# A model for annotating musical versions and arrangements across multiple documents and media

David Lewis
david.lewis@oerc.ox.ac.uk
University of Oxford e-Research
Centre
Oxford, UK

Elisabete Shibata
shibata@beethoven.de
Beethoven-Haus Bonn
Bonn, Germany

Mark Saccomano
mark.saccomano@uni-paderborn.de
Paderborn University
Paderborn, Germany

Lisa Rosendahl
lisarosendahl@gmx.com
Paderborn University
Paderborn, Germany

Johannes Kepper
kepper@edirom.de
Paderborn University
Paderborn, Germany

Andrew Hankinson
andrew.hankinson@rism.digital
RISM Digital Center
Bern, Switzerland

Christine Siegert
siegert@beethoven.de
Beethoven-Haus Bonn
Bonn, Germany

Kevin R. Page
kevin.page@oerc.ox.ac.uk
University of Oxford e-Research
Centre
Oxford, UK

## ABSTRACT

We present a model for the annotation of musical works, where the annotations are created with respect to a conceptual abstraction of the music instead of directly to concrete encodings. This supports musicologists in constructing arguments about musical elements that occur in multiple digital library sources (or other web resources), that recur across a work, or that appear in different forms in different arrangements. It provides a way of discussing musical content without tying that discourse to the location, notation or medium of the content, allowing evidence from multiple libraries and in different formats to be brought together to support musicological assertions.

This model is implemented in Linked Data and illustrated in a prototype application in which musicologists annotate vocal arrangements of the Allegretto from Beethoven's Seventh Symphony from multiple sources.

## CCS CONCEPTS

• **Applied computing** → *Sound and music computing*; **Arts and humanities**; *Digital libraries and archives*; • **Information systems** → *Multimedia information systems*.

## KEYWORDS

digital musicology, conceptual modelling, linked data, FRBR, music arrangements

## 1 INTRODUCTION

Much musicological research relies for evidence on documents from libraries and archives. These documents may come in different forms, and in various states of digitisation or transcription, and may be distributed geographically. Increasing digitisation of these resources makes their juxtaposition for comparison easier, but tools for supporting such research remain generic and cumbersome for the purpose. Furthermore, since information about observed difference is a significant component of the research, it is important to be able to share it, and in a reusable way.

The more of this research process that happens in a digital realm (rather than, for example, by printing resources and laying them out on a table), the more the observations made can be explicitly used to support the musicologist's further research, but also, the easier they are to capture for preservation, publication and dissemination. To be useful in this way, the core semantics of the observations and assertions being made must be clear and well-defined, even if aspects of their content remain human-readable only, in the form of prose.

In this paper, we derive a set of functional requirements for systems supporting investigations where multiple musical resources are compared, aligned and annotated. We consider where existing work satisfies these requirements before describing our own model. To demonstrate the model's viability, we introduce a simple, distributed application for comparing multiple arrangements of the same musical work – as images or encodings of notated music – that

implements a subset of the model, reading and writing annotations to a user's shareable online storage.

## 2  BACKGROUND

### 2.1  Motivating scenario: studying the practice of editing and arranging

We introduce this musicological scenario, which is an area of active research, to illustrate the wider motivations for our technical contributions which follow and to provide the use case for our prototype implementation.

Until the advent of audio recordings and radio broadcasts, many people encountered new works more as performances at home or with friends than at the concert hall or opera house. This audience for music in a domestic space demanded new compositions, suitable for the scale, skills and forces available, but they were also enthusiastic consumers of adaptations of works originally conceived for larger public performance. Despite the central role these activities played in musical culture for centuries, domestic music has remained under-explored, compared to that of concert and opera. The music is infrequently digitised, and cataloguers often struggle to provide clear records of arrangements and albums.

Arrangements of concert music were frequently reviewed by critics in the press or in specialist journals, while composers often took great care over arrangements of their works (including Beethoven, who insisted on rewriting his publisher's original arrangement of his *Große Fuge*[8, pp. 227-8]). The financial and reputational importance of these editions was considerable, and yet little attention has been given them as music, whether in academic or performance circles[1]. Besides the obvious difference in status between professional and amateur music, one clear reason for this difference in attention is that comparing versions of the same work is very labour intensive, and the volumes containing them, which have often been of low prestige within musicology and may have had quite localised print runs and sales, can be widely distributed amongst libraries and private collections. Although the challenge of comparing music in multiple documents has been made easier by the possibilities of digitisation and the Internet, the strong musical similarities and the scale and complexity of the works involved makes studying these questions time consuming.

Arrangements were not just reductions that aimed (and were promoted) for fidelity, bringing the experience of the large-scale work into a smaller space. Many drew on well-known works, to make them easier or shorter, sometimes recontextualising by introducing titles, verses and tempi (see Figure 1 for an example of these). They might even be transformed more dramatically, for example turning a symphony into a song or an aria into a virtuosic rhapsody. Other arrangements brought attention to more obscure music that would be less likely to reach the audience through their local concert halls – this includes the *Große Fuge* that appears to have sold well when arranged as a piano duet, despite very few public performances in its original form. In all these forms of adaptation, active and creative decisions are being made by the arranger or

---

[1]Exceptions in academic writing include scholarship that considers arrangements as evidence for performance practice (e.g. [7, 9], as well as those concerning historical, technical and social aspects (e.g. [8, 13]).



**Figure 1: An abridged arrangement by Jules de Sivrai (a pseudonym of Jane Roeckel, 1834-1907) of the Allegretto from Beethoven's Seventh Symphony, Op. 92, given a picturesque title and some accompanying verse (private collection Luigi Bellofatto, copy in Beethoven-Haus Bonn, C92 / 50).**

editor, and these will reflect the perception of the source work and its audience, as well as the skills and labour of the arranger.

### 2.2  Gathering and annotating musical evidence

When looking at different versions, arrangements or reconceptions of what is, on some level, the same work, it is important to be clear about how the different instances of the music relate to one another and, when talking about a feature of the music, which versions have it in common.

Most attempts to model the relationship between versions of an intellectual product derive from the Functional Requirements for Bibliographic References (FRBR)[11]. This provides not only the ability to group instances together as resulting, directly or indirectly, from the same intellectual endeavour, but it also provides a model for abstract intellectual products, Works, that get progressively more concrete as Expressions and then Manifestations, down to Items as individual products of Manifestation reproduction. This conceptual framework is extremely helpful, but is applied almost exclusively at the level of the complete piece or library volume.

The Music Encoding Initiative format[1] for representing music notation can also structure its metadata using FRBR, and it has some cross-reference facilities for grouping multiple files, along with a text-critical module that supports encoding multiple sources in a single edition. Other functionality provides for relating recordings and digitised sources to the encoded music. Thus, MEI can be used as a hub format for organising related music.

Given the ability to draw relationships between pieces of musical evidence, we then need to be able to record observations in a way that cannot be prescribed, since the annotations are likely to be novel research. Annotation of music in some form or another has a long history in digital musicology, but has been an increasing focus of research in recent years, with examples including Dezrann[3], the Winterreise Dataset[16], CRIM[2] and work built around the MELD (Music Encoding and Linked Data) framework ([5] for example). One standard that underlies several of these is the Web Annotation ontology for Linked Data ([10]), which provides structures for annotating a Target (which can be anything with a URI) with a Body (the annotation itself in whatever form it takes, including arbitrary text). Annotations can include information about motivation, and record who or what created them.

Whatever form an annotation takes, it must have a target – some circumscribed area of interest within one or more documents. Various strategies may be used to achieve this. Depending on the format of the document, it may be possible to indicate that area within the document itself (for example, using a Section element with an xml:id in MEI). This is not always possible and, even where it can be achieved, it requires the annotator to have write access to the document or to copy and republish it. Another method for XML-based formats such as MEI is to use the xml:ids of all the elements of interest as fragment identifiers of a URI. Gathering all these together allows any combination of elements to be targetted.

More broadly, media files such as images and audio can have regions targeted using the media fragment standard[2], which allows time and spatial co-ordinates to be used in URIs; while the IIIF APIs also support similar selections. For music notation, EMA[12] can provide an encoding agnostic way to refer to parts of documents by bar, beat and voice (providing the notation and its representation supports these concepts in a consistently retrievable way). EMA also allows regions to be combined, for example to highlight a figure as it moves around an orchestral texture.

Even given these available technologies, there are other challenges for modelling annotation of music accurately. In several arrangements of the Allegretto from Beethoven's Seventh Symphony, for example, the opening chord (present in Figure 1) is omitted. What should an annotation target when music or musical symbols are absent?

It is also important, especially where annotations may be exchanged between applications, to be clear exactly what purpose the music being annotated serves. For example, in the string quartet masterclass described in [6], annotations of the score refer specifically to practical notational and performance suggestions made at the time the annotation was made. They apply very directly to that score – which was projected on a screen during the masterclass and on which the annotations were visible. By contrast, in the digital companion to a musicological study of Wagner's use of motifs described in [5], the annotations refer not to the symbols in a single edition but, more broadly, to the work. In almost all cases, we would expect the same motifs, in the same orchestration, to appear in all editions of this work (at least all those based on the same Expression).

It will be clear by this point that annotation, especially where it covers multiple editions or arrangements is a complex concern, with multiple issues to resolve. Before we describe our own model, we clarify the functional requirements of an annotation system, derived from our experience and specific use cases. This allows critique of these separately from our modelling and implementation.

## 3 FUNCTIONAL REQUIREMENTS DERIVED FROM MUSICOLOGY

Although different musicological investigations will necessarily have different, and specialist, needs, we focus on what we consider a superset of functional requirements optimised for research that works with juxtaposing different, but related, pieces based on documents that come in different modes of presentation and formats. While it is unlikely that any single use case will require all of these,

we would expect some degree of common ground between many of them.

**R1. Evidence resources**  The system must be capable of referring to evidence in the form of (a) fully- or partially-encoded scores; (b) images with musical content; (c) audio; (d) video.

  **R1.1**  The system must be extensible to accommodate non-musical materials or musical materials in forms other than CMN.

  **R1.2**  All items must be addressable (having a stable identifier) and dereferenceable (it must be possible to convert that identifier to a form of the resource).

Since it would be unnecessarily onerous to be required to prepare a full edition of an entire work simply to make a comment about a single passage, we must be able to accommodate partial encodings (where only the areas of interest are provided) and images. Audio and video materials are of less interest for the narrower enquiries being used as use cases by the authors, but are of considerable importance to the wider community. We would also expect most musicological enquiries to work with materials beyond those that directly include sound or music notation. This is beyond the scope of the current work, but we add a subsidiary requirement, R1.1, here. Since these resources are to be used as evidence, it is important that they can be identified and, depending on access restrictions, interrogated (R1.2).

**R2. Reference**  The system must be capable of addressing arbitrary components of evidence resources.

  **R2.1**  The system must be able to refer to a subject of interest without requiring alteration to the source material.

  **R2.2**  If the subject of interest being considered is split across multiple resources, it must be possible to address it as a single object.

  **R2.3**  The system must be able to reference an absence, including the location where the subject of interest is absent.

Almost any partition of a primary source may be of interest. The discussion may be purely notational – based, for example, on the transformation of a single dynamic marking from $fff$ to $ff$ – or it might be more music-analytical – perhaps considering a single melodic contour that is passed from voice to voice, referring to a different bar from each instrument in the orchestra.

Many resource formats, including MEI and IIIF, allow the publisher to explicitly demarcate regions of interest. This can be useful, but would either require the musicologist to work with the publisher or to copy the resource, neither of which is likely to be practical in most cases. For this reason, we add a requirement (R2.1) that any partition of or reference to resource can be published independently of the resource.

The remaining two requirements, R2.2 and R2.3, cover common practical issues. Cases covered by R2.2 will be particularly common for music that is printed in parts, especially where the parts are not digitised as a single entity, or if they are distributed between institutions. Decisions by arrangers of what to include and what to omit are of considerable interest, and can only be articulated in machine-readable form if absence can be explicitly represented (R2.3).

**R3. Grouping**  The system must be able to refer to the subject of interest with a single identifier. Subjects of interest can include:

---

**R3.1** A partition of the music as it appears in a single form (e.g. bar 8 of the violin and viola part of the first movement of Steiner's string quartet arrangement of Beethoven's Eighth Symphony Op. 93, published in Vienna in 1816).

**R3.2** A partition of the music as it appears in a single version, but in different forms (e.g. the same bar of the arrangement, but as an image, a digital edition and a recording).

**R3.3** Partitions of different version of the same music, referring to parallel content (e.g. bar 8 of several arrangements of the first movement of the Eighth Symphony).

**R3.4** Partitions of the music, referring to recurring material in the same work (e.g. the opening theme of a sonata recurring at the beginning of the recapitulation).

These particular groupings were identified as important in our work with use cases: an extract of a single version of a work (as it appears in one resource – R3.1 – or multiple resources – R3.2); an extract of notationally 'the same' extract from multiple versions (R3.3); and, perhaps more familiar, an extract that returns motivically in the same work (R3.4).

**R4. Annotation** The system must be able to apply annotations to any subject of interest.

**R4.1** An annotation must be able to record arbitrary textual comments.

**R4.2** The system must distinguish between annotations that associate information with an object of interest (or the same information with multiple objects) and those that are concerned with the contrast between objects of interest, drawing attention to the differences between them.

**R4.3** The system must distinguish description (making an attribute in the source material explicit in the system) from commentary (explaining or justifying that attribute).

Annotations can refer single or multiple subjects of interest; in some cases, multiple subjects will all share a characteristic being described, and in other cases the annotation will set out to juxtapose them, drawing attention to differences. The difference between these must be clear.

**R5. Responsibility** The system must be able to record responsibility information (information provenance) for all annotations and identified subjects of interest.

Since any statement in this model is an assertion, it should always be possible to record who created it and when.

## 4 ASSESSMENT OF PRIOR WORK

In this section, we relate prior work to these functional requirements. Since the requirements are specific to a range of musicological applications, a mismatch between these approaches and our requirements does not invalidate them, although it does indicate where they would not currently satisfy our needs or, perhaps, a gap in their modelling.

Dezrann has sophisticated selection and annotation functionality, but little contextual metadata. It allows annotators to attach labels to scores (R1a) and audio waveforms (R1c) and, although the system architecture currently appears to require evidence resources to be stored and published with the annotations, the annotations

themselves are separable (R2.1). Dezrann labels can apply to multiple resources that share a common timeline, so can be applied to a score and its recording (R3.1, R3.2). Nonetheless, annotations appear to apply to a single point or region in the music, rather than multiple arbitrary ones (R2, R2.2), while absence could be indicated using a point label only with semantic extensions (R2.3). The idea of versions or related resources does not appear to exist natively in the system, so although common patterns (R3.4) can be recorded, parallel content (R3.3) can not.

Since Dezrann labels are typed, the nature of an annotation could be recovered, though this is not explicit and arbitrary text seems to be permitted (R4[3]). Information provenance does not appear to be recorded (R5).

The Winterreise Dataset is, by its nature, less concerned with infrastructure and more with reusable data. Annotations can refer to score positions, which are then related to scores (R1a), images (R1b) and recordings (R1c), with some structure provided to relate these (R3.2). As with Dezrann, little structure is provided to relate or group instances or to clarify their relationships (R3). Since annotations are simply separate CSV files, there is no alteration to evidence resources (R2.1). A time-slice based format means arbitrary segmentation and grouping of resources is not possible (R2.2, R2.3) and, although without prescribed syntax, any semantics of annotation *could* be recorded, the format is largely used for formal and analytical labels (R4).

The CRIM Project (which at the time of writing is under active development) provides tools for structural and motivic analysis and annotation of polyphony, combined with intertextual labelling of musical citation. Since the project is ongoing, assessment can not be made of a final published state. Annotations use the Web Annotation standard and can target arbitrary selections in a musical encoding without altering the resource (R1a, R2.1, R3.1). Extension beyond music notation (R1.1) is planned.

It does not appear to be possible to create a single CRIM reference across multiple resources (R2.2) or directly indicate absence (R2.3) and, rather than modelling a subject of interest as a single entity across multiple versions (R3, R3.1-4), the vocabularies annotate precise transformations that occur between versions, such as quotation, melodic inversion, etc.

CRIM annotations support arbitrary text labelling and record responsibility (R4, R4.1, R5). Since the system is designed for music-theoretical labelling, the question of annotation types (R4.2, R4.3) are beyond the scope of the project.

Applications built using the MELD framework are diverse in terms of materials and purpose, and so we consider two examples. The Delius quartet masterlass used RDF as to annotate an MEI score (R1a), and audio (R1c) and video (R1d) recording, with only partial use of shared timelines (since the score was not used linearly from start to finish in the masterclass). This approach does not alter published resources (R2.1) nor require them to be published on the same server. Temporal points and regions were used, as were points and regions of the score, but the user interface involved limited these in a similar way to Dezrann, to vertical slices of score or voices (R2, R2.2). Because of the fixed nature of the score, absence

---

[3]This judgement depends on the structure of the label's 'type' field, but we believe could be interpreted in this way.

is not considered (R2.3). This application takes each resource as separate, so no groupings are explicitly coded (R3) and, although web annotations are used throughout, which can support arbitrary content, the application operates with a heavily prescribed range of annotations, (R4). Web ontology semantics are used on a basic level in the application to record information provenance (R5).

The Lohengrin TimeMachine, built on the same MELD infrastructure, allows arbitrary stand-off selections (R2.1, R3.1) of MEI, audio and video (R1a, c and d), grouped using FRBR-derived classes that support grouping multiple embodiments of an extract (R3.2). A 'motif' class defined as a subclass of a FRBR work is used to support reference to varied but recurring content (R3.4). Versions are not considered (R3.3). Absences (R2.3) and divided resources (R2.2) are not considered. Textual and object-based commentaries are supported (R4, R4.1) using web annotations, but no classification of annotations is made (R4.2, R4.3), nor is any record of their originator (R5).

## 5 A NEW THREE-LAYER MODEL FOR MUSIC ANNOTATION

The functional requirements for reference (R2) and grouping (R3) depend on a separation between a resource and the musical ideas that are exemplified there. This separation is realised in a new model structured around a three-layer abstraction (see Figure 2). The two outer layers of this model are an explicit statement of structures seen in most other models – the lowest layer ('Evidence Objects') relates to addressing digital resources that provide our evidence, while the highest layer ('Musicological Objects') supports musicological annotation. A new middle layer ('Musical Objects') separates these two concerns and allows annotations to point to explicitly musical structures rather than just document structures with implied musical relevance.

### 5.1 The Evidence Objects layer: Referring to resources

Many musicological assertions will be supported by reference to materials used as evidence. To support this (R1.2), RESOURCES in our model should be URIs that point to media in standards-compliant forms such as MEI (for music notation), audio and video, and images, particularly using IIIF APIs.

One reason for restricting the form that evidence materials can take is that we must be able to refer to parts of those materials as well as the whole. These are the objects that we call REFERENCES in our model. REFERENCES are also URIs, but will use existing standards to select regions of the whole resource, usually using the URI fragment identifier. The mechanisms for these will be XML:IDS (MEI), media fragments and the IIIF Image API. The EMA schema for referring to music would also be acceptable, given a suitable interpreter.

The mechanisms in this layer are purely mechanical ways to refer to and partition documents, but there is no model provided for arbitrary combinations of partitions, or reuse of those partitions. To do this, we need to group and name these objects, a role that is filled by the Musical Object layer.

### 5.2 The Musical Objects layer: Identifying musical units

The Musical Objects layer (see Figure 2b) in our model serves two key purposes. Firstly, it supports arbitrary combination and division of objects from the Evidence layer in a way that commonly goes beyond what the URI schemes for document fragments support. Secondly, the layer provides clarity in what we have called Grouping (R3), allowing annotations to refer explicitly to musical substance rather than simply a single form that substance takes in one resource. This first role is carried out by the SELECTION object, which gathers references and resources to create a single URI to which other structures can point. All the URIs referenced by a single selection will be from a conceptually singular musical resource, but this may be distributed – for example, an edition of a symphony might have each instrumental part digitised separately and published at a different URI. Thus, the SELECTION is a FRBR MANIFESTATION. Where FRBR usually operates at the level of entire artistic works, here we use it (with the same meaning) at the level of a subset of the musical content. A subclass of SELECTION, AB-SENCESELECTION allows an insertion point or region to be combined with a directional ABSENCEQUALIFICATION (e.g. BEFORE) to indicate the location of missing material. This is particularly important for object-based encodings, such as MEI, where one can only directly address elements that are present in the resource.

The remaining objects in the Musical Object layer serve the second purpose, identifying the conceptual level at which commentary is operating, and grouping musical evidence into as appropriate. The first of these is the EXTRACT. We define an EXTRACT as a musical fragment or region for a single version of a work, but since we include multiple reproductions or performances, EXTRACT is a subclass of FRBR EXPRESSION. This allows, for example, the opening 8 bars of an image of a source, a transcription and a recording to be referred to as a single object.

The next layer up in our model groups EXTRACTS, allowing parallel passages in different versions of a work to be associated. This MUSICALMATERIAL corresponds to (and is a subclass of) the FRBR EXPRESSION. This is a broad and useful concept – on the one hand, it can be used to indicate aspects of orchestration, reduction or reworking between different versions, but on the other, it can be used quite neutrally to provide alignment information to support parallel presentation of different musical versions in an application.

The highest-level Musical Object, the MUSICALIDEA is an abstract entity that refers to the musical thought behind a musical structure, especially a recurring one[4]. Had the Lohengrin TimeMachine discussed earlier used our model, its *Motifs* would have been implemented using this object. The MUSICALIDEA is a subclass of a FRBR WORK applied at a local level within a piece of music.

These structures provide a FRBR-based set of concepts to allow more explicit and precise separation of a musical idea – or even a notation quirk – from a single exemplifying resource. It provides this in a medium-independent structure that supports the explicit connection of musical version, borrowing and cross-reference. Making these structures explicit and addressable also allows other structures to refer to them. This is not limited to the annotations provided by the MUSICOLOGICAL OBJECTS layer, but

---

[4]We derive our term with reference to Hanslick's *Musikalische Ideen* (see [4], ch 3).
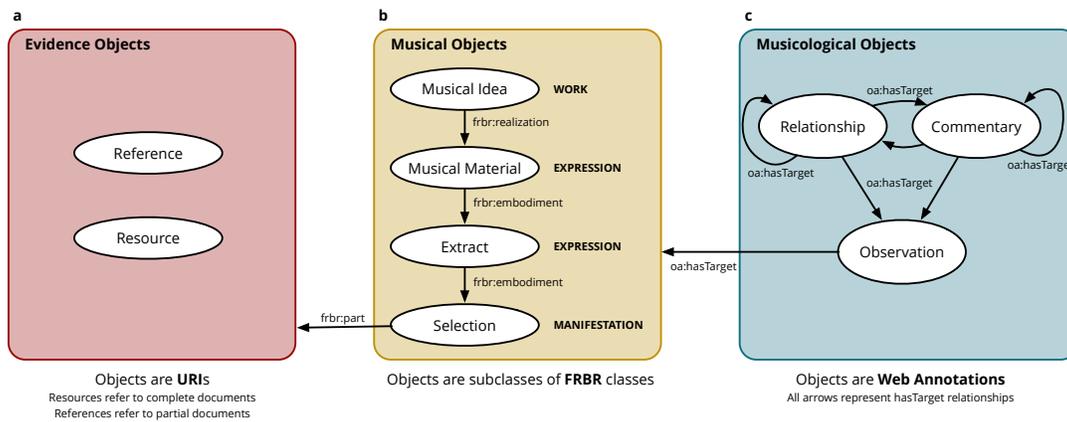
Figure 2: An overview of the 3-layer model: a. Evidence Objects refer to resources used by on which our scholarship is based, in the form of web-accessible resources; b. Musical Objects, which represent abstractions of the musical content; and c. Musicological Objects, annotations that describe and comment on Musical Objects.
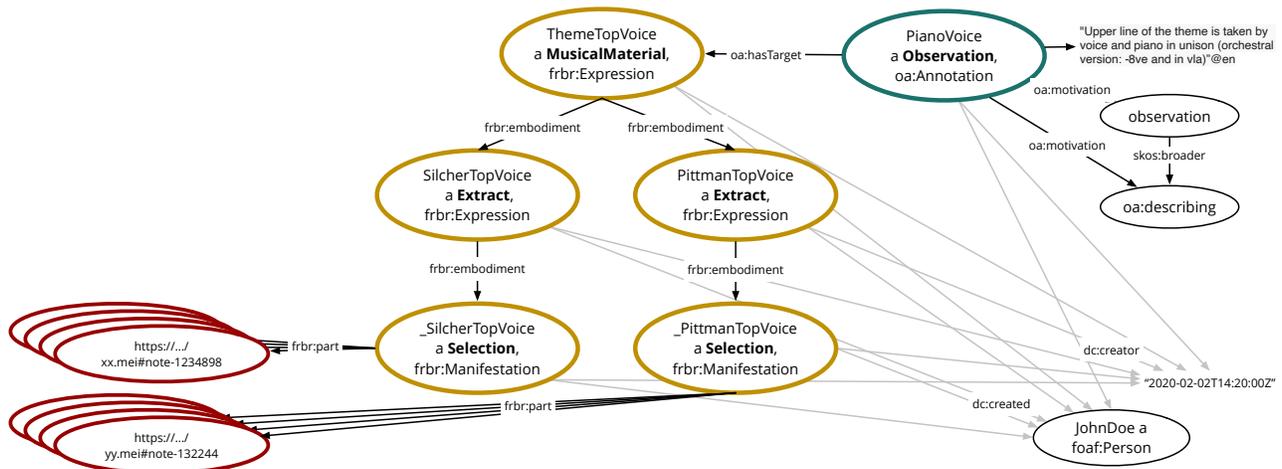


Figure 3: Instance data encoded using the model (selected and simplified from the prototype described in Section 6).

also allows the use of the same Dublin Core terms that Web Annotations use to record information about their creator and creation date.

## 5.3 The Musicological Objects layer: Annotating Musical Objects

The Musicological Objects layer provides structures for documenting and publishing musicological research. These objects are used to describe, compare and record historical context associated with musical objects, note music theoretical observations about them, present hypotheses, link them to non-musical sources, or make scholarly commentary. All the objects can be implemented using Web Annotations, providing structures for the annotations and their subjects, as well as their motivations, authorship and creation date.

We distinguish, for the purposes of our model, between Observations, Relationships and Commentaries. Observations have as their target one or more objects in the Musical Layer, and select or draw attention to aspects of the music that will be relevant for further study. These objects are meant to be descriptive and (relatively speaking) factual rather than explicitly interpretative.

Relationships and Commentaries both involve a greater degree of interpretation and musicological intervention – the former by drawing attention to a change or a difference between multiple Observations or other Musicological Objects, and the latter by applying a comment or interpretation to them (see Figure 2c).

To illustrate some of the model's structures in practice, a simple example, drawn from two arrangements of the Allegretto from Beethoven's Seventh Symphony, is given in Figure 3. Here, two arrangements of the movement for voice and piano are taken, and the opening melody, which is doubled between voice and piano, is annotated with an Observation. The Observation targets the MusicalMaterial that represents these parallel passages, each of which is represented by a Extract. If this example included

facsimiles of each of these sources, then we would have two Selections for each Extract, but here, we only provide connections to the MEI-encoded form of each. Every object can have information provenance recorded, such as the editor responsible and a timestamp.

## 6 APPLYING THE MODEL – A PROTOTYPE ANNOTATION TOOL

We now describe a prototype implemented to support the musicological investigation described in section 2.1 (Figures 4 & 5). Although we have described requirements for what a musicological annotation system should be able to express, we have not addressed requirements for systems that create or process the resulting objects. This is the subject of ongoing research, from which one clear requirement is that the musicologist should not be expected to encode objects directly, for example by hand-authoring RDF as was the case for The Lohengrin TimeMachine. More strongly, the scholar should engage with the structures only to the extent that it is helpful at the time, and in a form that is intuitive for them. Complex chains of objects should be created automatically where possible, and displayed using appropriate visual or textual language.

Aiming to support our own future musicological research, we have implemented the model in a prototype browser application[5], exploring its fit to the functional requirements and, less formally, its practical robustness and usability. We start from an evidence base of IIIF and MEI editions of music, connected and contextualised using Linked Data. Traversal of these structures uses the MELD 2.0 framework[15], alongside additional interactions and visualisations which do not.

The musicologist is required to have an account with a Solid Pod provider (discussed in [14]) in order to make new annotations, although a read-only mode works without a log-in. The scholar can navigate through a catalogue of pieces to identify a subset of work to compare. These can then be presented and annotated. The model is not directly presented to the user – labelling a group of regions of music as 'Parallel Passages', for example, dynamically creates and stores the MusicalMaterial, Selections and Extracts without these being presented as such to the user (they are simply highlighted as the passages under consideration). A debug view allows us to view all the objects being created by the system.

## 7 ASSESSMENT AGAINST REQUIREMENTS

No single application would be expected to use all elements from the requirements list, but this example does demonstrate a broad selection of them in practice. We refer to scores (R1a) and images (R1b) via URIs (R1.2), but do not implement audio, video or non-musical materials (R1c-d, R1.1) – all of which can be accommodated within the model. Our selection tool (Figure 4) permits arbitrary selection of score elements (R2) and these selections are published separately from the evidence resource, as Linked Data stored and published through the Solid infrastructure (R2.1). Musicological annotations, musical objects and the resources themselves stand separately from each other, and can be distributed and published in different places and with different levels of access. The application is designed to accommodate multiple resources (R2.2), and supports

a single selection across multiple resources (such as orchestral parts as different image sets). Absences (R2.3), which are accommodated by the model, are not, as yet, supported by the tool.

The application transparently handles single-source and multiple version annotations (R3, R3.1, R3.3). The ability to indicate multimedia renditions of the same version (R3.2) and recurring motifs (R3.4) are available in the model, but not yet in this implementation, although this will be added to future iterations. For all structures, responsibility for their creation is recorded (R5), with user information made available through Solid.

In a future iteration, annotations with textual content will be made through the application and attached to any unit that can be selected for the purpose (R4, R4.1), with a distinction between observation and commentary (R4.3), however these facilities are not present in the current prototype.

## 8 DISCUSSION AND SUMMARY

The model and implementation combine to produce a decentralised system, where musical structures and observations about them can be made separately from the resources being used as evidence, but can themselves be shared and reused by their authors or by others. This especially supports research on multiple resources that are stored and published in different places, benefitting especially from frameworks such as IIIF. Although the model is implemented and expressed in RDF, as an abstract model, it may prove useful in other implementations.

The concepts that we have defined here are deliberately neutral in music-theoretical terms, and we anticipate that some may choose to define sub-classes for specific structures such as 'theme', 'recapitulation' or 'riff'. These would depend substantially on the intended application, but are not expected to alter the relationships between our parent classes. For our own current research, we benefit from the greater flexibility of more general-purpose groupings; early stage investigations in particular would struggle to prescribe a vocabulary of units that will be labelled.

Our modelling thus far has focussed most heavily on the Musical Object layer, as a prerequisite to making meaningful annotations. We continue to investigate the Musicological Object layer in parallel with musicological research, evaluating the need for further specialisation of web annotations.

A distributed annotation system poses interesting opportunities and challenges, both technical and social. The ability to control who can access user-created structures, and when, is particularly valuable to researchers, but more research is required into the user interfaces and technical mechanisms for controlling their sharing and discovery. Meanwhile, there are sustainability and reproducibility challenges for research built on distributed structures such as these, requiring, for example, explicit strategies for static snapshots of the graph that forms the basis for completed research.

We have demonstrated the viability of our model through a working implementation, creating a user interface that aims to present intuitive concepts and interaction modes whilst reading and writing the generic objects of our model. What we would expect is that information sharing between applications can be facilitated by the use of a powerful shared model. Future work will test the model in wider settings and, crucially, with a range of

---

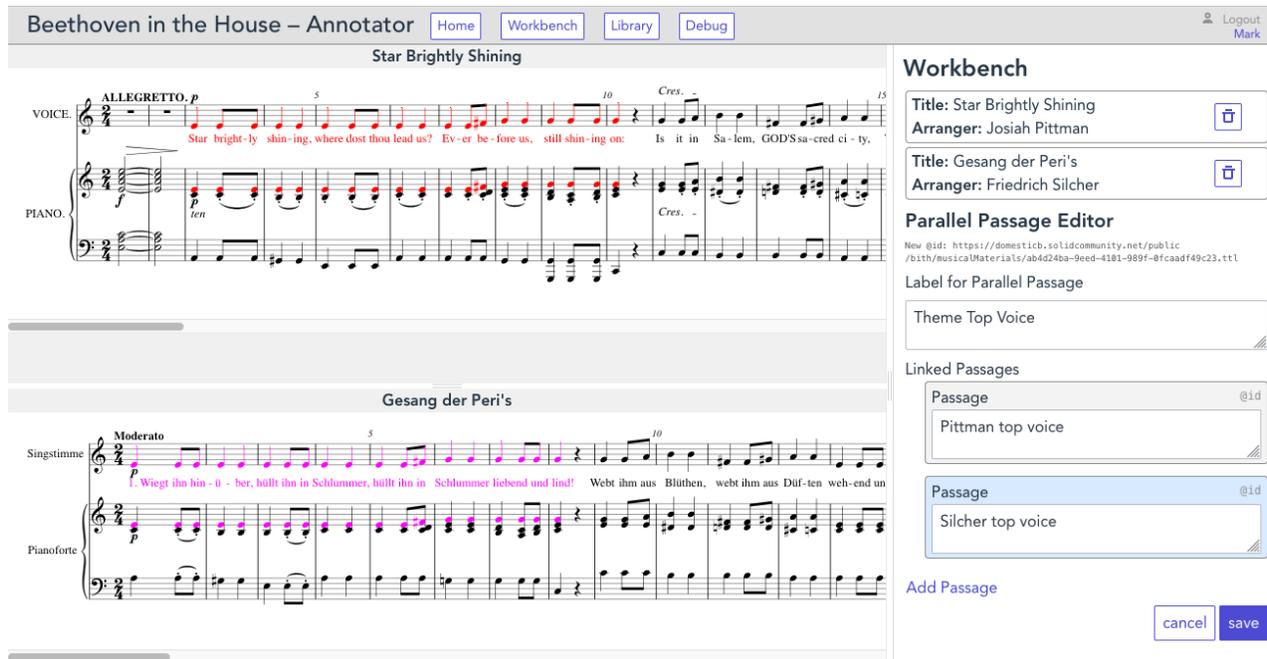[5]Available on GitHub at https://github.com/DomesticBeethoven/bith-annotator.

**Figure 4: Screenshots of the prototype application, showing a user indicating parallel passages between two different arrangements – an action that creates a series of objects (a MUSICALMATERIAL, two EXTRACTS and two SELECTIONS) which are saved to and accessible from their Solid pod.**
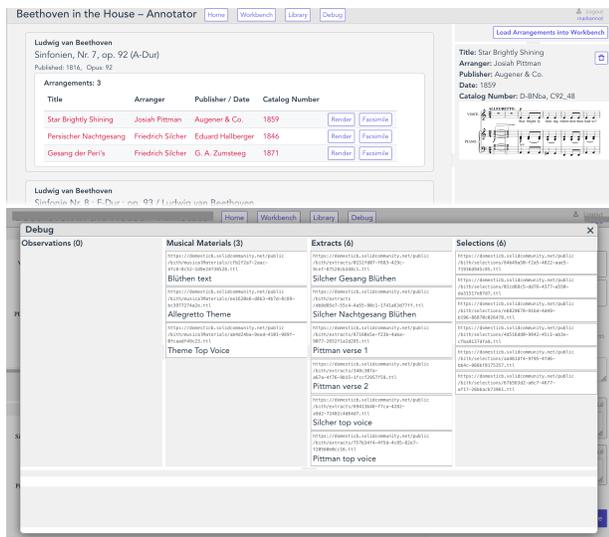


**Figure 5: Screenshots of the prototype application, showing selection of relevant arrangements to compare (above) and a debug view that shows the RDF objects that have been created in the background (below).**

musics, to better understand any flaws or limitations; however we are confident of the applicability and suitability of our model well beyond the research questions explicitly considered so far.

Central to the value of this approach is in allowing a musicologist to make scholarly observations about music as distinct from musical documents. Different documents, whether recordings, encoded scores or images, can be treated in a uniform way at the musical and musicological level because we recognise that our discourse is often directed at the things the documents give evidence for. The separation inherent in our model gives a consistency of structure so that shared information between quite different endeavours can still retain expressive richness despite operating over different musics and media, and in doing so creates a strong foundation for the inclusion of digital library resources in digital musicology research.

## ACKNOWLEDGMENTS

## REFERENCES

[1] Tim Crawford and Richard Lewis. 2016. Review: Music Encoding Initiative. *Journal of the American Musicological Society* 69, 1 (April 2016), 273–285. https://doi.org/10.1525/jams.2016.69.1.273

[2] Richard Freedman, Philippe Vendrix, David Fiala, Micah Walter, Emilio Sanfilippo, Raffaele Vigianti, and Daniel Russo-Batterham. 2020. Citations: The Renaissance Imitation Mass (CRIM). *LE STUDIUM Multidisciplinary Journal* 4 (2020), 22–26.

[3] Mathieu Giraud, Richard Groult, and Emmanuel Leguy. 2018. Dezrann, a Web Framework to Share Music Analysis. In *Proceedings of the International Conference*

*on Technologies for Music Notation and Representation (TENOR 2018)*. 104–110.

[4] Eduard Hanslick. 1854 [third ed. 1874]. *Vom Musikalisch-Schönen : Ein Beitrag zur Revision der Ästhetik der Tonkunst.* J. A. Barth, Leipzig. https://archive.org/details/vommusikalischsc00hans

[5] David Lewis, Kevin R. Page, and Laurence Dreyfus. 2021. Narratives and exploration in a musicology app: Supporting scholarly argument with the Lohengrin TimeMachine. In *8th International Conference on Digital Libraries for Musicology, Virtual Conference, July 28-30, 2021*, Claire Arthur (Ed.). ACM, 50–58. https://doi.org/10.1145/3469013.3469020

[6] David Lewis, David Weigl, and Kevin Page. 2019. Musicological Observations During Rehearsal and Performance: A Linked Data Digital Library for Annotations. In *6th International Conference on Digital Libraries for Musicology* (The Hague, Netherlands) *(DLfM '19)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3358664.3358669

[7] Stephen Mould. 2014. Fidelity to 'Fidelio': The convention of unwritten appoggiaturas in Beethoven's only opera. *Music and Letters* 95 (2014), 17–38.

[8] Nancy November. 2021. *Beethoven's Symphonies Arranged for the Chamber: Sociability, reception and canon formation.* Cambridge University Press, Cambridge, UK.

[9] Graham Pont. 2007. French Overtures at the Keyboard: The Handel Tradition. *Early Music* 35 (2007), 271–288.

[10] Robert Sanderson, Paolo Ciccarese, and Benjamin Young. 2017. *Web Annotation Data Model.* W3C Recommendation. W3C. https://www.w3.org/TR/2017/REC-annotation-model-20170223/.

[11] Barbara Tillett. 2005. What is FRBR? A conceptual model for the bibliographic universe. *The Australian Library Journal* 54 (2005), 24–30.

[12] Raffaele Viglianti. 2016. 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* (New York, USA) *(DLfM 2016)*. Association for Computing Machinery, New York, NY, USA, 57–60. https://doi.org/10.1145/2970044.2970056

[13] Alan Walker. 1981. Liszt and the Schubert Song Transcription. *The Musical Quarterly* 67 (1981), 50–63.

[14] David M. Weigl, Werner Goebl, Alex Hofmann, Tim Crawford, Federico Zubani, Cynthia C. S. Liem, and Alastair Porter. 2020. Read/Write Digital Libraries for Musicology. In *7th International Conference on Digital Libraries for Musicology, Montréal, Canada, October 16, 2020*. ACM, 48–52. https://doi.org/10.1145/3424911.3425519

[15] David M. Weigl and Kevin R. Page. 2017. 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*, Sally Jo Cunningham, Zhiyao Duan, Xiao Hu, and Douglas Turnbull (Eds.). 221–228. https://ismir2017.smcnus.org/wp-content/uploads/2017/10/190_Paper.pdf

[16] Christof Weiß, Frank Zalkow, Vlora Arifi-Müller, Meinard Müller, Hendrik Vincent Koops, Anja Volk, and Harald G. Grohganz. 2021. Schubert Winterreise Dataset: A Multimodal Scenario for Music Analysis. *J. Comput. Cult. Herit.* 14, 2, Article 25 (May 2021), 18 pages. https://doi.org/10.1145/3429743