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, UK [2] BeethovenHaus Bonn, Germany
[3] RISM Digital Centre, Switzerland [4] Paderborn University, Germany
david.lewis@oerc.ox.ac.uk

ABSTRACT

Digital musicology research often proceeds by extending and enriching its evidence base as it progresses, rather than starting with a complete corpus of data and metadata, as a consequence of an emergent research need.

In this paper, we consider a research workflow which assumes an incremental approach to data gathering and annotation. We describe tooling which implements parts of this workflow, developed to support the study of nineteenth-century music arrangements, and evaluate the applicability of our approach through interviews with musicologists and music editors who have used the tools. We conclude by considering extensions of this approach and the wider implications for digital musicology and music information retrieval.

1. INTRODUCTION

Digital humanities research often extends and enriches an evidence base – in the form of digital, machine-accessible corpora – as it progresses, mirroring a methodological process of evidence gathering and preparation that is common and accepted in analogue research. Rather than assuming complete corpus encoding as a prerequisite for digital scholarship, we anticipate that research subjects will more usually be found in un-transcribed and only minimally-catalogued documents. A researcher or team can thereby more effectively support their work by digitising, transcribing, and annotating a corpus incrementally. Resource limitations will generally mean that this is most efficiently carried out in an incomplete way, producing partial editions of short extracts or individual instrumental parts, instead of a complete corpus as an outcome of the investigation. To support this mode of digital scholarship, we propose that incremental workflows, which manipulate and analyse incomplete sources, should be an explicit consideration for applied MIR assemblies.

We present an example of this approach, from the *Beethoven in the House* project, where musical arrangements and miscellaneous music publications aimed at a domestic market are the subject of the scholarship. In this case, little of the music has been published in modern editions, and no digital editions existed at the start of the research process. Some sources had been photographed and published online before the project began, and the remainder were digitised at the request of the project. Our data model abstracts the musical structures from the surface presented by digital representations themselves, so that our tools can switch transparently between working with digital scores and facsimile images, with measure detection supporting the transition. We also use Linked Data and user-authored, web-based storage, which supports the enrichment of institutional data resources, such as library images, without requiring that scholars have write access to those servers. We focus on chained components and data compatibility rather than trying to build end-to-end tools. Our ambition is that, at the end of the process, the digital tools support our own research, as well as supporting reusability and transparency, since the 'working materials' can be published along with the finished results.

In this paper, we consider a research workflow which assumes an incomplete and incremental approach to data gathering and annotation. We describe tooling implementing this workflow, and evaluate the applicability of the approach through interviews with musicologists and music editors who have used the tools. We conclude by considering extensions of this approach and the wider implications for digital musicology and MIR.

2. MUSICOLOGISTS AS DIGITAL RESEARCHERS

Most Information Retrieval implementations are optimised from the perspective of a 'whole' or 'complete' corpus, produced by some prior acts of digitisation, being interrogated by a user motivated by a single, explicit information need. This approach facilitates the optimisation of retrieval tool engineering, since the elements of the system are well known, and the quality of tools can be transparently

quantified, assessed, evaluated, and compared. Meanwhile Bates's model of berry picking [1] is based on the observation that information needs often develop during the user's interactions with a system, as a part of a research process that takes new findings into account in the search. Different information-seeking strategies and their modes of search and scope of application (whether based on content, features or metadata) are further teased out by Weigl et al [2]. While this does not replace or reject the engineering of MIR tools based on concrete requirements, design, and evaluation, it does suggest we should consider such tools being recomposed as components within a multitude of individualised workflows – where the overall object of the composite workflow cannot be determined a priori. We reflect that this is especially true when MIR tools are used as a means to undertake curiosity-driven research, as they often are in support of digital musicology.

A similar pattern is identified for data as well as tools. Fenlon et al. [3] note the role of selecting and gathering data in the research process so that, as the investigation evolves, so may the subset of the corpus being studied. More recently, Oberbichler et al. [4] have observed that the separation between the management of digital materials and their analysis is less clear for humanities scholarship. They note that the clarity and separation of workflows and responsibilities in digitising, organising, and interrogating collections that make for efficient, maintainable solutions may be problematic in these domains.

For notated music, where data entry remains expensive in terms of time and effort, separating between digitisation, digital editing, metadata organisation, and research can mean that much musical heritage is ruled out from digital research, as digital editors become unwitting gatekeepers of our history. This can have the effect of channelling research into canonical composers and works, and diverting it from less-well-represented areas and niche and regional music [5]. The lack of encoded corpora appropriate to their research has long been identified as an important problem for musicologists [6, 7]. Although it is true that these issues could be addressed by comprehensive and complete mass digitisation and encoding, in the absence of this, an alternative strategy may be required.

We have seen that berry picking can be extended to accept that the research process involves partial and changing research questions, and even that during the investigation, the researcher may add to, correct or enrich the metadata [2]. An alternative interaction model might extend berry picking to acknowledge that this is true for the data itself. Clearly, this may pose problems for statistical evaluation of IR tools, and necessitates consideration of alternative approaches to system and workflow design. Nonetheless we can demonstrate that it is a mode of use aligned with the needs – and limited resources – of digital musicologists.

Given limited resources, we cannot assume that a single scholar, or even a funded research project, can transcribe the complete corpus of music that might be relevant to their investigations – including any comparison or control groups – prior to research commencing, and even producing a complete digitisation by the end of their investigations may prove impractical. Creating an expectation of the prior existence of these primary objects of study may feed the sense of "disconnect between this research strand and musicological users' needs and requirements" identified by Inskip and Wiering [8]. A better approach would accommodate images or partial editions – transcriptions of only a few bars or one instrumental part – created incrementally as the research progresses.

Many of the basic tools that already exist could be made to accommodate this approach well, indeed the extra information that may be available in a digital environment at a later research stage may help them, supporting a bootstrapping approach to training or parameter tuning. Without musicologist-facing, high-level tools built on these, researchers are more likely to resort to less machine-accessible approaches, such as pen and paper or local spreadsheets.

In this paper we explore this interaction model through a set of prototypes. In the next section, we describe a workflow and tooling designed to support musicological research in previously digitally unavailable music, and discuss how an incremental approach can be supported, before evaluating the approach in subsequent sections.

3. SUPPORTING RESEARCH WITH INCREMENTAL AND INCOMPLETE CORPORA

Musicology, and indeed research more broadly, may involve many activities and strategies, whose selection will be informed both by research topic (see [2]) and the stage at which the research stands. For example, a researcher may start with a literature exploration, then start reviewing music scores through a catalogue, selecting a set of potential subjects to look at more closely and then focus down later. The researcher may scan through the scores, selecting works or passages for further consideration, and rejecting others. This might be followed by closer engagement with the chosen texts, often relating them to extra-musical information. Finally, their investigations will be written up formally.

Teasing apart the steps of this example, and when they are most likely to happen in a research life cycle, we can see the following:

1. Literature exploration (early phase)
2. Catalogue exploration (early phase)
3. Workset selection (early phase)
4. Content exploration (mid phase)
5. Content analysis (mid phase)
6. Connecting music with extra-musical material (mid phase)
7. Visualisation and reporting (end phase)

This is not intended as a complete catalogue of research steps, but illustrates common components, and helps ground our observations. Each of these steps will decompose into tasks that may or may not be carried out

Phase	Step	Example activity	Example tools and media
Early	1. Literature exploration	Makes up-to-date literature survey	RILM, JSTOR Google Scholar Physical browsing
	2. Catalogue exploration	Explores the repertory; identifies a superset for more attention	Library catalogues RISM IMSLP, CPDL Physical browsing
	3. Workset selection	Looks at the music, scans through scores to identify works or passages for detailed consideration	RISM IMSLP, CPDL Specialised corpora Physical sources Image digitisation
Mid	4. Content exploration	Close reading of scores, identifying distinctive parameters that support an emerging thesis	OMR Measure detection Sonic visualiser Piano
	5. Content analysis	Lists spacing and instrumentation of chords at cadences	Humdrum toolkit Music21 Sonic visualiser Spreadsheet Paper
	6. Making Connections	Associates particular orchestration approaches with review and theory texts	Spreadsheet Paper
End	7. Visualisation & reporting	Writes and publishes a journal article	Journal, Published edition Recording, Dataset

Table 1. A typical set of steps in a research lifecycle, with example activities and tools. Although this appears as a list, scholars may jump between these, or pursue several at the same time. The **Beethoven in the House Annotator** supports stages 4 and 5, producing data suitable for stage 7.

in a digital environment, and although broadly sequential, a musicologist may jump backwards at any point to supplement the data they already have. To support such flexible research patterns, we believe it is important to create an ecosystem of tools that read or write compatible data, facilitating researcher-directed methodologies for tool selection and task ordering.

4. THE BEETHOVEN IN THE HOUSE ANNOTATOR: A TOOL SUPPORTING MID PHASE RESEARCH

To investigate the feasibility of an 'incremental' interaction model with MIR tools, we have developed a tool to support an active musicological investigation which also embodies the 'mid phase' of the research life cycle described above, focussing particularly on steps 4 and 5. The tool's main purpose is to bring together digitised resources in the form of images and digital scores, and allow a musicologist to view them in a browser, selecting specific extracts for study and then annotating those with scholarly commentary. The resulting annotations are stored, and can be shared and published, including references to the pertinent selections from the digital music resources.

A user entering the **Beethoven in the House Annotator** first selects items to explore in a 'library' view – a listing which displays metadata about available musical works, their arrangements and digital resources available. Because the annotator is designed to handle comparisons of the same passage of music as it is realised in different versions, the selected resources are displayed one above another to aid analysis.

Once works are selected and loaded into the display pane, a musicologist can point and click on individual notes and measures, or click and drag to select larger regions, whether the resource is a facsimile image or a rendered score encoding. Individual selections can be annotated, but also parallel passages in different versions of a work ('Musical Material') can be identified (figure 1, left), and these structures themselves annotated (figure 1, right).

Previous annotations can also be viewed and themselves annotated. Thus, the tool can be used for quick browsing or juxtaposition of music and metadata and for detailed labelling of the content.

The **Beethoven in the House Annotator** is built as a web application, and implemented as a MELD (Music Encoding and Linked Data) application [9] [1] . As a baseline provision we assume the materials underpinning the musicologist's investigation are available in image form via IIIF [2] as welll as digital MEI scores when these are available. We further assume that the musicologist has the tools and skills to optionally transcribe whole pieces or extracts and save or convert them as MEI (this can be carried out using music typesetting packages such as Sibelius or Mus-

[1] More precisely, we use data models and the graph traversal library from MELD, with Vue-based application code.

[2] The International Image Interoperability Framework provides for standardised image delivery through APIs for content and presentation. Although increasingly widespread in use by research collections in particular, it is not yet comprehensively adopted. For the purposes of our research project, required digitized images were provisioned via a local (private) IIIF server where they were not already available over IIIF from the holding collection.

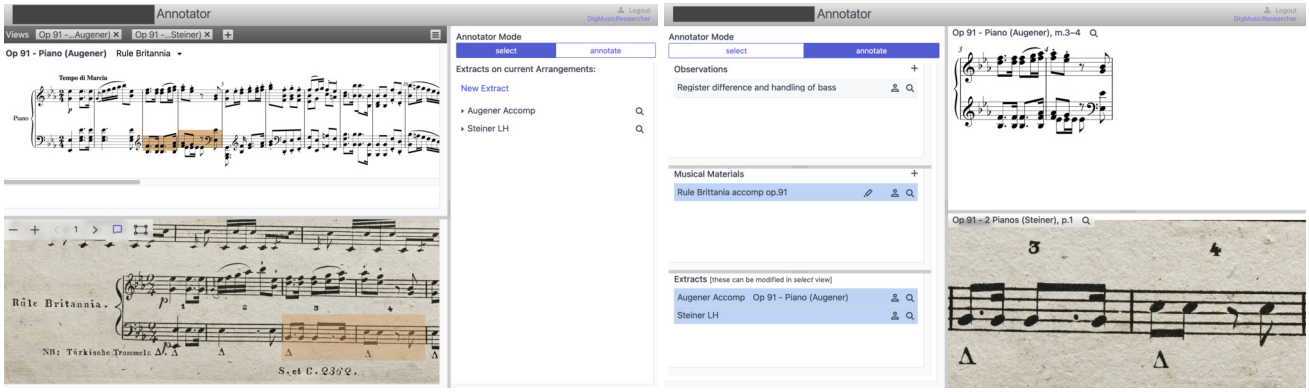

Figure 1. Two screenshots from the **Beethoven in the House Annotator**. **Left:** 'Musical material' – Parallel passages recorded as occurring in two different arrangements of Beethoven's *Wellington's Sieg* (Op. 91). Selections need not be contiguous or limited to a single part. The upper version here has been retrieved from an MEI file and is displayed using Verovio. The lower version is from a IIIF file for which measure locations have been separately detected using the Cartographer tool and stored, along with links to the image, in an otherwise minimal MEI file. **Right:** The annotation view, showing an observation recorded about the musical material shown left. In both cases, structures are saved to the musicologists personal Solid pod, with their login shown upper right.

eScore with the help of plugins).

We also assume the prior existence of well-formed and self-describing catalogue metadata, and we base our prototype on the Linked Open Data published by the Gemeinsame Normdatei (GND) of the Deutsche Nationalbibliothek. We do this with the intention that these could in future be loaded directly where records exist. [3]

Direct image annotation is possible within our tool. The musicologist may prefer to use a labour-saving measure detection tool, such as Cartographer [4] or MEI Friend [10], both of which can output MEI with empty measures and image co-ordinates, and which have been successfully tested with our tool. When provided with MEI and IIIF resources such as these, our annotation tool allows the researcher to annotate the image measure by measure – giving a semantically-richer anchor for the annotation with relatively low input of manual intervention (see the lower pane in figure 1, left). If the researcher needs a finer level of annotation, then they may fill in additional music notation in the MEI, and can indicate the selective nature of the encoding in the MEI header, a process supported by tools such as MEI Friend.

Our application supports textual Web Annotations [11] made onto conceptually abstracted musical extracts rather than directly onto elements or regions of the image or encoding, allowing parallel material occurring in different arrangements of a work to be annotated together and, at a more basic level, allowing the model to remain agnostic to the different types of media used as evidence (figure 1, right, illustrates an example of an annotation on a passage that has been identified in two arrangements, in one case using the MEI transcript, and in the other a IIIF image after a process of measure detection). This uses the Music Annotation Ontology described by Lewis et al [12].

In order to promote data sharing between tools rather than a single monolithic application, user data is stored as Linked Data in Solid Pods [13], distributed online data storage with fine-grained access control, and for which the user can choose provider. This provides a simple mechanism for data portability between applications, given compatible data structures. The structures written can refer to resources anywhere on the web, and traversal carried out by the MELD library will draw them into the application.

In summary, the **Beethoven in the House Annotator** described above supports our proposed workflow in several ways. Firstly, it is conceived as part of a pipeline of tools publishing compatible Linked Data and MEI, and is already interoperable with existing tools. Secondly, it is intended to provide a low barrier for including evidence materials, allowing the use of any web-published IIIF images, complete or partial MEI files, and GND metadata rather than requiring extensive data entry and local servers. Thirdly, it supports the sharing of source data and metadata, along with intermediate observations, within a research team. Finally, as currently implemented, annotations are minimally structured. This supports an evolving research agenda, trading expressiveness against semantic structures.

5. EVALUATING THE BEETHOVEN IN THE HOUSE ANNOTATOR AND ITS WIDER APPLICATION

Whilst the tool's internal development was aimed at satisfying researcher needs within our own project, two rounds of wider evaluation were carried out, timed to coincide with two phases of application development. These evaluation rounds were carried out as semi-structured interviews following shortly after a combination of a presentation about the Annotator and period of time freely ex-

[3] In practice, the GND is not currently usable for client-side applications due to access control headers. This would still allow the use of a server-cached version of the data. Where other metadata is needed, we draw on the WikiData model.

[4] `https://cartographer-app.zenmem.de`

ploring its functionality over a pre-loaded musical library. In the first round interviews were conducted with musicologists recruited via a Studienkolleg (summer school) located at Beethoven Haus, Bonn, in September 2022. In the second round in March 2023, volunteers from staff at the Beethoven Haus were interviewed. In the first round, we interviewed 9 scholars, and 7 in the second round, of whom 2 had previously been interviewed. This allowed us to assess progress with new and returning users.

5.1 Workflow as data pipeline, low barriers for evidence gathering

The application was regarded by all interviewees as useful in the context of larger musicological research projects and editorial work. Since our interviewees were musicologists and editors rather than engineers and, since we did not demonstrate or present any tools for other steps in the process, this support is based primarily on the interviewer's description of the intended wider context for the app rather than concrete experience. Interviewees did raise important concerns regarding the workflow itself, and these are discussed in 5.5 below, and as further work.

5.2 Sharing of evidence and findings

Users that we spoke to were strongly attracted both by the idea of sharing data and annotations and the option of keeping these private or controlling access – either during the research process or separating draft and publishable work. They immediately identified the equivalence of this approach to paper based methods of publication and regarded using publicly shared annotations as "similar to quoting published books", although there are concerns about how to verify and attest its quality. It is clear that these features would be easier to realise given user interfaces optimised for these tasks, since the default management interfaces of Solid providers, our principle medium for publication, generally present usability barriers to newcomers. Nonetheless, one user evaluates that the application has the potential to "bring everything together in a way I haven't experienced before" in terms of gathering and sharing knowledge about musical works. This would support the "Nachprüfbarkeit", or verifiability (literally reviewability) of a conclusion by collecting the evidence in a single place.

Although musicology can appear – at least from its outputs – as the activity of lone scholars, sharing between scholars in an informal way is common, as is the use of student assistance, both of which can benefit from controlled data sharing. Certainly, several participants were explicitly open to a wider set of contributors, one noting that, depending on the quality of the community, "more knowledge can be obtained". Beyond this, other musicological use cases identified by participants are more commonly team or group activities, such as scholarly music editing or pedagogical uses, with sharing either between teacher and student or between students within a class.

This sharing approach is well supported for our own Linked Data structures, but there are concerns with the boundaries of that sharing. For example, a Linked Data structure that is publicly shared could annotate a part of an image or score that is not itself publicly available (perhaps for copyright reasons). This would not render the information in the Linked Data unusable, and the URI itself would remain uniquely identifiable, but for some uses would become unavailable. There is no clear way to deduce that one identified element in an MEI file occurs earlier in the piece than another purely from the URI since these semantics are located in the MEI score. Our use of the Music Annotation Ontology brings more aspects of musical selection into the Linked Data domain, but we do not attempt to export musical meaning encoded in MEI into RDF.

5.3 Minimally-structured annotations

The open nature of the annotations and the **Beethoven in the House Annotator** more generally was very clearly important in allowing the musicologists to identify a wide range of contexts in which it would be useful to them. These covered the full range from studying stages in the development of a particular music edition ("Plattenstadien"), systematic musicology, historical approaches, philology and pedagogy. Participants also identified the ease of linking material, both music and annotation material, which is evidence that our low structure approach may have reduced barriers to use. Beyond this (sometimes implicit) validation of our approach, participants identified some structures in annotations to support navigation and discovery.

In the **Beethoven in the House Annotator**, annotations are edited and viewed separately from the score view. In our first version, this view was purely textual, making them harder to navigate, and placing a strong reliance on user-provided labels. Adding musical previews for the second version enhanced findability, but multiple participants noted that an informal taxonomic labelling, such as tags, would enhance this – especially where annotations are shared between users. Data currently available to the application includes metadata and musical locations (where annotations are made on transcribed sources or images on which measure detection has been run). These are not currently used in the annotation listings, but could be, allowing the navigation by measure and source requested by several participants.

5.4 Application-specific responses

The **Beethoven in the House Annotator** builds on rich underlying data models and a complex range of data sources and technologies. An aspect that emerged from the interviews is that the terms chosen for defining key elements in the model did not translate well when designing a user interface. Most users had difficulty navigating the application because their expectations about the terms used did not match with the meaning given to them in the context of the model. Although the learning curve can be conquered, the interviewees expressed that substituting and simplifying the language (in certain cases, hiding structures) would be more beneficial to a quick acclimation into

the application. Although the general-purpose nature of the tool makes the choice of task-specific language difficult, use of clear domain-specific and task-appropriate terms would have been better received by the musicologists and required less detailed briefing.

Although the functionality of the **Beethoven in the House Annotator** is distinctive in ways that were recognised and appreciated, those participants who have worked with comparable applications commented on affordances that they missed from the other tools. In particular, familiarity with EDIROM tools left some participants missing the more advanced navigation system, with, for example, jumping to measure numbers.

5.5 The workflow outside the application

Our workflow acknowledges the poverty of encoded scores but does not, currently, accommodate the lack of digitised images. These, too, have been created according to particular priorities, which may not reflect those of researchers. Libraries and archives must weigh up rarity, value, appearance, physical condition, use and public impact among many other factors when deciding their digitisation policy. Interviewees expressed particular concern for sources located in institutions for whom the burden posed by digitisation in the first place and publication as IIIF in the second is too great, while private collectors may have no desire to engage in digitisation at all. Even following the suggestion of one interviewee, and supporting user upload of static images – whether to their own Solid Pods, or some public IIIF server operating for the common good – could fall foul of institutional restrictions. This may indicate a need to point our structures at musical regions even where no digital proxy exists at all, something which would require a semantic representation of musical location. Although some progress has been made towards such a representation (see, for example, [14]), further modelling is needed to make a robust system.

Similarly, interviewees speculated about how additions are made to the library that the application presents. Currently, we have no application to support the selection of items from a published catalogue to create and operate on a selected workset (steps 2 and 3 in table 1) or the discovery and data transformation this would require. This has not been the focus of the current research, but it does mean that we have relied on some manual technical interventions that would be unsuitable for the sort of musicologists we target here.

6. CONCLUSIONS AND FURTHER WORK

The research workflow we describe here is one in which a scholar adds and edits data and metadata, and in which research priorities develop throughout. We assert this is expresses, albeit schematically, a common approach in musicological research. Rather than trying to create tools to manage the whole process, we have advocated for smaller tools that can comfortably handle mixed, incomplete and partial data, and accumulate results in a way that is data-compatible with other applications and IR tools that the researcher might use.

The musicologists interviewed identified a range of contexts for the **Beethoven in the House Annotator**. That these went not only beyond our design for it, but also beyond its capabilities provides evidence of the need for and dearth of tools that support such activities and the diversity of approaches that can and should be considered.

Our interviews also point clearly to further work, with early-phase support – in the form of digitisation, search and retrieval, and workset gathering – being priorities that would help researchers prepare their materials for use with mid phase applications such as our own. Candidates for components of such tooling, such as Cartographer, but also Sonic Annotator and MEI Friend, often already exist, and often have elements that directly support their role in an ecosystem of tools, particularly in terms of data compatibility, but are often seen either as entirely standalone tools or built into workflows in task-specific ways that are not generalised.

Our investigation demonstrates that the workflow into which the **Beethoven in the House Annotator** fits is recognised and valued by musicologists. The flexibility of the annotator tool, in terms of data and functionality, presents many opportunities in support of musicological research. Importantly, we show that it supports or replaces activities currently taking place in forms – such as Word documents, spreadsheets or on paper – that provide few opportunities for scholars to take advantage of either MIR tools in data analysis on the one hand or digital transparency and sharing of results on the other. Thus, it is recognised as going beyond reproducing existing methods, by enhancing and extending them.

7. ACKNOWLEDGMENTS

This research was undertaken by the project ‘Beethoven in the House: Digital Studies of Domestic Music Arrangements’, supported by a UK-Germany funding initiative: in the UK by the Arts and Humanities Research Council (AHRC) project number AH/T01279X/1 and in Germany funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project number 429039809; with additional support from the UK Software Sustainability Institute Phase 3, funded by the UK Engineering and Physical Sciences Research Council (EPSRC) project number EP/S021779/1.

8. REFERENCES

[1] M. J. Bates, “The design of browsing and berrypicking techniques for the online search interface,” *Online Review*, vol. 13, no. 5, pp. 407–424, May 1989. [Online]. Available: https://www.emerald.com/insight/content/doi/10.1108/eb024320/full/html

[2] D. M. Weigl, K. R. Page, P. Organisciak, and J. S. Downie, “Information-seeking in large-scale digital libraries: Strategies for scholarly workset creation,”

in *2017 ACM/IEEE Joint Conference on Digital Libraries, JCDL 2017, Toronto, ON, Canada, June 19-23, 2017*. IEEE Computer Society, 2017, pp. 253–256.

[3] K. Fenlon, M. Senseney, H. E. Green, S. Bhattacharyya, C. Willis, and J. S. Downie, "Scholar-built collections: A study of user requirements for research in large-scale digital libraries," in *Connecting Collections, Cultures, and Communities - Proceedings of the 77th ASIS&T Annual Meeting, ASIST 2014, Seattle, WA, USA, October 31 - November 5, 2014*, ser. Proc. Assoc. Inf. Sci. Technol., vol. 51, no. 1. Wiley, 2014, pp. 1–10. [Online]. Available: https://doi.org/10.1002/meet.2014.14505101047

[4] S. Oberbichler, E. Boroş, A. Doucet, J. Marjanen, E. Pfanzelter, J. Rautiainen, H. Toivonen, and M. Tolonen, "Integrated interdisciplinary workflows for research on historical newspapers: Perspectives from humanities scholars, computer scientists, and librarians," *Journal of the Association for Information Science and Technology*, vol. 73, no. 2, pp. 225–239, 2022. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.24565

[5] A. Kijas, "What does the data tell us?: Representation, canon, and music encoding," 2018, Keynote at Music Encoding Conference, Maryland, 24 May 2018. [Online]. Available: https://medium.com/@kijas/https-medium-com-kijas-what-does-the-data-tell-us-926ba830702f

[6] F. Wiering, "User needs and challenges in digital musicology," 2014, Digital Music Lab Workshop on Analysing Big Music Data, City University London, 19 March 2014. [Online]. Available: https://webspace.science.uu.nl/~wieri103/presentations/WieringLondonDigitalMusicLabFinal.pdf

[7] N. Cook, "Towards the compleat musicologist?" 2005, Invited Keynote for the 6th International Conference on Music Information Retrieval (ISMIR), London 2005. [Online]. Available: https://ismir2005.ismir.net/documents/Cook-CompleatMusicologist.pdf

[8] C. Inskip and F. Wiering, "In their own words: Using text analysis to identify musicologists' attitudes towards technology," in *Proceedings of the 16th International Society for Music Information Retrieval Conference, ISMIR 2015, Málaga, Spain, October 26-30, 2015*, 2015, pp. 455–461.

[9] D. M. Weigl and K. R. Page, "A framework for distributed semantic annotation of musical score: "take it to the bridge!"," in *Proceedings of the 18th International Society for Music Information Retrieval Conference, ISMIR 2017, Suzhou, China, October 23-27, 2017*, S. J. Cunningham, Z. Duan, X. Hu, and D. Turnbull, Eds., 2017, pp. 221–228. [Online]. Available: https://ismir2017.smcnus.org/wp-content/uploads/2017/10/190_Paper.pdf

[10] D. M. Weigl and W. Goebl, "Alleviating the last mile of encoding: The mei-friend package for the atom text editor," in *Proceedings of the Music Encoding Conference, Alicante, 2021*. Humanities Commons, 2022. [Online]. Available: https://hcommons.org/deposits/item/hc:45977/

[11] R. Sanderson, P. Ciccarese, and B. Young, "Web annotation data model," W3C, W3C Recommendation, Feb. 2017, https://www.w3.org/TR/2017/REC-annotation-model-20170223/.

[12] 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 *DLfM '22: 9th International Conference on Digital Libraries for Musicology, Prague Czech Republic, 28 July 2022*, L. Pugin, Ed. ACM, 2022, pp. 10–18.

[13] D. M. Weigl, W. Goebl, A. Hofmann, T. Crawford, F. Zubani, C. C. S. Liem, and A. Porter, "Read/write digital libraries for musicology," in *7th International Conference on Digital Libraries for Musicology, Montréal, Canada, October 16, 2020*. ACM, 2020, pp. 48–52. [Online]. Available: https://doi.org/10.1145/3424911.3425519

[14] R. Viglianti, "The music addressability api: A draft specification for addressing portions of music notation on the web," in *Proceedings of the 3rd International Workshop on Digital Libraries for Musicology*, ser. DLfM 2016. New York, NY, USA: Association for Computing Machinery, 2016, p. 57–60.