

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

SoundCloud (Musii), London, UK

Solutions Engineer

July 2022 - March 2023

- Successfully onboarded multiple enterprise clients to Musii's AI-as-a-Service APIs, providing clear functional guidance, rapid troubleshooting and efficient catalogue management via API using Python.
- Identified causes of unintentional API overuse. Proactively and collaboratively designed a diagram to align API implementation with pricing, successfully eliminating overuse.
- Created accurate AlaaS API documentation, hosted on Confluence, building a centralised knowledge base.

Musii (Acquired by SoundCloud), Singapore (Remote)

Freelance Software Engineer

August 2021 - November 2021

- Developed a Python script for segmenting audio files using AI-generated timestamps and producing a CSV file for streamlined QA, utilising advanced data manipulation and cloud storage integration.
- Refined internal API client in Python for consistent CSV output and file name parsing, enhancing efficiency.
- Labelled datasets for autotune and vocal quality, adhering to precise definitions to ensure classifier accuracy.

Kobalt Music, London, UK

Royalties Distribution Manager

July 2013 - July 2020

July 2018 - July 2020

- Managed royalty transaction data ingestion, resulting in accurate mapping and transformation of data from various structured files into SQL databases according to external specs and internal requirements.
- Developed an SQL-based royalty transaction data mapping algorithm, successfully matching over £500k in royalties previously held in suspense.
- Proposed a use case and curated the dataset to build a musical composition matching algorithm, contributing to winning the Kobalt Hackathon 2019 End User Impact award.

Assistant Royalties Processing Manager

July 2016 - June 2018

- Performed daily exploratory data analysis on 10+ million row datasets using SQL focusing on range, outliers, and distributions. Dealt with missing values and enhanced ETL algorithms to achieve over 80% automation.
- Proactively pre-processed and extracted vital metadata from a non-machine-readable format for 100k+ onboarded copyrights, reducing manual workload by 40%.

PROJECTS

- **XGBoost Spam Filter:** Developed an XGBoost-based 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 about data preparation, model evaluation, classification and regression algorithms, ensemble learning, clustering, dimensionality reduction and NLP techniques. Independently developed a spam filter using documentation and 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.