Selective Encoding: Reducing the Burden of Transcription for Digital Musicologists

Mark Saccomano, Lisa Rosendahl, David Lewis, Andrew Hankinson, Johannes Kepper, Kevin Page, and Elisabete Shibata

ABSTRACT

One of the largest barriers to digital musicology is the time required to create an encoded music file. While tools exist to automate parts of the process, most of the symbolic content—pitches and rhythms—still needs to be entered manually, note by note. To facilitate the creation of corpora for digital analysis, we have developed a procedure for encoding only the portions of a score relevant to a particular study. These encodings can then be extended at a later time, by any scholar who has access to them. Currently, there is no standard way to record metadata that details which specific sections of a score have been encoded. This paper will introduce a pair of possible methods, constructed in the course of our research and tool development, to enhance the ability of MEI to accommodate these selective encodings.

The first method takes advantage of MEI’s capacity to create customized schemas. The second, simpler method takes advantage of element entailments within current MEI structures and consists of additional documentation to clarify existing usage. Our research project serves as a case study that illustrates the key assumptions underlying these two methodologies, and how project-based considerations can lead to the adoption of one approach over another.

INDEX

Keywords: Music Encoding, Metadata, Arrangements, Data Modeling, Corpora, Digital Workflows,

ACKNOWLEDGEMENTS

The project Beethoven in the House: Digital Studies of Domestic Music Arrangements investigated domestic music arrangements of the nineteenth century through two complementary studies, a digital research environment developed alongside the studies, and the innovative application of digital musicology methods within this environment. Forschungszentrum “Beethoven-Archiv,” Beethoven-Haus Bonn, Oxford e-Research Centre, University of Oxford, RISM Digital, and Paderborn University partnered for this project. The project ran from 2020 to 2023 and was jointly funded by DFG (429039809) and AHRC (AH/T01279X/1). For more information and to view our publications, please visit the project website at https://domestic-beethoven.eu

1. Introduction: Encoding as Bottleneck

1 As Lewis, et al. (2023, 796) describe, the nature of research in the humanities entails circling back to the source material in an iterative process, with each pass bringing fresh perspectives as scholars return with new facts, new data, and new conjectures. In the digital version of this scenario, research often includes a collection process in which transcriptions are made of sources for inclusion in a data set. This collection process will likely be accomplished incrementally, as a project’s needs dictate and its resources permit. The ability to accommodate incomplete sources, Lewis concludes, ought to be an explicit requirement for MIR tool development. Since workflows and methodologies in digital musicology are still evolving, it may not yet be necessary to firmly

advocate for the institution of any single procedure to facilitate this goal. However, it is advisable that researchers carefully consider this prospect and come to an agreement on an appropriate encoding scheme, and equally important, be able to articulate its rationale.

2 As anyone who has ever done any music encoding knows, transcribing a score into a digital format is a laborious, time consuming process. Transcribing multiple scores for the creation of a viable workset then becomes a significant barrier for digital musicological research. Even when assisted by notation software, making a single, edition-quality MEI transcription of even the shortest score can take up many valuable research hours. Yet in a digital context, these files often serve as the primary data set for scholarly research (ibid.). This makes the time required to create an encoding one of the single largest barriers to digital musicology. While tools exist to automate parts of the process (format converters for music notation software, for example, and optical music recognition technology), most of the symbolic content—pitches and rhythms—still needs to be entered manually, note by note. Even for an experienced encoder, it can take days or weeks to produce an encoding of a single symphonic movement. And while larger projects may have a budget for assistants to encode materials, the quantity of music that can be made ready for digital research methods remains limited.

3 To facilitate the creation of corpora for digital analysis, and to make them amenable to reuse, we have developed a procedure for creating encodings of only the portions of a score relevant to a specific research project—encodings which, if desired, can then be extended at a later time by any scholars who have access to them.

4 Currently, there is no structured way to record metadata that specifies which sections or which aspects of a score have been encoded. There is a `<samplingDecl>` element available in `<meiHead>` that accepts text descriptions of sample sizes and any selection principles that were followed, but it is optional and has no other internal structure besides generic `<head>` and `<p>`aragraph elements. (The `<mei:samplingDecl>` element is modeled on the Text Encoding Initiative's `<tei:samplingDecl>` and has been adopted unchanged for use in MEI.)

5 What we present here are two different methods for creating valid MEI editions of musical works in which only a portion of the work has been encoded—and which can then be reused and enriched for other purposes in subsequent research projects. Both were developed to work with the project's

web application (Kepper 2023) that adds digital annotations to encoded music. While quite distinct in their approach to metadata, both methods aim at hewing as close as possible to FAIR principles for data management (Wilkinson et al. 2016).[1]

6 Our more explicit, detailed method takes advantage of MEI's capacity to create customized schemas. It features both new and repurposed elements and attributes to model a digital music encoding as the complete set of a source's encoded and unencoded sections, with the aim of providing readily accessible information about the extent of the source score's representation. Our second method is a simpler approach. It works within already existing MEI structures and is essentially a clarification of existing usage, taking advantage of element entailments to identify and retrieve the information regarding its extent. This second solution is thus a matter of *documentation*, rather than modification or customization of MEI.

7 We introduce both these methods within the context of a single digital musicology project: *Beethoven in the House*, a study of Beethoven's symphonic works arranged for performance in the home.[2] Our project serves as a case study that illustrates some of the assumptions that underlie different encoding methods, and how project-based considerations can lead to the adoption of one approach over another. As is the case with many research projects, time and resources were limited, and encoding entire symphonic scores was neither feasible nor necessary, as musicologists on the project were only interested in comparing certain passages of the large scale works and how those passages were realized in various arrangements. Thus, we knew from the start that one of our tasks on this project would be developing and documenting methods in MEI for creating what we call "selective encodings."

8 In addition, our efforts anticipated an increase in the availability and use of applications that facilitate the transcription and editing of music sources (for example, the web app mei-friend developed by the Department of Music Acoustics at the University of Music and Performing Arts in Vienna, and version 4.2 of the recently retooled MuseScore, which now provides support for MEI files), further expanding access to encoding technology and increasing the appeal of MEI for a wider variety of digitally-based musicology projects. By relieving researchers of the burden of learning MEI's strict guidelines for structuring data before embarking on or contributing to a digital musicology project, such editing applications have the potential to generate a much greater number of encodings for individual projects than was previously possible. Although these

encodings may begin their life as project-based resources with limited application beyond an individual study, they can be extended and enriched. They also may go through several stages, as newly transcribed portions are added to a document. To facilitate the discovery and encourage the reuse and enrichment of these encodings, we needed a standard, unambiguous way to identify what parts of a score have been encoded in each file.

9 Initially, we surmised that we would need to create a detailed metadata scheme that included new, specially purposed elements in order to provide a proper accounting of the transcribed material contained in selective encodings. We show here what such a scheme would look like and what steps would be necessary for its implementation. This metadata structure is somewhat complex, not only due to its encoding requirements, but also in regard to the procedure for proposing and implementing changes to the MEI standard itself. So we also devised a simpler method in which extent could be signaled to an application indirectly, through preprocessing with OMR technology. Below is a description of the two schemes, and an account of how we came to choose one method over the other.

2. Encoding Methods

10 In the context of the project, we developed two methods to facilitate the selective encoding of scores. Both of the methods make use of the `<samplingDecl>` (sampling declaration) element. This element originates in TEI, where it is used in the `<encodingDescription>` element to describe the criteria and methods used in compiling texts for a corpus or collection. It was transferred to MEI unchanged and allows only paragraph and heading elements. It can be extended with a few attributes, but it is a fairly simple element that contains mostly text.

2.1 First Method

11 For the first method, we have modified the structure of the `<samplingDecl>` element. Instead of just defining the encoding practice, we wanted it to provide a comprehensive listing of encoded and unencoded sections of the source work in the header, similar to how the notesStatement acts as a parent element to collect individual annotation elements.

12 The introduced or repurposed elements for this method are the following:

- `<samplingDecl>`: The sampling declaration contains a collection of sample and gap elements that represent all parts of a source.
- `<section>`: The section element delineates a part of the music we are encoding and gives it an xml:id. This distinguishes arbitrary segments of music.
- `<sample>`: Using the section elements, each sample element can now correspond to an ID. A number of attributes provide additional information about the selection.
- `<gap>`: Between the encoded sections in the body, as well as between the sample elements in the sampling declaration, we use gap elements to indicate that some content in the source has not been realized in the encoding.
- `<range>`: To specify the number of measures within the sample element, the range element can be used.

13 These elements allow the organization of the encoded sections (see Example 2) and make the information about the sample machine-readable, except for any explanatory text in the `@aspect` and `@reason` attributes. It is possible to see what measurements are encoded and why by looking at the header rather than searching through the body, making the encoding more human readable.

Example 1. Example for the first method (header).

```
  <meiHead> [...] <encodingDesc>
    <samplingDecl xml:id="n69a6223-2b20-4aed-a2b0-896663052493">
     <p>Only measures for comparison with other arrangements are
     encoded.</p>
     <sampleList>
      <sample n="1" xml:id="m278540e-22ae-4fa0-91e5-155fdd7a94e2"
corresp="#f3107d22-44ad-479f-a71d-0c855b6bb278" aspect="fully encoded"
reason="comparing instrumentation at beginning" resp="#LR" isodate="2022-04-24"/>
      <sample n="2" xml:id="e375f67e-f47f-4b79-a870-426948259d7b"
corresp="#d1c6af8e-b2b8-48de-b293-78b93dcd13aa" aspect="fully encoded"
reason="crescendo to rare fff at bar 190" resp="#someoneNew" isodate="2022-04-26"/
>
      <gap xml:id="s38027c2-2dba-4db2-b261-d6c6c7f66cba"
corresp="#d6ea886f-99d9-4433-a1a8-2ab761e99a70" extent="191-360" unit="measure"/>
```

```
      <sample n="3" xml:id="s6f8d2af-8798-4382-84b1-f8c92ec18a72"
corresp="#m19fc126-7849-48f6-85ee-352fc6799103" aspect="measure positions, but no
notes" resp="#cartographerApp" isodate="2022-05-27"/>
     </sampleList>
    </samplingDecl>
   </encodingDesc> [...] </meiHead>
```

Example 2. Example for the first method (body).

```
  <body> [...] <section xml:id="f3107d22-44ad-479f-a71d-0c855b6bb278"
decls="#m278540e-22ae-4fa0-91e5-155fdd7a94e2" resp="#LR">
    <measure> [...] </measure> [...] </section>
   <section xml:id="d1c6af8e-b2b8-48de-b293-78b93dcd13aa" decls="#e375f67e-
f47f-4b79-a870-426948259d7b" resp="#someoneNew">
    <measure> [...] </measure> [...] </section>
   <section xml:id="d6ea886f-99d9-4433-a1a8-2ab761e99a70"
decls="#s38027c2-2dba-4db2-b261-d6c6c7f66cba">
    <gap extent="148" unit="measure" reason="sampling"/>
   </section>
   <section xml:id="m19fc126-7849-48f6-85ee-352fc6799103"
decls="#s6f8d2af-8798-4382-84b1-f8c92ec18a72" resp="#LR">
    <measure> [...] </measure> [...] </section> [...] </body>
```

14 All content within the body of a selective encoding should be contained in at least one section so that it can be referenced in the header. But there is a lot of flexibility in this method: The sections themselves can point to the samples and gaps in the <samplingDecl>. It is also possible to name the author of the encoded sections either in the <body> itself or in the <samplingDecl> by pointing up to the <responsibilityStatement>. The extent of the gap can be described either with an attribute in the body or with an extent element in the sampling declaration. If the extent attribute is used and contains a number, it should be accompanied by a unit attribute. It is possible to use the reason attribute here to declare that the reason for the gap is not in the source itself, but rather is the result of editorial decisions. However, this should not be necessary, since the header already contains this information.

15 The <samplingDecl> itself has not only a short paragraph about the sampling method, but also a <sampleList>—an element we created for collating the encoded and unencoded sections of a score. It lists the <sample> and <gap> elements that are linked to the <section>s in the <body>. It is possible to add attributes that give the reason for choosing a sample or define the particular aspect of the music that is encoded (for example only the piano part). If there is more than one encoder, it is possible to link each sample to the <responsibilityStatement> and add the encoding date here. This provides a comprehensive overview of the encoding status, making it easy to retrieve the data.

2.2 Second Method

16 The second method is a simpler approach that works within the current MEI structures and seeks to clarify and refine existing documentation by using element entailments and text descriptions to identify and retrieve encoding information from an MEI file. The usage of the <samplingDecl> here is in line with its intended use for TEI: it entails a few sentences explaining the editorial reasons governing the selections—which remain the same for every encoding in the project—rather than giving the specific reason for choosing each sample. This avoids confusion with the annotations about the comparison between different arrangements that were added later for the project.

Example 3. Example for the second method (header).

```
<samplingDecl>
  <p>The project "Beethoven in the House" aimed to provide a digital study of
  Beethoven’s symphonic works arranged for home performance. As time and
  resources were limited, encoding entire symphonic scores was neither
  feasible nor necessary. Instead, the project focused on comparing certain
  passages of the large-scale works and their realisation in various
  arrangements. This led to the development of a procedure for "selective
  encodings," which involves encoding only the portions of a score relevant
  to a particular study.</p>
  <p>This is a Selective Encoding that includes measures with empty layers. By
  contrast, a measure that is empty in the source would be encoded using
  mei:mSpace. Encoded measures are separated from empty measures with
  section elements with an @xml:id. For the encoded sections the attribute
  @resp is used to indicate the transcriber.</p>
  <p>For more information, read the documentation at
```

```
  https://doi.org/10.5281/zenodo.7870625.</p>
 </samplingDecl>
```

17 Example 4 shows the body of an encoding of a piano arrangement of *Wellingtons Sieg*. Instead of the `<gap>` element, empty `<layer>` elements explicitly signify unencoded content, which is material that exists to be transcribed, but for sampling reasons, is not part of the encoding yet. Compared to the `<measureSpace>` element, a measure with empty layers is agnostic about what is in the source. Encoded areas are separated from the areas with empty bars by section elements, which are referenceable by IDs and linked to the `<responsibilityStatement>`. The bars themselves are linked to the facsimile, which would not be possible by using the `<gap>` element instead. Encoding the bars with empty layers makes them machine-readable and allows an editing application to add new content, thus making it possible to add a IIIF image to each measure by using the `@facs` element, as shown in the example. In our project the empty layers for the sections without content were created using an XSLT and then loaded into the Edirom application Cartographer-Online.[3] The Cartographer app not only recognises the measures in the image, but is also able to associate them with the measures of an MEI encoding.

Example 4. Example for the second method (body).

```
 <section xml:id="m053a3dc0-d6db-428c-8a7e-928d12bc70cf" resp="#DM">
  <measure xml:id="b59e443c2-c2e9-4406-ae04-937660a0e1b2"
facs="#d1b3d6a86-554e-4697-8d9d-5ded94e4b2bd" n="382">
   <staff xml:id="m6295d1e5-8ed6-4f65-a1d0-f33d271d6517" n="1">
    <layer xml:id="m7bce773f-62c6-492a-a6e6-e1d5653b0438" n="1">
     <chord xml:id="m4adb7cf8-a8c5-47cc-b59a-a7381880de58" artic="stacciss"
dur="8" stem.dir="up"> [...] </chord>
    </layer>
   </staff>
  </measure> [...] </section>
 <section xml:id="b677f8a9-0504-4343-b34e-4ac52cd845d7">
  <measure xml:id="bab5634dc-0dfc-495e-b937-4d44e00ab7bc"
facs="#d1b827878-691e-49cd-8934-bcf2d24cc71d" n="386">
   <staff n="1">
    <layer/>
   </staff>
   <staff n="2">
```

```
      <layer/>
     </staff>
    </measure>
    <sb/>
    <measure xml:id="bd63bbbcc-97ef-4ef6-8c83-620936fc0796"
facs="#dbc1e35fc-4d1b-40dc-acb8-94a7e63f95e3" n="387">
     <staff n="1">
      <layer/>
     </staff>
     <staff n="2">
      <layer/>
     </staff>
    </measure> [...] </section>
```

18 In the end, the choice between these two methods depends largely on the specific needs of the project, the resources available and the intended use of the encodings. Our project benefited greatly from the development of these selective encoding methods. Although we didn't end up using the first method for the project, the act of searching for the right elements and attributes and discussing different approaches was useful to find our own solution. This is why we hope that our documentation of the decision process can help other projects to find their own approach.

3. Conclusion: Different Paths to FAIR

19 An important benefit of designing and formalizing an approach to selective encodings is that it facilitates the digital study of new repertoires, including works and composers that may otherwise go unexamined (Lewis et al. 2023). As Anna Kijas (2018) points out, many currently available encodings are a result of University initiatives and funding strategies that focus on composers already well represented in existing editions and music scholarship.[4] Such projects affirm and propagate a canon dominated by Western musical traditions, with white male composers as exemplars.

20 The designation of "selective encoding" is also an acknowledgment that digital research methods often extend and enrich their source material during the course of a project. As Lewis et al. also note, when dealing with material that is found in untranscribed documents, an iterative approach is often the only practical means by which research can proceed. The digital resources

produced by such projects, at each stage, are selectively encoded music editions, and like all digital editions, they are always capable of being enriched, no matter what degree of "completeness" is achieved. This is an important point in the context of the FAIR principles, a key and often required component of open scholarship. Nick Thieberger writes that we don't often think of our own work as being archivable resources; however,

> if your work creates new documents … [or] provides a new interpretation, then that interpretation and those documents need to be made available to the audience you are writing for. Creating well-formed 'archive-ready' research materials makes them reusable, and archiving ensures that … others can benefit from work you have done.
>
> (Thierberger 2018)

21 To facilitate reuse of materials, we maintain that clear details regarding the encoded portions of the work referenced by the title statement should be presented in an easily findable and easily readable location in the document. As stated in the FAIR Principles: "Metadata and data should be easy to find for both humans and computers."[5]This is what our first solution set out to accomplish. The series of `<sample>` elements in the header give a potential file user an overview of the current state of the encoding. This placement is preferable to looking through the contents of the `<body>` tag, making note of any sections with the `@decl` attribute, and then scrolling through the file in order to uncover the extent of its encoding. Even when a document is intended for use in an editing application, it would be helpful to be able to determine the file contents without loading and processing it first. This additional structure in the header also makes it easier to devise machine processing methods (e.g., XSLT or XQUERY) for retrieving the sampling information.

22 In order to fully realize the potential for the reuse and enhancement of project-based MEI encodings, this method would need to be incorporated into the MEI format itself, and documented in its Guidelines, thus ensuring interoperability. This is a major drawback of such a procedure, as it requires the MEI community and its technical team to fully review the proposal before adopting it. The review process is a lengthy one and ultimate approval is not guaranteed. In addition, MEI tends to be conservative with respect to changes in the standard. It is important to remember that an early adopter of the format was the digital editions community, and much of MEI's development

has proceeded with the needs of this particular community in mind: stability and consistency fosters continued adoption and use of the standard, an important consideration that helps ensure maximal reusability of encodings.

23 In cases where music is encoded for the purpose of annotation, or where corrections are required after importing encodings from notation software, or processing them with Optical Music Recognition software, editing applications will necessarily be recording the extent of the encoding. If a user is then manually adding this quantitative information inside a `<samplingDecl>`—information derivable from countable elements within the document—then this constitutes a duplication of information. Such duplication is a clear violation of the programming mandate "Don't Repeat Yourself," inviting errors and inconsistency into the data. It might be possible to create an XSLT that would automatically calculate the extent of the encoding and insert that information in the proper format into the header, but that would present a highly complex, labor-intensive programming task, bringing us right back to the limited-resources problem we were trying to solve. Since we were designing a process for musicologists, not expert encoders, it made sense to develop a system that would allow users to directly enter notes and not have to worry about `<scoreDef>`s and `<staffDef>`s and `<staffGrp>`s and `<layer>`s—or header elements like `<revisionDesc>`. This is a strong argument for allowing the `<samplingDecl>` to be limited to a simple generic statement about why it consists of empty `<layer>` elements. While we had consulted with the wider MEI community for input and feedback on the more detailed structure, we soon discovered that the introduction of new elements and attributes in MEI can be at odds with an archival philosophy that prioritizes the preservation of materials and interoperability of digital resources. The *Beethoven in the House* project, by design, brought together MEI's users and innovators with the theorists and technicians who are tasked with its stewardship. What resulted for us was not a clash between two competing approaches, but rather *clarity* by having to articulate what was being proposed and why. In this way we could also anticipate the pitfalls that might lie ahead, and plot a prudent course going forward.

24 Ultimately, this led to a reframing of the entire concept of selective encoding: Note that we specifically call these encodings *selective*, rather than *partial*, or *sparse*, as though something were lacking. That's because *all* digital transcription work is about increasing knowledge by bringing materials to light and sharing them with others. And to make the resulting encodings Findable,

Accessible, Interoperable and Reusable (which is what both proposals aim to do), is to enrich these files, not to complete or repair them and then label them as *finished*. Every encoding will always be in some sense partial. Some features are going to be prioritized at the expense of others. Researchers have different needs and different purposes for their encodings. Not every encoder will assign significance to every mark on the page. And some degree of interpretation is necessary when deciphering a set of symbols: editorial decisions are always being made at every level. The choice that transcribers have is how explicit to be about what has been encoded.

25 The solution that our project adopted through a process of interdisciplinary dialog also benefited from reframing. Since it introduced no new elements, and no changes in existing structure, we came to view this solution as merely a question of documentation, further easing the degree of direct intervention into the MEI schema that is required to selectively encode scores: we simply specify aspects of structure that had been left undefined in the MEI Guidelines.

26 As disciplinary work continues to become less siloed, and as finding and relating resources with Linked Data becomes more practicable, the progressive enrichment of encodings will inevitably increase, creating an ever-growing number of usable transcriptions for future researchers. While a project may have needs that can be easily addressed with additional data structures, it can be worthwhile to consider instead adapting to the data model of an existing standard, thus better ensuring that a project's research contributions can be shared and its data reused. [6]

BIBLIOGRAPHY

"FAIR Principles." GO FAIR. https://www.go-fair.org/fair-principles/ (February 20, 2024).

Kepper, J. *DomesticBeethoven/bith-annotator: Release 2023–04*. Zenodo, April 28, 2023. https://doi.org/10.5281/zenodo.7877741.

Kijas, A. "What does the data tell us?: Representation, canon, and music encoding." Keynote at Music Encoding Conference 2018, Maryland, May 24, 2018. http://dx.doi.org/10.17613/yp01-mw44.

Lewis, D. Shibata, E Hankinson, A. Kepper, J. Page, K. R. Rosendahl, L. Saccomano, M. Siegert, C. "Supporting musicological investigations with information retrieval tools: an iterative approach to data collection." In *Proceedings of the 24th International Society for Music Information Retrieval Conference (ISMIR)*, Milan, 2023.

Rosendahl, L. *Beethoven in the House: Selective Encodings of Arrangements of Beethoven's opp. 91, 92, and 93*. Zenodo, April 28, 2023. https://doi.org/10.5281/zenodo.7875059.

Thieberger, N. "Research Methods in Recording Oral Tradition: Choosing Between the Evanescence of the Digital or the Senescence of the Analog." In *Research Methods for the Digital Humanities*, edited by Lewis Levenberg, Tai Neilson, and David Rheams, 233–241. Springer International Publishing, 2018. doi:10.1007/978-3-319-96713-4_13.

Wilkinson, M. D. et al. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data*, 3:60018, 2016. doi:10.1038/sdata.2016.18 .

NOTES

1 FAIR: Findability, Accessibility, Interoperability, Reuse. https://www.go-fair.org/fair-principles/.

2 https://domestic-beethoven.eu/

3 https://github.com/Edirom/cartographer-app

4 "Universities often focus their efforts on large-scale digitization of hegemonic texts … which again reinforces canonization. The collections in libraries and archives, primarily those in first-world countries with access to digital imaging equipment and digital library infrastructure, as well as institutional or grant funding, perpetuate not only canonization, but also *colonization*." (Kijas, 2018).

5 https://www.go-fair.org/fair-principles/. The GO FAIR website serves as a concise web reference to the FAIR Principles.

6 The project encodings are published on Zenodo and can be reused (Rosendahl 2023).

AUTHORS

MARK SACCOMANO

Mark Saccomano is a music theorist at Paderborn University and a research associate for the Beethoven in the House project. He previously taught music history and music theory at Columbia University in New York and was adjunct professor of music at Montclair University in New Jersey.

LISA ROSENDAHL

Lisa Rosendahl is a research associate on the project Beethovens Werkstatt at Beethoven-Haus Bonn. With master's degrees in history and musicology, as well as a certificate in digital humanities, she brings an interdisciplinary approach to her research on music and social history of the eighteenth and nineteenth centuries.

DAVID LEWIS

David Lewis trained as a historical musicologist at Kings College, London. He is currently a researcher at the University of Oxford e-Research Centre and lecturer in Computer Science at Goldsmiths.

ANDREW HANKINSON

Andrew Hankinson is a researcher and software developer for the RISM Digital Center in Bern. He has held positions on the technical group and board of the Music Encoding Initiative.

JOHANNES KEPPER

Johannes Kepper studied music and media science as well as computer science at the Musicology Seminar Detmold/Paderborn and the University of Paderborn. Since 2006 he has been active in the development of the Music Encoding Initiative (MEI) and is the German PI of the Beethoven in the House project.

KEVIN PAGE

Kevin Page is an associate faculty member and senior researcher at the University of Oxford e-Research Centre. He is co-founder of the Digital Libraries for Musicology conference, teaches digital musicology and linked data methods for the Master's programme in Digital Scholarship at Oxford, and is the UK PI of the Beethoven in the House project.

ELISABETE SHIBATA

Elisabete Shibata focuses on the connection between music and new technologies. She is currently pursuing her PhD under Prof. Dr. Frank Hentschel at the University of Cologne, where she is investigating the digital representation of arrangements using Beethoven's music as an example.