Supporting musicological investigations with information retrieval tools

An iterative approach to data collection

David Lewis[1], Elisabete Shibata[2], Andrew Hankinson[3], Johannes Kepper[4]
Kevin R. Page[1], Lisa Rosendahl[2], Mark Saccomano[4], Christine Siegert[2]

[1]University of Oxford | [2]BeethovenHaus Bonn | [3]RISM Digital Centre | [4]Paderborn University

Musicologists as Digital Researchers

Musicological research can be iterative and incremental

Especially in exploratory phases, musicologists may work with incomplete corpora as they explore and prioritise their topic.

Materials for research may not be pre-defined or contained within existing corpora

Supporting this means building tools to accommodate mixed-media workflows with partial digitisation carried out incrementally according to need and available effort.

Researchers work with the digital materials that are available, enriching those that they discover during their research to need more complete metadata or digital editions.

Musicological research is highly diverse. No single tool will cover the entire research process for all cases

Instead, we conside a 'toolbox' of applications that interact only through the data they import and export and which combine to support as much as possible of the research workflow and subsequent publication of results.

For tools to work together, data standards must be followed

We use MEI, which we have extended to describe explicitly selective encoding of parts of a musical work[4]. We also use existing Linked Data standards, particularly Web Annotations[2], and a new ontology for Music Annotation[1].

A scholar might start by browsing a catalogue of musical sources published as Linked Data (such as RISM or GND). They shortlist interesting material, then start looking at facsimiles digitised and distributed as IIIF by a holding library (such as digital.bodleian.ox.ac.uk).

Data is assembled during musicological research rather than before it begins

The scholar annotates these images, and runs measure detection to speed up their annotation. Where transcription would be helpful they generate selective editions of extracts using mei-friend or an OMR system and further annotate these, or use them as input for MIR tools.

The data produced by this process is potentially valuable research data, and we would hope scholars will publish this alongside more conventional outputs

Supporting Research in Practice

The **Beethoven in the House Annotator** was developed to support comparing and annotating different versions of the same music – particularly for studying arrangements of Beethoven and his contemporaries for domestic use.

The application supports annotation of library-published IIIF images and digital scores in MEI, with a shared data model for the musical regions and annotations[1], and a similar user interface. It allows a user to mark passages of interest and indicate where parallel passages occur in different versions.

Annotations to these passages can then be made using Web Annotations[2].

https://domestic-beethoven.eu/
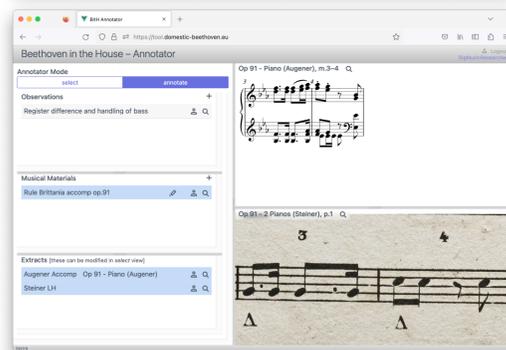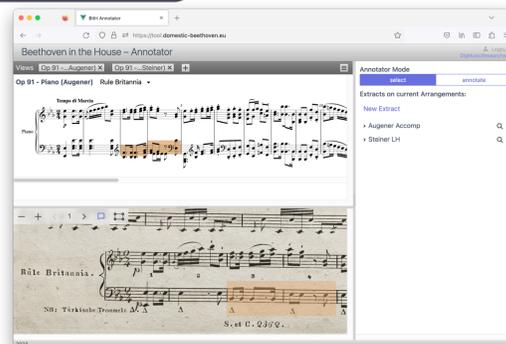
If measure detection has been run (using Cartographer or mei-friend), this is used to support whole-bar selection in images.

All these objects are written to the user's Solid pod – cloud storage on an open-source, standards-based model, which gives the user control of who can access their data.

The tool is built as a web application, based on the MELD framework[3].

Our application was evaluated through two rounds of semi-structured interviews with volunteers at the Beethoven Haus Bonn. The response was generally very positive, including the observation that the app has the potential to "bring everything together in a way I haven't experienced before".

[1] D. Lewis, E. Shibata, M. Saccomano, L. Rosendahl, J. Kepper, A. Hankinson, C. Siegert, and K. R. Page, "A model for annotating musical versions and arrangements across multiple documents and media," in Proc. DLfM 2022, ACM, 2022, 10–18.
[2] R. Sanderson, P. Ciccarese, and B. Young, Web annotation data model. W3C, W3C Recommendation, Feb. 2017.
[3] D.M. Weigl and K.R. Page. A framework for distributed semantic annotation of musical score: "Take it to the bridge!". In Proc. ISMIR 2017, Suzhou, China, October 2017.
[4] M. Saccomano, L. Rosendahl, D. Lewis, A. Hankinson, J. Kepper, K. Page, E. Shibata, Selective Encoding: Reducing the Burden of Transcription for Digital Musicologists. Joint MEC & TEI Conference 2023.

Browse & select sources → Refine sources in corpus → Annotate & structure corpus → Analyse corpus content → Evidence scholarship → Publish