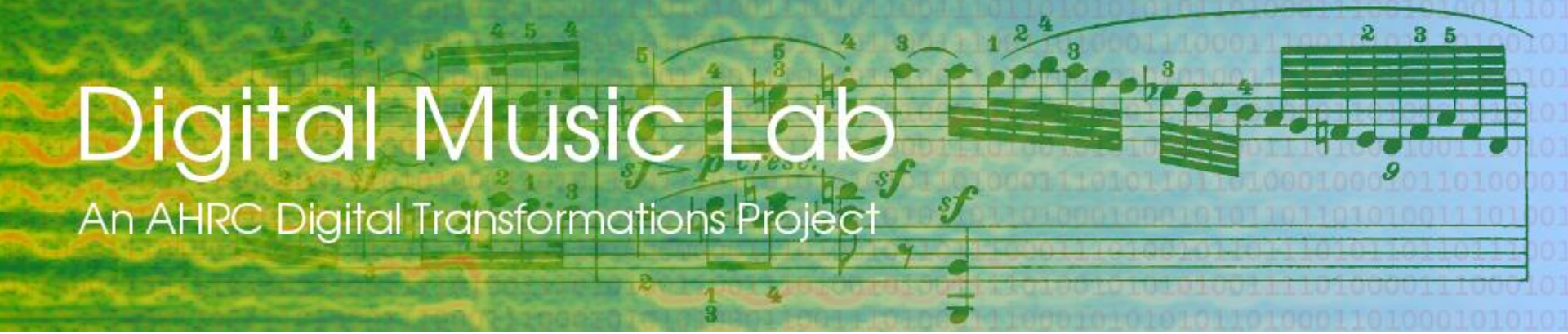


Analysing Big Music Data
Introduction

Tillman Weyde, Stephen Cottrell, Jason Dykes, Emmanouil Benetos, Mark Plumbley, Simon Dixon, Nicolas Gold, Mahendra Mahey, Aquiles Alencar-Brayner, Daniel Wolff, Alexander Kachkaev, Dan Tidhar, Mathieu Barthet, Stephen Hargreaves, Samer Abdallah, Adam Tovell

Arts & Humanities Research Council

CITY UNIVERSITY LONDON

EST 1894

Queen Mary University of London

UCL

BRITISH LIBRARY

Outline

- Digital transformations
- Why music and big data
- The Digital Music Lab project
- Research questions
- Data
- Methods
- Results
- About this workshop

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

Digital Transformations

- Music has gone digital
- What about musicology?

- Data has become a central part of culture
- How can we study digital culture?
  - Technologies
  - Opportunities
  - Challenges

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

Digital Transformations in Musicology

- Gap between musicology and music technology (music information retrieval)
- Large heterogeneous data collections
- Need for software infrastructure
- Audio and symbolic music processing
- Connecting resources (semantic web, linked data)
- Tools and visual interfaces
- Methods for gaining musical insight from data

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

The Digital Music Lab project

- Started 1st January 2014
- Ends 31st March 2015
- City University (Dpt of Computer Science, Dpt of Music)
  - Tillman Weyde, Stephen Cottrell, Jason Dykes, Emmanouil Benetos, Daniel Wolff, Dan Tidhar, Alex Kachkaev
- Queen Mary UoL (Centre for Digital Music)
  - Mark Plubmley, Simon Dixon, Mathieu Barthet, Steven Hargreaves
- University College London (Dpt of Computer Science, Centre for Digital Humanities)
  - Nicolas Gold, Samer Abdallah
- British Library (BL Labs)
  - Aquiles Alencar-Brayner, Mahendra Mahey, Adam Tovell

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

Research Questions

- How can **music research** use audio transcription and analysis on **large data collections**?
- How can we provide an **infrastructure** that enables **researchers to make use** of large data collections **and create reusable open datasets**?
- How can **computational tools** be made usable **for music researchers**, musicians and other users (who are **not necessarily computer scientists**)?

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

Objectives

- Develop a networked infrastructure to bring computation to the data
- Avoid copyright problems by design
- Integrate audio feature extraction and transcription
- Demonstrate application of analysis tools based on this infrastructure
- Interactive visual interfaces
- Musicological applications

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

What has been done

- Identified needs and questions
- Infrastructure with Semantic Web interfaces and Middleware
- Dataset preparation with features & transcription
- Visual interfaces
- Musicological studies on temperament, and chord sequences, and melodic/harmonic patterns

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

Workshop: Musicological Questions

- Automatic analysis of scores
- Structural analysis from audio (e.g. recognise repetitions of an expostition)
- Analysing styles, trends over time
- New similarity metrics e.g. performace-based
- Work across different heterogenuous collections
- Utilise external metadata and annotations

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

Infrastructure

- Feature Extraction
  - Vamp plug-ins
  - Spark and other techniques for parallelisation
- Middleware
  - Semantic Web server (RDF with Prolog using ClioPatria)
  - Music Ontology
  - Manages aggregation and collection level analysis
  - Provides SPARQL endpoint

CITY UNIVERSITY LONDON

Queen Mary University of London

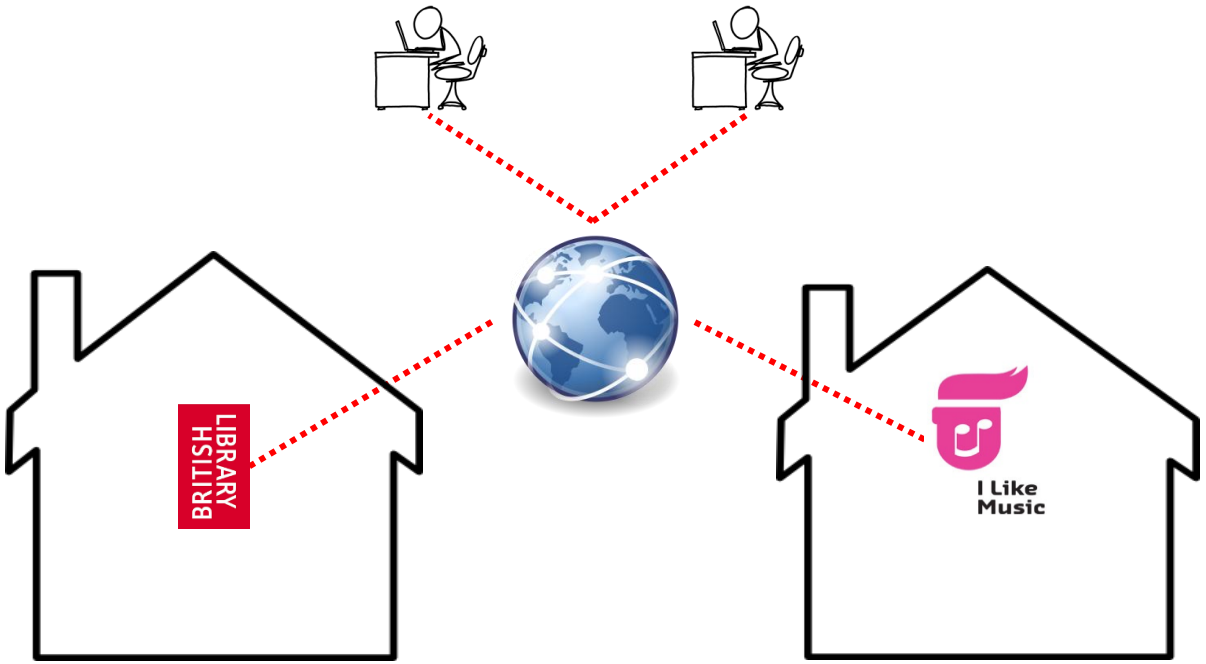
UCL

BRITISH LIBRARY

Arts & Humanities Research Council

Infrastructure

- Derived data from 2 collections
  - Accessible via the web



CITY UNIVERSITY LONDON

Queen Mary University of London

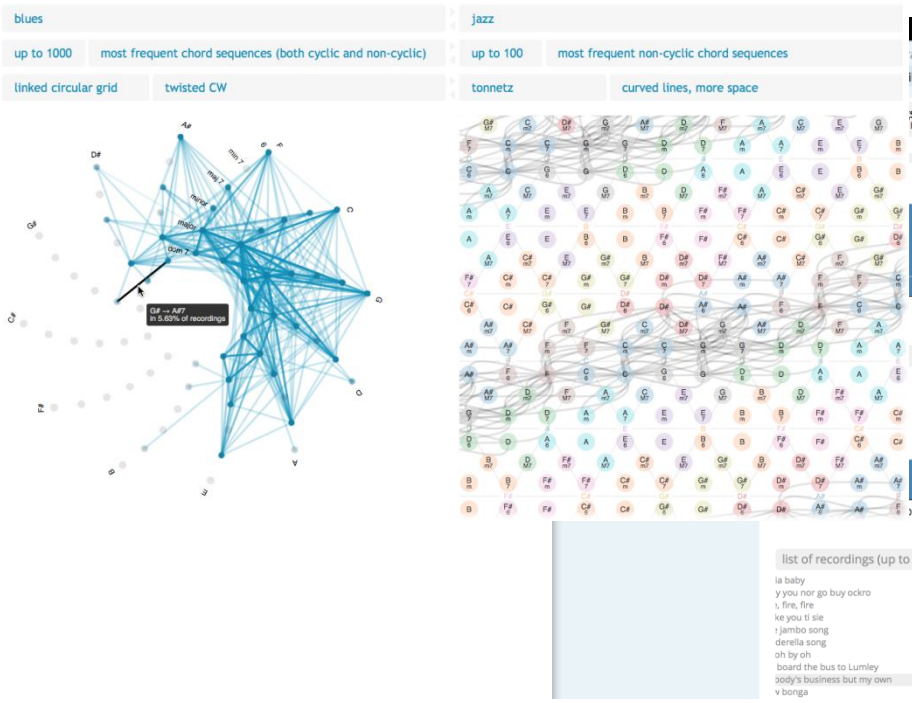
UCL

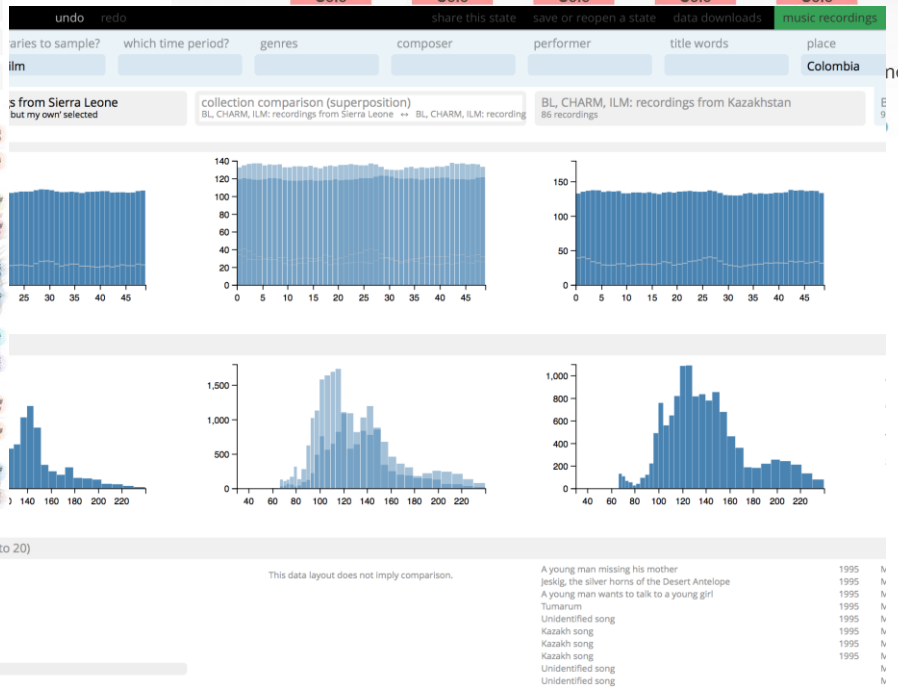
BRITISH LIBRARY

Arts & Humanities Research Council

Interfaces and Visualisations

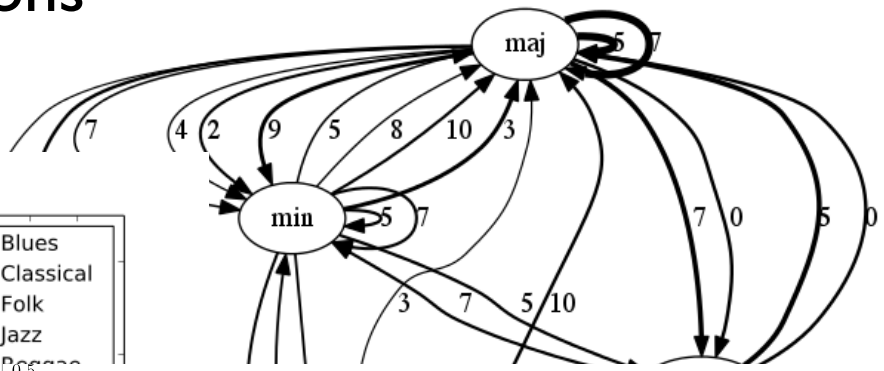
- Audio collections
- Chord sequence patterns
- Tag crowd-sourcing

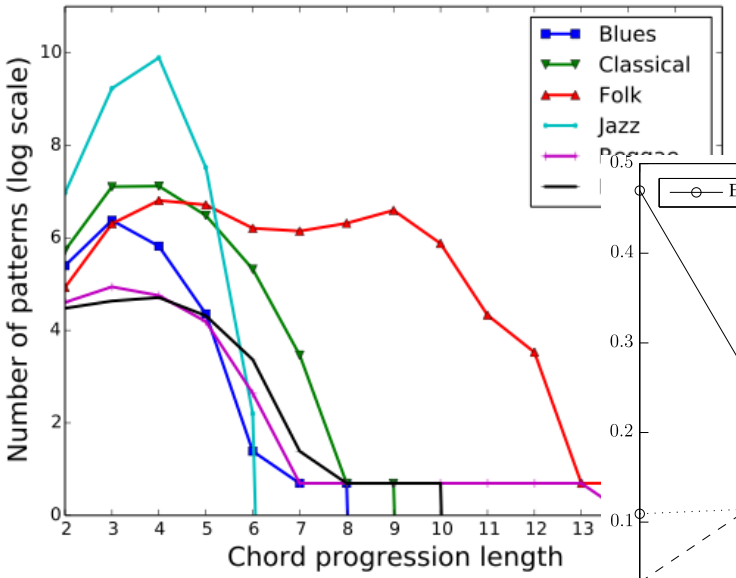


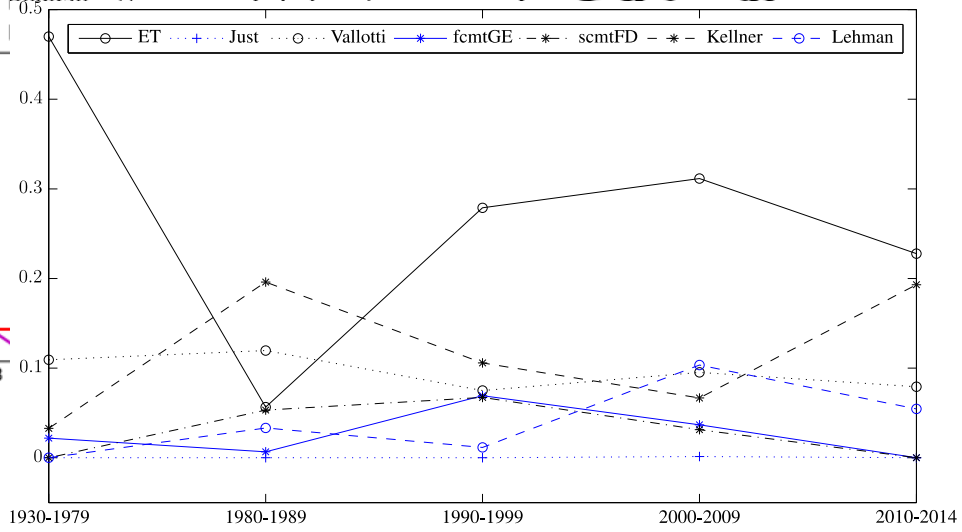


Studies

- Temperament
- Chord progressions







CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

Outputs

- Curated datasets and derived data (>4 Terabytes)
- Web service with visual interfaces
- Publications (more to come)
- Redistributable virtual machine images (in preparation)

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

What we would like to learn from you!

- How would you like to use data for research?
- How could you use our collections and features?
- Are the suggested methods useful to you?
- Are the tools useful to you?
- What data, methods and tools, would you like to have next?
- Where is the journey going?

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

The rest of today

- 10.20 Keynote Prof Lorna Hughes
- 10.50 Short talks (DML team)
- 12.05 Demos and instructions
- 12:40 Lunch (in situ)
- 13.30 Hands-on workshop
- 14.30 DML applications
- 15.00 Coffee break
- 15:30 Panel and Conclusion
- 16.30 End of workshop
  Further informal discussions

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council

Housekeeping

- WiFi
  - Select "BL Visitor" network
  - Enter your name and e-mail
  - Check your e-mail within 15min to get your credential
- Hash tag
  - #digitalmusiclab
- Collaborative document
  - Link: http://bit.ly/18hm1T9
  - For running commentary and feedback

CITY UNIVERSITY LONDON

Queen Mary University of London

UCL

BRITISH LIBRARY

Arts & Humanities Research Council