# Accelerating embedded computational analysis of Web data about music in UK universities

Q and A session
31 July 2023
Zoom

https://www.acceleratingdigitalskills.org

Eamonn Bell
Durham University

Annaliese Micallef Grimaud
Durham University

# Web Data for Music Researchers

# Have you ever wondered?

- Where do I find and how can I use online databases, datasets, and other sources to effectively support my music research?

- Should I learn to code? What can I achieve with what I already know about working online?

- How do I identify and retrieve content on social media (e.g. Twitter, TikTok, YouTube) that is relevant to my research focus?

- What is an API and how can I use them to support my research?

- How to access archived Web data from sites that are no longer online? How to preserve and manage your use of Web data for future use?

# What can be done?

- Mapping the social construction of music genres online (Born and Haworth 2017)

- Digging the construction of the aesthetics and affects of ASMR on YouTube (Spencer 2020)

- Analysing the reception of Top 40 pop music in 1M+ YouTube comments (Bell 2023)

- Understand timbre in UK EDM by looking at fan discourse online (Perevedentseva 2022, PhD thesis)

- Situated ethnographic research on hip-hop producer economies on Twitch (Ng and Gamble 2022)

- For more inspiration see:
  - [Music and the Internet Conference](#) (2023)
  - [Information Overload? Music Studies in the Age of Abundance (birmingham.ac.uk)](#) (2021)
  - [Youtube and Music Conference](#) (2020)

# Project overview

# Concept/Aims

"Web data is especially useful for the cultural study of music as the internet plays host to novel and complex cultural and social forms, like music streaming, ripping and sharing, and online co-production of music. However, **specialised digital skills are required to retrieve, process, and analyse this data [...] The project will pilot a new digital skills training programme for music researchers, which will train approximately 60 UK-based researchers at a variety of digital skill levels and representing various career stages in topics relevant to using Web data as a primary source.** The training offered will be of direct relevance to music researchers, as all workshop leaders themselves are music researchers, and they will be invited to illustrate how digital skills training has allowed them to answer exciting, transformative research questions."

# Objectives

- Needs mapping amongst **music researchers (MRs)** through survey programme and focus groups
  - In progress
- **Six one-day workshops (online)**
  - **Batch 1 - August 2023 - three workshops**
  - **Batch 2 - October 2023 (t.b.c.) - three workshops**
- Mini-grant programme to support targeted development of materials in line with user needs
- Six "The Carpentries"-style lessons published online
  - Early 2024
- Peer-reviewed journal article reporting on needs mapping, delivery of workshops/lessons, and evaluation of pilot concept
  - 2024

# Project staff

**Durham University**
- **PI - Dr Eamonn Bell (Computer Science)**
- **PDRA - Dr Annaliese Micallef Grimaud (Computer Science)**
- Co-I - Dr Kelly Jakubowski (Music)
- Co-I - Dr Robert Powell (Computer Science)

**University of Birmingham**
- Co-I - Dr Christopher Haworth (Music)

**Contracting partners**
- Three (to five) external workshop leaders (WLs)
- Sustainable Software Institute (SSI)

# Online footprint

- Website (https://www.acceleratingdigitalskills.org)
  - You can register for workshops here (Workshops (acceleratingdigitalskills.org)
- Mailing list (DIGITALSKILLS-MUSIC-ANNOUNCE@JISCMAIL.AC.UK)
  - Please sign up for project updates if you have not already done so!
  - JISCMail - Subscribe or Unsubscribe to the DIGITALSKILLS-MUSIC-ANNOUNCE List
- Twitter (@DigSkills4Music)


Shout-out: Music and Online Cultures Research Network (MOCReN)

https://mocren.org

# Workshops

# Processing Text-Based Corpora for Musical Discourse Analysis

- Tuesday 15th August 2023, 10:00-16:00 BST
- Online (Zoom)
- Dr Maria Perevedentseva (Salford)



- introduction to some of the computational tools for analysing large text-based datasets about music

- browser-based tools including Voyant and TensorFlow's Embedding Projector,

- live-coding in Python in order to explore various properties of the datasets

- information about their contents and understand the quantitative and qualitative relations they reveal

- **foundational concepts of natural language processing and corpus linguistics that will enable them to design and implement larger projects of their own**

# Ethical and Methodological Essentials for Music Researchers using Web Data

- Monday 21st August 2023, 10:00-16:00 BST

- Online (Zoom)

- Dr Steven Gamble (Bristol)



- key ethical and methodological considerations for using Web Data

- **examine best practice principles and ethical guidelines developed in neighbouring fields in addition to specific case studies in existing research on music and the internet.**

- legality, privacy, fairness, inclusion, and social justice in designing research projects including Web data.

# Command-Line Basics and Web Data Fundamentals for Music Researchers

- Wednesday 30th August 2023, 10:00-16:00 BST

- Online (Zoom)

- Dr Eamonn Bell (Durham)



- basics of interacting with command-line applications

- **a variety of command-line tools that are useful for the retrieval and storage of large amounts of data about music that circulates online**

- introduction to HTTP APIs

# Second batch (c. October 2023)

- Working with web archives (Bell)
- From Excel to R - A gentle introduction to programming for data analysis (Micallef Grimaud)
- Working with social media platforms (tbc)

# Questions and Answers

# What are The Carpentries

- A community-led 501(c) "teaching foundational computational and data science skills to researchers in academia, industry and government"
- Origins in Software Carpentry (1998) and Data Carpentry (2014); current form established 2018
- Lessons are made and reviewed by the community at The Carpentries Incubator
- Consistent format for lesson material
- Ethos of engagement and live-coding
- Attention paid to accessibility in lesson design and delivery
- Materials can work in live sessions and/or as self-paced