Huw Cheston

Website: https://huwcheston.github.io/
GitHub: https://github.com/huwcheston

Location: London, UK
Email: hwc31@cam.ac.uk
Phone: +44 7474 654583

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** 4 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*

- **Thesis title**: Computational Modelling of Jazz Improvisation
- **Advanced state-of-the-art** in music computing tasks including automatic performer identification
- Published **3 peer-reviewed** papers in major (top-20) scientific journals
- Fully-funded with £75k competitive research grant from Cambridge Trust
- **Service:** peer review for *Transactions of the International Society of Music Information Retrieval*, *Music & Science* journals

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, scipy, scikit-learn, pytorch, statsmodels, numpy, librosa, ggplot2, captum
Domain Knowledge: music categorisation, classification, & retrieval, audio signal processing, explainable AI

Selected Publications

Journal Articles

Cheston, H., Cross, I., & Harrison, P. M. C. (2024). Trade-offs in coordination strategies for duet jazz performances subject to network delay and jitter. *Music Perception*, *42*(1), 48–72. https://doi.org/10.31234/osf.io/z8c7w

Cheston, H., Schlichting, J. L., Cross, I., & Harrison, P. M. C. (2024a). Jazz trio database: Automated annotation of jazz piano trio recordings processed using audio source separation. *Transactions of the International Society for Music Information Retrieval*, *7*(1), 144–158. https://doi.org/10.5334/tismir.186

Cheston, H., Schlichting, J. L., Cross, I., & Harrison, P. M. C. (2024b). Rhythmic qualities of jazz improvisation predict performer identity and style in source-separated audio recordings. *Royal Society Open Science*, *11*(11), 240920. https://doi.org/10.1098/rsos.240920

Preprints

Cheston, H., Balen, J. V., & Durand, S. (2025). Automatic identification of samples in hip-hop music via multi-loss training and an artificial dataset. *arXiv*. https://doi.org/10.48550/arXiv.2502.06364

Conference Proceedings

Cheston, H., Bance, R., & Harrison, P. (2024). Characterizing jazz improvisation style through explainable performer identification models. *DMRN+19: Digital Music Research Network Workshop*. https://www.qmul.ac.uk/dmrn/media/dmrn/DMRN-19-proceedings--January.pdf

Cheston, H., Cross, I., & Harrison, P. (2023a). An automated pipeline for characterizing timing in jazz trios. *DMRN+18: Digital Music Research Network Workshop*. https://www.qmul.ac.uk/dmrn/media/dmrn/DMRN-18-Proceedings.pdf

Cheston, H., Cross, I., & Harrison, P. (2023b). Coordination strategies in networked jazz performances. *17th International Conference on Music Perception and Cognition (ICMPC)*. https://icmpc17.com/proceedings-in-zip/ICMPC17-APSCOM7-e-Proceedings.zip/

References Available on Request