

Archie Dobiss

Address: London, UK

Phone: +44 (0)7917 332 892

Email: archie.dobiss@gmail.com

Portfolio: adobiss.github.io

LinkedIn: linkedin.com/in/adobiss

Professional Summary

End-to-end Data Specialist with over 5 years of experience managing and processing large datasets, ensuring data integrity and high levels of automation. Skilled in Python, Pandas and SQL, with hands-on experience leveraging cutting-edge AI/ML techniques for innovative solutions in data processing and semantic search applications. Adept at collaborating with cross-functional teams and solving complex data-related challenges. Seeking to leverage my experience and skills to streamline and enhance data management efforts across various sectors.

Professional Experience

Independent AI Product Development, London, UK

End-to-End Data Specialist | April 2023 - Present

- **Project Ownership & Strategy:** Identified use case and developed a comprehensive project roadmap, including scalability and production-level deployment considerations.
- **Data Engineering & Exploration:** Executed data engineering tasks, including data pre-processing and feature engineering for a 700k sample dataset using Pandas, Matplotlib and Seaborn, ensuring optimal data integrity and transformation for NLP-based tasks.
- **AI Integration & Solution Design:** Designed and implemented PoC AI solutions using Sentence-BERT language model for vector embeddings and FAISS library for similarity search, achieving a 97% accuracy rate. [\[Portfolio\]](#)

SoundCloud, London, UK

Solution Engineer, Musiio by SoundCloud | July 2022 - March 2023

- **Python, API:** Successfully onboarded multiple enterprise clients to Musiio's AI-as-a-Service APIs, providing rapid troubleshooting and efficient catalogue management via API using Python.
- **Documentation:** Created accurate AI-as-a-Service API documentation hosted on Confluence, establishing a centralised knowledge base.

Musiio (Remote)

Software Engineer (Freelance) | August 2021 - November 2021

- **Python, AWS S3:** Developed a Python script to dynamically segment audio files using AI-generated timestamps, enrich metadata, upload to the cloud and generate a CSV for streamlined quality assurance.

Kobalt Music, London, UK

Royalties Distribution Manager | July 2018 - July 2020

- **Data Pipeline Management:** Created mappings for all types of royalty transaction data file formats (delimited, fixed-width, forward-reference) into SQL databases according to external specs and internal requirements.
- **Communication:** Centralised royalty transaction data ingestion, securing stakeholder buy-in and enabling scalable ingestion error resolution.
- **SQL:** Developed an SQL-based royalty transaction data mapping algorithm, successfully matching over £500k in royalties previously held in suspense.

Kobalt Music, London, UK

Assistant Royalties Processing Manager | July 2016 - June 2018

- **Data Analysis:** Utilised SQL on datasets exceeding 10 million rows to identify royalty matching issues, consistently surpassing the 98% accuracy benchmark.
- **ETL Processes:** Proactively introduced critical data point transformations, maintaining over 80% automated royalty processing during a period of rapid growth in royalty volume.
- **Problem-solving:** Proactively introduced matching model improvements, eliminating processing workload duplication for major sources and reducing overall manual processing by 40%.

Education

The University of Huddersfield

AI Bootcamp | November 2021 - March 2022

The University of Bedfordshire

BA (Hons) Advertising and Marketing Communications | 2:1

Skills

- **Programming Languages:** Python, SQL.
- **Data Processing:** Pandas, NumPy, Matplotlib, Seaborn.
- **Data Warehousing:** ETL processes, Data pipeline design.
- **Cloud Services:** AWS Cloud Practitioner (CLF-C02) - Udemy course (in progress).
- **AI/ML Frameworks:** Hugging Face, PyTorch, Scikit-learn, NLTK.
- **Tools:** Git, Jupyter, VS Code, Command-line interface.
- **Soft Skills:** Problem-solving, Communication, Team collaboration.