Huw Cheston

Website — Google Scholar — GitHub

London, UK

hwc31@cam.ac.uk

Experience

Research Scientist Intern (Audio Intelligence) June – August 2024

Spotify *London, UK*

- Developed end-to-end deep learning model for identifying copyrighted samples in catalogue audio recordings, **driving downstream product applications** in music content categorisation & plagiarism detection
- Results **outperformed competing system ("Shazam") by 9x**, improved upon internal model by **13%**
- Managed pipelines for generating artificial training data at **petabyte-scale** using Google Cloud & Apache Beam
- Deployed training runs on distributed GPU clusters using Ray & Kubernetes
- **Presented results** to senior company stakeholders and in a scientific paper [accessible at arXiv:2502.06364]

Music Computing Lecturer + Supervisor 2021 – 2024

University of Cambridge *Cambridge, UK*

- Delivered **100+** undergraduate supervisions and lectures on modelling and visualisation of audio data
- **Managed** four undergraduate students on music-related data science and machine learning projects

Data Science Instructor 2023

Sutton Trust *Cambridge, UK*

- Delivered workshops on music + data science for secondary-age students from state-educated backgrounds
- Designed interactive coding and statistics exercises on Google Colab, hosted on GitHub Pages

Teaching Assistant 2020 – 2021

Kingswood School *Bath, UK*

- Planned & taught music technology lessons, both in-classroom and virtually during COVID
- Managed recording studio and music technology suite, produced promotional audio-visual material for the school

Professional Musician 2016 – 2020

Freelance *UK*

- Worked with internationally recognised acts including *Clean Bandit*, *Everything Everything*, *Dinosaur*

Education

University of Cambridge 2021 – 2025

Ph.D, Music Computing *Cambridge, UK*

- **Advanced state-of-the-art** in music computing tasks including automatic performer identification
- Three **peer-reviewed computer science** publications in major (top-20) scientific journals
- Fully-funded with £75k competitive research grant from Cambridge Trust

University of Oxford 2016 – 2020

BA + MSt., Music Psychology *Oxford, UK*

- Graduated with **highest mark in year**, 85% average
- Fully-funded masters study with £25k research grant from Linacre College, Oxford

Skills

Stats: multi-level modelling, NHST workflows, time-series analysis, experiment design, optimisation
Inference: simulation, AB testing, hypothesis testing, bootstrapping, causal inference, dimensionality reduction,
Machine Learning: explainability, model selection, artificial datasets, language modelling, neural networks, big data
Languages: Python, JavaScript, HTML/CSS, R, SQL
Tools: Git, Unix, LaTeX, Google Cloud Platform, Apache Beam, Docker, Ray
Libraries: pandas, matplotlib, plotly, seaborn, scikit-learn, pytorch, statsmodels, numpy, librosa, ggplot2, captum
Domain Knowledge: music content categorisation + retrieval, audio signal processing, explainable AI, applied stats

References Available on Request