TROMPA: Towards Richer Online Music Public-domain Archives

Deliverable 5.2

Score Edition Component

Grant Agreement nr	770376
Project runtime	May 2018 - April 2021
Document Reference	TR-D5.2-Score Edition Component v1
Work Package	WP5 - TROMPA Contributor Environment
Deliverable Type	Report
Dissemination Level	PU
Document due date	30 April 2019
Date of submission	Asked for extension
Leader	GOLD
Contact Person	Tim Crawford (GOLD)
Authors	Emilia Gomez (UPF), Cynthia Liem (TUD), Sonia Espi (UPF), Aggelos Gkiokas (UPF), Tim Crawford (GOLD), David Weigl (MDW)

Reviewers	Bauke Freiburg (VD), Emilia Gomez (UPF)

Executive Summary

Scores are used throughout TROMPA, and are an indispensable means of communicating and preserving the musical content of classical music works. Music notation evolved since about the tenth century and is universally understood by professional musicians, i.e. those most likely to generate the performances discussed and appreciated by their audiences. At the same time, it is not necessary to understand music notation to appreciate and enjoy classical music, so TROMPA does not expect its community of music-enthusiasts to read it, although many members of this community will be able to do so.

Digital scores, either rendered from MEI encodings via the TROMPA Data Infrastructure, or from PDF or other graphic files representing physical library documents (such as the historic printed scores available via IMSLP, or the early music prints in EMO), will find extensive use throughout WPs 3, 4, 5 & 6.

The Digital Score Edition component of TROMPA is to a large extent based on the Music Encoding and Linked Data (MELD) [1] technology developed in prior projects FAST[1] and Transforming Musicology[2]; this will be adapted and augmented for the specific purposes of TROMPA. It combines use of the Music Encoding Initiative (MEI) format for representing musical documents with a semantic approach tailored to the requirements of musicians and music scholars, using Linked Data.

This document describes the essentials of the Digital Score Edition component and summarises how they will be put to use in various use-cases. It largely focuses on new elements being introduced for TROMPA, in particular, methods for user-contributed and automated annotations which preserve provenance and maintain user control; to a large extent this is ground-breaking work, so some research questions (see 1.2 Annotations and links) remain open.

The Digital Score Edition component of TROMPA closely interacts with the Contributor Environment (WP5 - TROMPA Contributor Environment) via TROMPA's Data Infrastructure (TR-D5.1-Data Infrastructure v1). We expect it to find use to a greater or lesser extent in the following Tasks: 3.2 (Music description); 3.4 (Visual analysis of scores); 3.6 (Multimodal cross-linking); 5.3 (Multimodal integration of music data); 5.4 (Music performance assessment mechanisms); 6.2 (Pilot for music scholars); 6.3 (Pilot for orchestras); 6.4 (Pilot for instrument players); 6.5 (Pilot for singers)

---

[1] http://www.semanticaudio.ac.uk

[2] https://tm.web.ox.ac.uk/

Version Log		
#	Date	Description
v0.1	27 April 2019	Initial version submitted for internal review
v0.2	30 April 2019	Revised version after internal review
v1.0		Final version submitted to EU

1. Introduction

Task 5.2 focuses on the integration of data and technologies (WP3–WP4) for the collaborative edition of digital scores. In this functionality, users will be able to access digital scores that can be annotated and linked; this further leads to detailed linking and annotation across and between score documents and audio recordings. This alignment is enabled by the adoption within TROMPA of the Music Encoding Initiative (MEI) as the principal storage and exchange medium for musical data, and the MELD (Music Encoding and Linked Data) framework for data management. The corresponding deliverable D5.2 is a Digital Score Edition component for flexible on-screen display (and audio playback where possible), within a standard web browser, of musical scores, online user-interaction (e.g. passage-selection and highlighting for further processing or search-query building) and annotation (including linkage to other documents or score-locations). The first release will allow annotation (at the measure level) of scores and metadata; the second release will permit detailed linking and annotation across and between score documents and audio recordings.

There have been problems of recruitment at GOLD, due to teaching pressure on staff time, and the fact that the programmer, Ryaan Ahmed, has recently been headhunted by MIT so has not been able to complete the development of the Digital Score Edition for M12. However, the design and specification of the Digital Score edition described in this document is in a more advanced state than was anticipated in the project proposal, so we expect that our new programmer, Oscar Dub, will be able to make rapid progress and produce a more fully-featured first version within a six-month extension to M18 (October 2019). The plan is for a basic functioning version, working with hard-coded data, by the end of May 2019, to which will be added a user-annotation interface and linkage to the Contributor Environment (including metadata display) by the end of July 2019. Further testing and linkage to other TROMPA components will be carried out during the next two months to ensure delivery of robust software by M18.

The present document is a report on the design and requirements of the software of the first release, which will be developed within the proposed extension to M18, and made available via the TROMPA web-site.

For the first release D5.2-1, described in this document, the interface displays a page of score (previously specified via the Contributor Environment, usually by a metadata query) on which the user can:

I. select one (or maybe more) regions, whose encoded musical contents (or data such as locations of musical objects, depending on task) can be specified and passed to other TROMPA components via the data infrastructure's API (see Granularity of selection/interaction, below);

II. see helpful labels indicating the presence of existing annotations, which may contain links to other resources, within TROMPA or external, add new annotations, and edit annotations which they 'own' (see Annotations and links, below);

III. follow links (from within the annotations) to other TROMPA musical entities or score-pages or to external resources, while maintaining a trail which will enable users to return progressively to the starting-point (see Annotations and links, below).

1.1 Granularity of selection/interaction

For D5.2-1 this will be at the measure level, corresponding with the initial granularity specified for other Tasks at this stage of development (e.g. T3.4, 'Visual analysis of scanned scores', and T3.5, 'Multimodal music information alignment', which both use bar line recognition as their basic level for useability, though finer granularity is likely to be achieved in many cases, depending on the format and cleanness of the input data). The TROMPA data model can accommodate higher granularity, and in terms of the MEI encodings, granularity of targeting can technically be at any level of the hierarchy.

1.2 Annotations and links

The MELD framework permits a fully adequate range of annotation types which have been deployed in earlier projects (see 3. Music Encoding and Linked Data, below); it can easily be extended with new styles, new annotation types, etc., though this is not likely to be necessary for the first release described here. However, as MELD does not provide any authoring environment for user-generated annotations, a user-friendly interface to achieve this for music scholars and other uses will be developed within the TROMPA Digital Score Edition component. Associating annotations with credentialled users (so that they maintain control of access, sharing and display of their own annotations) is new ground requiring significant original research, and, since it further requires significant interactions with the TROMPA Contributor Environment, is not planned for the first release of the Digital Score Edition component. MELD operates upon Linked Data employing standard web URIs (where linking to web resources; e.g., an MEI file or an audio recording) or media fragment identifiers (where linking to resource fragments; e.g., a particular measure of an MEI file, or a particular time range within an audio recording). This provides for transparent interlinking between resources hosted within the Contributor Environment (via its REST API) and resources hosted anywhere on the web.

For score display in a browser, we will support other formats, e.g. PDFs / bitmaps, only when we have reliable information about image regions and their relationship with the underlying musical data model; normally this would have been determined in an OMR process, and bounding box data maintained in the TROMPA system by means of the unique xml:id identifiers within the initial MEI encodings.

1.3 Audio linkage

Linkage to audio recordings is not part of the specification for D5.2-1 (M12), but it is eventually required for D5.2-2 (M34). Therefore a few proof-of-concept examples of such audio/score linkage will be incorporated to demonstrate the power of the concept. Ultimately we intend to achieve automatic score-scrolling synchronised with playback, but a simple 'play-from-here' button makes the interface immediately more attractive to the user by means of alignments generated by the TROMPA multimodal music information alignment component via the alignment data model proposed in D3.5.

1.4 MELD and TROMPA score-interface requirements

The MELD-clients-core library provides a starting point for the development of the various interfaces required. Existing MELD components cater for these requirements, but the alignments need to be generated. Automated (rather than manual/hard-wired) generation of alignments forms part of Task 3.5, 'Multimodal music information alignment', and will be added as it becomes available.

1.5 Uses of the Digital Score Edition component within TROMPA

Scores are used extensibly within the domain of classical music. Historically, they are the medium of preservation and communication of music between composers and performers, providing the information needed to realise the composer's concept in sound. Music notation evolved since about the tenth century and its modern form is universally understood by professional musicians, i.e. those most likely to generate the performances discussed and appreciated by audiences. The interpretation of the often incomplete or ambiguous data in historical scores has in former times been considered work that must be done by music scholars, but since the 'early music revival' of the 1970s, and with the general availability of online facsimiles (many of which can be found in IMSLP), performers, both professional and amateur musicians, increasingly take on this role directly. Part of the TROMPA mission is to aid such musicians by providing extra contextual information about the music and how other people perform it. Throughout TROMPA, scores will be a standard means of visualising music, but, at the same time, it is not necessary to understand music notation to appreciate and enjoy classical music, so TROMPA does not expect its community of music-enthusiasts to read it, although many members of this community will be able to do so.

Several TROMPA Tasks are certain to make frequent use of the Digital Score Edition component, including the following:

WP3: Task 3.2: Music description; Task 3.4: Visual analysis of scores; Task 3.5: Alignment of musical resources; Task 3.6: Multimodal cross-linking

WP5: Task 5.3: Multimodal integration of music data; Task 5.4: Music performance assessment mechanisms; Task 5.5: Annotation tools

WP6: Task 6.2: Pilot for music scholars; Task 6.3: Pilot for orchestras; Task 6.4: Pilot for instrument players; Task 6.5: Pilot for singers

2. Layout/features

In the diagrams which illustrate this section the score-edition interface appears as a 'frame' (the large rectangle in the diagrams) inside whichever TROMPA interface provides the context for its use (this context may have its own menus and even its own toolbox, e.g., for selecting a score from the corpus; it also displays basic composer/title information from metadata provided by the Contributor Environment).

The score-edition frame gives data about the current display (bar-number range), and provides basic in-score navigation buttons as well as its own palette of context-dependent 'Action' buttons. These might include various kinds of view, search and audio/MIDI playback (where a suitably linked audio or MEI file is available).

2.1 PDF and Digital Score Views

2.1.1 PDF View

'PDF View' simply renders the required portion of the musical score from an existing PDF, which may derive from a scan of a historical music edition or from a modern, computer-engraved score. Score-location links (such as page/system/bar numbers) need to be aligned externally with page-image offsets (pixel-based); these will initially be hard-coded, but provided ultimately by the use of optical barline-detection methods from the Visual Analysis of Scanned Scores (D3.4) and Multimodal cross-linking (3.6) TROMPA deliverables. It will be necessary to allow users to zoom in to the scores to examine details (especially in the case of complex music), or to zoom out to see the context of a detailed view. Similarly, user-adjustments to the default presentation (horizontal or vertical scrolling) may be necessary depending on the nature of the PDF scores and the level of annotation detail. (We do not expect to allow multiple page-views, except possibly in the case of adjacent pages, e.g. where the PDF scan is of a two-page spread.)

2.1.2 Digital Score View

'Digital Score View' renders a score from its MEI encoding (where available) using the javascript library Verovio[3]. This library offers a number of advantages to TROMPA, being based on the data-model of MEI and fully compatible with the Linked Data requirements of MELD, and has been used for this purpose in a number of related projects. It has the further advantage that it is capable of rendering both standard modern common music notation and the earlier form of mensural notation used in 16c music prints, both of which will be used within TROMPA.

The user needs to be able to switch easily from PDF View to rendered Digital Score View; as this user-selection is made, the possible further actions may change too, as shown in the following mock-up illustrations. (See 2.1.3 PDF View with nothing selected.)

2.1.3 PDF View with nothing selected

---

[3] http://verovio.org

Figure 2.1 - In PDF View (nothing selected), a user can Play/Pause the score (if possible)

In PDF view the user can select (within the display) only to the measure level; in digital-score view selections may be of any group of displayed items (this will be constrained according to the current capabilities of the system). Actions are of two general types: internal (affecting the display itself) and external (in which data will be sent from the interface to an external TROMPA process – usually some kind of search – and results returned to the user in a separate window/tab; such a result may be another score-display). (See 2.1.4 PDF View with a region of score selected)

2.1.4 PDF View with a region of score selected

Figure 2.2 - In PDF View (bars 11-13 selected), users can play from the beginning of the selection, or just the selection, and perform other processes on the selection (here, a 'Find Similar' search)

The main internal actions in both views will involve navigation within the score ('Next/previous measure', 'next/previous system', 'Go to measure/system x'), and display-zoom, controlled mainly by buttons.

External actions are constrained by the nature of the selection: in PDF view (selections are measure-based), a search can only be for similar measures within the corpus (and depends further on the availability of a suitable MEI encoding and/or index entry extracted therefrom).

However, in Digital Score View, a selected note-sequence could be used as a more generic search for a musical motif or phrase, which in turn might be further constrained: e.g. pitch only, rhythm only, pitch and rhythm combined, pitch contour (up/down/same), diatonic pitch (note-names), etc. (See 2.1.5 Digital Score View with specific notes selected)

2.1.5 Digital Score View with specific notes selected

Figure 2.3 - In Digital score View, users can select individual notes.

3. Music Encoding and Linked Data (MELD)

MELD[4] is a semantic framework and a set of client rendering and interaction libraries [1] designed to simplify the creation of web applications providing multimodal access and interaction with music score encodings. The Digital Score Edition component described in this document offers a good example of the range of possibilities enabled by MELD, which handles the integration of TROMPA score data which can thus be displayed (as in the case of PDF score documents), rendered (as in the case of MEI-encoded scores which are displayed within the browser using the Verovio Javascript library), played back (as in the case of pre-aligned audio files, or even video), and passages within the scores can be selected to act as queries to a search-system or for other external processing, with

---

[4] http://github.com/oerc-music/meld

links and provenance maintained throughout. MELD is thus enabling technology for the Digital Score Edition providing the data infrastructure on which the score-display interface can easily be developed in Javascript.

The semantic framework combines and builds on W3C web standards including the RDF[5] model for data interchange, the Web Annotation Data Model[6], the Linked Data Platform (LDP)[7], and the Media Fragments URI specification[8], established Semantic Web ontologies including the Timeline and Segment Ontologies [4-5], and the Music Encoding Initiative's[9] comprehensive XML schema (MEI) which expresses the classes, attributes, and data types required to encode a broad range of musical documents and structures [6]. It provides for the abstract structural segmentation of musical works, and for the linking of structural segments to concrete multimedia representations, e.g., score fragments as collections of identified MEI elements; specified regions of audiovisual recordings; image regions; identified regions of marked-up text in HTML or TEI.

Through its basis in RDF Linked Data, each entity within this framework is identified by its own URI identifier. This supports the creation of Web Annotations targeting any combination of entities, including (fragments of) the abstract and concrete representations discussed above, as well as entities described within the TROMPA Contributor Environment (via its REST API), and any other URI-identified resources within the wider Semantic Web. The framework thus supports the authoring of annotations linking a passage in a scholarly article to a set of phrases within an MEI-encoded score; or, comparing fragments of two audio recordings, each capturing renditions of the same segment of a musical work; or, commenting on regions of scanned (printed) score images in relation to a digital encoding of the same score; or even associating a particular musical phrase occurring within multiple distinct musical works with the geospatial provenance relating to their composition.

The TROMPA alignment data model, which instantiates the MELD semantic framework, further enables the inclusion of feature data derived from musical representations by quantifying specific aspects of e.g. an audio signal, or of a symbolic score. This data model is discussed in D3.5, and illustrated in the figure below (See 3.1. TROMPA alignment data model based on the MELD semantic framework)

The MELD client libraries[10] facilitate the creation of web interfaces that render and facilitate interactions between instances of the different musical representations interlinked within the MELD semantic framework. They are implemented in a set of Javascript (React) components, actions, and reducers, which can be configured and flexibly extended to build applications for specific use contexts.

To date, MELD applications have been created around uses in scholarly publication [6], music performance [7], and composition [8], demonstrating the applicability of this technology to TROMPA's varied use cases. However, MELD has thus far lacked an authoring environment for the definition of user-defined structural segmentations (beyond automatic note-, measure-, and section-

---

[5] https://www.w3.org/RDF/

[6] https://www.w3.org/TR/annotation-model/

[7] https://www.w3.org/TR/ldp/

[8] https://www.w3.org/TR/media-frags/

[9] https://music-encoding.org

[10] http://github.com/oerc-music/meld-clients-core

level structurings retrievable from MEI encodings) and for annotation vocabularies (beyond those defined a priori by the application developer through manual creation of RDF triples). This limitation will be addressed by the TROMPA Digital Score Edition component, which will in turn be used to enhance the usability of MELD components within TROMPA user pilots (supporting, e.g., scholars in publishing scholarly communication, or instrumental performers in their rehearsal).

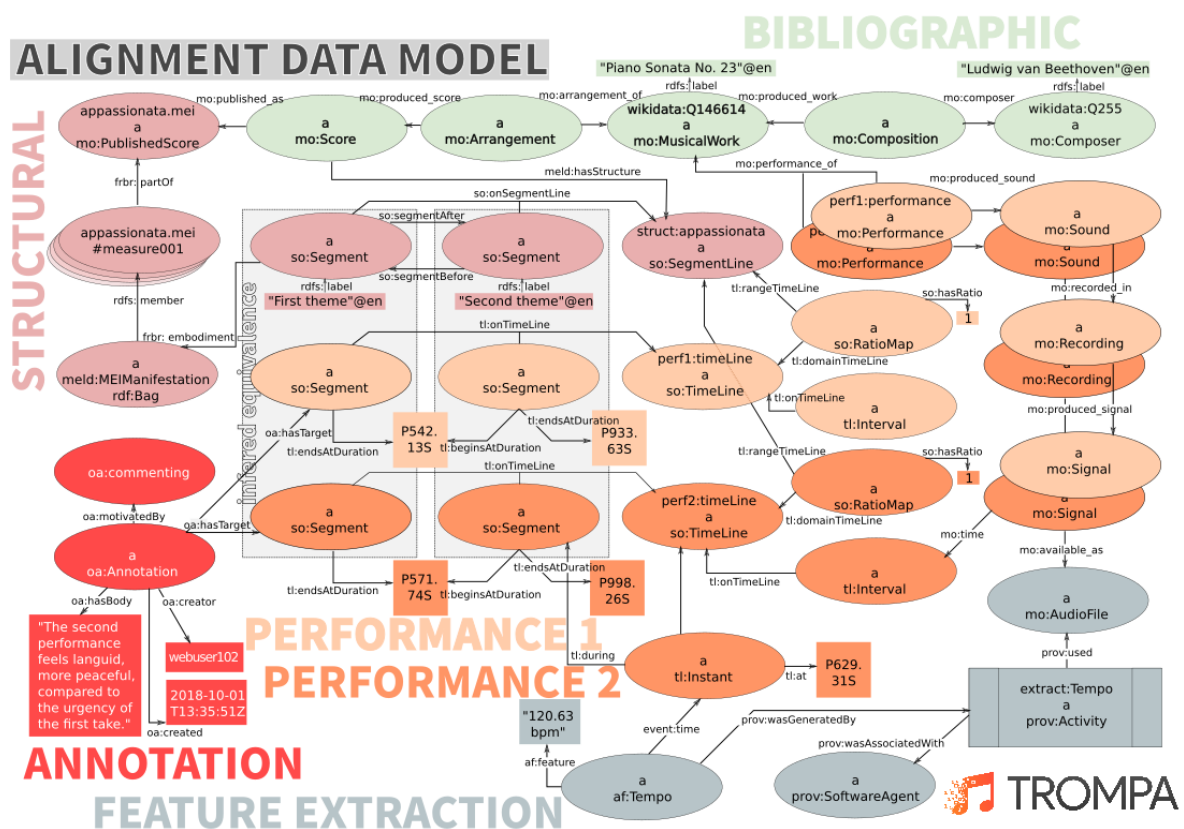## 3.1 TROMPA alignment data model based on the MELD semantic framework

Figure 3.1 - In Digital score View, users can select individual notes.

4. Digital Score Annotations

As mentioned above, MELD does not provide a user-facing interface for directly generating or editing annotations. However, such an interface is being developed for audio annotation in D5.5 (Annotation Tools) and this will be adapted for the purpose. The annotations to scores, both textual/verbal and graphical/non-verbal (e.g. rectangular regions of coloured shading), will be recorded and maintained using the mechanism described in D5.5, which was specifically designed for TROMPA's use cases for Performers, Choir Singers and Music enthusiasts. This makes use of the

W3C Web Annotation data model, as deployed appropriately within the TROMPA Contributor Environment and integrated within MELD.

The audio-annotation tooling within D5.5 is to some extent based on existing open-source projects such as the Audio Annotator javascript interface[11], so the score-annotation tooling for the Digital Score Edition interface will, at least initially, adopt the same strategy with suitable adaptation. The requirements for storage, maintenance, authentification, ownership and access-management of annotations for the Digital Score Edition are very similar to those for audio-annotation in D5.5. As with audio annotation, it may be necessary to indicate, as a form of annotation, locations of musically-similar passages[12] within a score, which can easily be achieved with colour-coded shading, for example.

At some future point within TROMPA we intend that a more unified approach to annotations in general capable of incorporating other modalities (e.g. video, image-regions, marked-up textual elements) will be developed; this will form a significant innovative research output for the project.

5. Digital Score Edition and the TROMPA Contributor Environment

The Digital Score Edition component interacts with TROMPA's Contributor Environment for the selection of scores for display, for saving annotations and for managing data-transmission for further processing (e.g. initiating searches external to the currently-displayed score). The Contributor Environment's API[13] is described in technical detail in the CE API manual (see D5.3). Specific CE components for the Digital Score Edition[14] and for Multimodal integration of music data[15] have also been provided.

6. Music repertories

For the typical repertories which need to be displayed frequently in the Digital Score Edition component, we give two examples from different periods which are contrasted in terms of their complexity and their notation.

1. Early music: 16th-century vocal music

    This repertory has the advantage of relative simplicity in terms of score-display requirements. Also, GOLD has long-term experience in this field and continues to engage in active research. The Early Music Online (EMO) resource of ~300 early printed music books has been used extensively for this work, and is freely available, not only in high-quality

---

[11] https://github.com/CrowdCurio/audio-annotator
[12] See: https://github.com/MTG/similarity-annotator
[13] https://github.com/trompamusic/ce-api
[14] https://github.com/trompamusic/ce-digital-score-edition-component
[15] https://github.com/trompamusic/ce-multimodal-component

digital scans from archival microfilms (c32,000 page images), but also in MEI encodings from the output of the OMR program Aruspix, and indexes derived therefrom. Many of the works in EMO are freely available online in score editions as PDFs, sometimes with digital encodings and even with individual voice parts in MIDI format.[16] (The latter scores can be displayed as PDFs or, if MEI encodings can be extracted, as Digital Scores using Verovio.)

Recent work on full-text musical-content searching of EMO allows fast retrieval of music to match a given example (currently, either as query code or as a page-image). Such searches will be activated by user-selection of a note-sequence within a rendered score to find similar music in the collection: the query is generated by extracting the sequence of MEI pitches from the selected notes which is then appropriately encoded and passed as query by URL to the EMO server's API.

2.   Beethoven piano scores combined with structural analyses

Later music than that described in 6.1, above, is much more complex in texture and in notation. An example which is central to the TROMPA Music performance assessment component (D5.4-1) is 19c piano music; this TROMPA component will eventually provide a comprehensive collection of digitised (MEI) Beethoven piano scores. These will be associated with simple structural analyses described by visual annotations (using colour-shading and text-labelling), to inform performers about salient segments of their own performance and those of others. When these annotations can be directly generated by scholars, they will provide a basis for scholarly discourse as well as helping the performance-assessment use case. This is an example of a non-verbal (i.e. non-textual) form of annotation, made possible through adoption of MEI and the data-sharing capacities of the TROMPA environment enabled by MELD.

7. References

[1] Weigl, D., & Page, K. (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.

[2] Lewis, D., Weigl, D. M., Bullivant, J., & Page, K. R. (2018). Publishing musicology using multimedia digital libraries: creating interactive articles through a framework for linked data and MEI. In *Proceedings of the 5th International Conference on Digital Libraries for Musicology* (DLfM '18). ACM, New York, NY, USA, 21-25. DOI: https://doi.org/10.1145/3273024.3273038

[3] Pugin, L., Zitellini, R., & Roland, P. (2014). Verovio: A library for Engraving MEI Music Notation into SVG. In *Proceedings of the 15th International Society for Music Information Retrieval Conference* (ISMIR 2014), pp. 107-112.

---

[16] Over 1,000 such works are available in PDF at https://www.uma.es/victoria/ together with digital encodings.

[4] Raimond, Y., Abdallah, S. A., Sandler, M. B., & Giasson, F. (2007). The music ontology. In *Proceedings of the 8th International Conference on Music Information Retrieval* (ISMIR 2007)*, pp. 417–422.

[5] Fields, B., Page, K., De Roure, D., & Crawford, T. (2011). The segment ontology: Bridging music-generic and domain-specific. In *2011 IEEE International Conference on Multimedia and Expo (ICME)*,pp. 1–6. IEEE.

[6] Hankinson, A., Roland, P., & Fujinaga, I. (2011). The Music Encoding Initiative as a Document-Encoding Framework. In *Proceedings of the 12th International Society for Music Information Retrieval Conference* (ISMIR 2011), pp. 293–298

[7] Kallionpää, M., Greenhalgh, C., Hazzard, A., Weigl, D., Page, K. R., & Benford, S. (2017). Composing and realising a game-like performance for disklavier and electronics. In *New Interfaces for Musical Expression* (NIME 2017).

[8] De Roure, D., Klyne, G., Pybus, J., Weigl, D. M., & Page, K. (2018). Music SOFA: An architecture for semantically informed recomposition of Digital Music Objects. In *Proceedings of the 1st International Workshop on Semantic Applications for Audio and Music* (pp. 33-41). ACM.

7.2 List of abbreviations

Abbreviation	Description
API	Application Programming Interface
CE	Trompa Contributor Environment
EMO	Early Music Online
HTML	HyperText Markup Language
IMSLP	International Music Score Library Project
MELD	Music Encoding and Linked Data
OMR	Optical Music Recognition
PDF	Portable Document Format
REST	REpresentational State Transfer
TEI	Text Encoding Initiative
URI	Uniform Resource Identifier

XML	EXtensible Markup Language