision and 99% recall. [\[Portfolio\]](#)
- **Audio Resampling Webservice:** Researched frameworks and learnt FastAPI to develop an asynchronous webservice in Python within 10 days, earning positive feedback and subsequent hiring. [\[Portfolio\]](#)

EDUCATION

- **Data Wrangling w/ Python Intensive (ODSC)** (Mar 24-Apr 24): Thoroughly completed a comprehensive course focusing on practical skills in data acquisition, cleaning, pre-processing, exploration and visualisation. Applied Python, Pandas, Matplotlib, Seaborn and Google Colab to real-world data for effective manipulation and lightweight statistical analysis.
- **AI Bootcamp, University of Huddersfield, UK** (Nov 21-Mar 22): Learnt data preparation, model evaluation, classification and regression algorithms, ensemble learning, clustering, dimensionality reduction and NLP techniques. Independently developed a spam filter using documentation and the learnt techniques.

SKILLS

- **Data Handling:** Data Wrangling, Data Exploration, Data Migration, ETL Techniques, Binary Classification.
- **Data Science Programming:** Python, Pandas, NumPy, Seaborn, Matplotlib, Scikit-learn, NLTK.
- **Database/ Tech Tools:** SQL, BigQuery; Google Colab, Jupyter, Version Control (Git), REST APIs, FastAPI.