SIMSSA DB

Symbolic Music Discovery and Search

Emily Hopkins
McGill University
Music Research
Montreal, QC, Canada
emily.hopkins@mcgill.ca

Yaolong Ju
McGill University
Music Research
Montreal, QC, Canada
yaolong.ju@mail.mcgill.ca

Gustavo Polins Pedro
McGill University
Music Research
Montreal, QC, Canada
gustavo.polinspedro@mail.mcgill.ca

Cory McKay
Marianopolis College
Liberal & Creative Arts
Montreal, QC, Canada
cory.mckay@mail.mcgill.ca

Julie Cumming
McGill University
Music Research
Montreal, QC, Canada
julie.cumming@mcgill.ca

Ichiro Fujinaga
McGill University
Music Research
Montreal, QC, Canada
ichiro.fujinaga@mcgill.ca

ABSTRACT

The SIMSSA DB is designed to meet the specific research needs of musicologists and others who could benefit from a unified repository and interface for discovering, accessing, and contributing symbolic music files. It was originally conceived as part of the optical music recognition workflow developed for the SIMSSA (Single Interface for Music Score Searching and Analysis) Project (simssa.ca). However, the popularity of specialized projects such as the Josquin Research Project, KernScores, and the SEILS dataset implied the need for a general-purpose database of symbolic files for researchers to discover, access, and contribute files in a range of formats for any genre or era of music. To this end, SIMSSA DB offers novel metadata and content-based search functionality, prioritizing matters of musicological importance, including provenance tracking, high-quality metadata harvesting, the ability to model complex musical relationships, and the experimental reuse of research datasets.

KEYWORDS

Symbolic music, metadata, feature-based search

1 INTRODUCTION

The current SIMSSA DB was originally conceived as part of the SIMSSA Project [1], which focuses on analyzing, organizing, and distributing the machine-readable symbolic music files derived from Optical Music Recognition (OMR). However, OMR remains error-prone for the moment, and most computational musicology research is currently conducted on symbolic music files that are hand transcribed or significantly corrected from OMR by researchers.

Current symbolic music databases include KernScores [23], the SEILS dataset [21], and Classical Music Archives [12]. These are valuable resources but have limitations in terms of search capabilities, formats supported, metadata quality, and scope. The Josquin Research Project [2] is another example of a symbolic music resource that addresses some of these limitations, and its popularity suggests that these resources are needed by the music research community [2, 16] Currently, best-case scenarios for making symbolic music files available involve researchers storing files on their own custom website (e.g., Measuring Polyphony [3]) or using general-purpose repositories (e.g., the Josquin Research Project's GitHub repository [4]).

The SIMSSA DB is the successor of an older database which was created as part of Cumming's Digging into Data Challenge grant (2012–14) [5], designed for use with the now-deprecated Counterpoint Web App [24]. This older database had a user-friendly interface, but the data model was not sufficiently complex. The first version of our new data model was developed in 2017 [15], and the SIMSSA DB data model is designed to be inclusive of as many eras and formats as possible, including MEI, MusicXML, MIDI, Kern, Sibelius, etc.

This paper will 1) give an overview of the data model for the SIMSSA DB, 2) introduce the web interface design including download, upload, and search, with emphasis on the potential of feature-based content search, 3) highlight the importance of provenance, and 4) address larger issues such as archiving research data sets, linked data, copyright, and contributing back to common resources.

2 THE "WORK" CONCEPT IN THE SIMSSA DB DATA MODEL

Building an effective database requires modelling a great deal of bibliographic metadata about symbolic music files. A key issue is the distinction between an abstract "musical work" and the file(s) instantiating it, which are what the database actually contains. An example of an "abstract work" is Beethoven's third symphony—not a particular score or edition, just the abstract idea of a combination of musical elements that make up "Beethoven's Third". This idea of a work existing outside material reality is a useful abstraction for organizing music; however, an abstraction cannot be digitized into a symbolic music file. One must digitize a particular edition, using a particular kind of software, and making various other decisions about how to represent this work. For example, "Beethoven's Third" could be stored as manifested in Breitkopf& Härtel's 1862 edition,

[1]https://simssa.ca/
[2]http://josquin.stanford.edu/
[3]https://measuringpolyphony.org/
[4]https://github.com/josquin-research-project
[5]"Electronic Locator of Vertical Interval Successions (ELVIS) (2012-2014; https://elvisproject.ca/)

as made available in a MusicXML file from the IMSLP/Petrucci Music Library site. [6]

The IFLA-LRM and other library-based conceptual models provide a good starting point for addressing such issues. In the IFLA-LRM, a work is "the intellectual or artistic content of a distinct creation"—so it maps on to "Beethoven's Third Symphony the abstract work," not a material object [22, p. 20]. The expression level is for "a distinct combination of signs conveying intellectual or artistic content" (e.g., a translation of a libretto or a piano reduction of a symphonic work), and a manifestation indicates a particular edition, by a particular publisher, for example, Breitkopf and Härtel in 1862. [22, p. 23] An item is then a specific object—not just a particular edition, but an actual copy, such as a file on the IMSLP/Petrucci Music Library website.

The SIMSSA DB model deviates from this slightly, as shown in Table 1. Essentially, the SIMSSA DB collapses works and expressions, instead describing that dynamic by relating works to other works. This is similar to BIBFRAME (Bibliographic Framework Initiative) [7], which was designed to transition away from MARC (MAchine-Readable Cataloging) and work with linked data. This is done in order to prioritize musical content equivalence. The SIMSSA DB notion of a "source" maps quite closely to a manifestation, and a file maps to an "item"—because, in the end, we are storing and describing individual files.

Table 1: The IFLA-LRM compared to the SIMSSA DB.

IFLA-LRM	SIMSSA DB
Work: Beethoven Symphony No. 3	**Work:** Beethoven Symphony No. 3
Expression: piano reduction, study score	**Related work:** piano reduction
Manifestation: particular edition (e.g., Breitkopf & Härtel, 1862)	**Source:** particular edition (e.g., Breitkopf & Härtel, 1862)
Item: specific object (the file as it exists stored on IMSLP)	**File:** specific object (the file as it exists stored on IMSLP)

This means that instead of defining one work and many expressions, we typically define works and related works. For example, a symphony could be made available in a piano reduction. A work can also be divided into sections, such as an opera aria, or a movement in a symphony. In this model, it is possible to relate sections to other sections as well. For example, musical material could be reused in different masses, or an opera aria made available with a piano part as a separate publication. The SIMSSA DB prioritizes showing when musical content is similar or equivalent—a primary concern for music researchers.

In addition to managing this mapping from files to works, we have many other metadata fields to capture information. Instead of a "composer" field, there is a "contributor" drop-down menu, allowing the user to select Composer, Arranger, Author of Text, Transcriber, Improviser, or Performer. It is also possible to indicate whether the attribution of a contribution is certain or uncertain. The database also allows works to be designated secular or sacred, for the music where this applicable. Currently, genre and form are addressed by separating them into two fields—genre as in style (e.g., Romantic) and genre as in type (e.g., minuet). These are a sample of some of the fields; the full data model is described in our developer documentation [6].

3 THE WEB INTERFACE

Anyone can search or download symbolic music files, no user account required. The search interface combines both metadata and content-based search, where the latter is implemented using jSymbolic to extract 246 features from each symbolic file on upload [14]. Features in this context refer to any information about a piece of music that can be conveyed with a numerical value, including pitch statistics; melodies and horizontal intervals; chords and vertical intervals; rhythm; instrumentation; texture; and dynamics. jSymbolic is still being actively developed, with new features being added regularly. This feature search allows us to combine searching for metadata and content in novel ways. Users can start with a keyword search, then refine results using metadata facets or feature value sliders.

In addition to searching with jSymbolic features, users can also download complete feature value files for each symbolic music file, for use in machine learning or statistical analysis experiments using many features per piece. For example, in [13], jSymbolic features were used to train a machine learning model to differentiate between Renaissance, suggesting ways to identify the composer of pieces with uncertain attribution, and to learn more about their respective stylistic characteristics empirically.

Users do need to create an account in order to upload files. Anyone can create an account (they do not have to be approved, or institutionally affiliated), making it easier for users to save groups of files they are interested in (either in a research corpus or download "cart"). Only the original uploader (or an administrator) can change or delete content. Requiring user accounts, instead of allowing anyone to make changes, is also a slight deterrent to spam or abusive uploads and will help correct any such activity more swiftly. Of course, more sophisticated oversight may be needed as the DB grows, such as refining user account permissions, content moderation, and guarding against denial-of-service attacks.

Building the upload interface is in some ways like training someone to catalog music on the fly, for an item that does not have cataloging rules. Users are not only adding files—they are entering metadata as well. Aiming for ease of entry as well as linked-data compatibility, the database is designed with a "preference for identifiers over textual strings"; where possible, metadata is harvested from external sources such as VIAF and the Library of Congress [25].

The SIMSSA DB also adopts an approach inspired by RISM's Muscat [8]. Of particular interest, RISM has an important relationship with VIAF. When entering persons in a new RISM record, Muscat defaults to the RISM records [8]. If the person cannot be found, Muscat allows the user to look it up in VIAF to try to complete the record from there. Users can still add a new entry even if no match can be found, and are prompted to verify the creation of a

[6]https://imslp.org/
[7]https://www.loc.gov/bibframe/
[8]http://muscat-project.org/

new item, helping to ensure that they really meant to add the new item and did not make an error. Allowing new persons or terms is an attempt at striking a balance between highly controlled quality and flexibility: users may have needs we cannot anticipate or be studying repertoire not well-described by existing vocabularies.

The first phase of adding new files to the database requires the user to enter title, sections, contributor, and genre metadata. The second phase allows the user to identify sources and upload files. Sections entered in the first phase are automatically included so the user can attach files to either the whole work or any of the given sections as needed.

4 PROVENANCE

A strong emphasis is placed on recording provenance information in the SIMSSA DB. This information is essential to musicologists, and provenance (both in terms of original source and digital encoding) impacts machine learning experiments, which may be skewed by quirks of individual programs and encoding workflows [5, 17]. The provenance of a symbolic music file is often multi-layered, comprising information about the particular, manuscripts, print editions, and/or digital files from which its music was transcribed. Information on provenance allows users to make their own judgments about the suitability of any given file for their needs.

The "source" field is one key aspect of tracking provenance; it describes the particular edition, manuscript, or other document (physical or electronic) from which a symbolic music file was generated. Sources can be related to other sources—so the edition that an encoder consulted could itself be based on an earlier manuscript. The DB also has a field to indicate a "collection of sources," such as a hymnal or fake book, and the "archive" field, indicating a repository of documents (physical or digital). Researchers assemble symbolic music datasets from already existing files (e.g., the Josquin Research Project), newly transcribed files (OMR, manual transcription), or a combination of the two, (e.g., files from an online repository edited according to another source.)

It is necessary to be able to represent different kinds of sources, e.g., original manuscripts, early printed reproductions, modern editions, and online collections. The current form allows one immediate source (e.g., IMSLP) as well as a chain of "parent" sources (e.g., Breitkopf & Härtel, 1862). An important priority for future work is modelling more complex relationships such as "sibling" sources at the same level. Even if a user may not need the entire history of a contributed file, it is still important that they be able to re-trace the original uploaders' process, so as to be able to understand their editorial decisions, correct errors, or look for additional pieces in the same repertoire.

5 MANAGING RESEARCH DATASETS

In addition to tracking where files come from and how they were generated, users can also group them within the database as "Research Corpora." The user can include any file in the database in a corpus, not only files they upload themselves. As in streaming music playlists, a piece can belong to many corpora simultaneously. Corpora are designed to be publicly viewable, so that others can use them to repeat, validate or expand upon research already done with them, but private collections may be added in the future. These corpora serve as an important discovery tool for finding other researchers' work or discovering the collections of files in the database.

Transcribing and correcting symbolic files is laborious. It is important to be able to preserve them well, both so that other researchers can use them and so that experiments can be repeated in the future. Long-term digital preservation is a complex problem [1]. For now, we approach this in three ways. First, while datasets are being worked on, they go in GitHub [9]. This allows us to track changes and collaborate easily, and prevents things from getting lost under layers of permissions in services like Google Drive [10]. Second, SIMSSA DB then serves as a useful place to upload finished files, add relevant metadata, store features, and make work discoverable. However, it poses some challenges for long-term storage of experimental data.

In many projects, researchers use symbolic files for machine learning or other large-scale automated analyses, where having a stable, accurate dataset is essential. Given that we have made it possible for users to edit their own contributions to the database, it is not possible to guarantee that a file will remain identical over time. Therefore, we also make it possible to integrate the SIMSSA DB with external repositories, which can store a complete, static version of an experimental dataset. Currently we are working with Zenodo [11], a service developed by CERN (European Organization for Nuclear Research) and popular in open science communities.

Zenodo can also mint DOIs (Digital Object Identifiers). A DOI is a persistent identifier for a given digital object (whereas a URL, or Uniform Resource Locator, indicates only a location). Creating a DOI for a dataset makes it easier for other researchers to cite it, and increased citations is an important part of encouraging scholarly activity in terms of dataset creation. These DOIs will also be used to connect the SIMSSA DB and external repositories, allowing users to discover published datasets through our main search interface.

The database is currently seeded with three high-quality collections of files from recent projects:

1. RenComp7, 1584 pieces [13]
2. Josquin-La Rue Secure Duos Dataset (JLSDD), 108 duos [5]
3. MS Florence 164–167, 116 pieces [4]

There are relatively few high-quality symbolic datasets for early music available, so these files represent a meaningful initial contribution, and the database will continue to expand.

Dataverse [12] is another option for data storage. Unlike Zenodo, a single centrally-hosted repository, various Dataverse instances are hosted by institutions, and use the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) [13] to promote interoperability between instances. SIMSSA DB will be part of the Dataverse pilot project at McGill University, collaborating with McGill librarians. This is in line with the "inside-out" model of library services, where libraries are increasingly concerned not only with bringing external resources in for their local audience ("outside-in" library service) but to support the "inside-out" sharing

[9]https://github.com
[10]https://www.google.com/drive/
[11]https://zenodo.org/
[12]https://dataverse.org/
[13]https://www.openarchives.org/pmh/

of resources created at an institution, making them available to both local and external audiences [7].

6 LINKED DATA

Linked data involves storing information in a way that makes the relationships between entities explicit, thereby enabling semantic search. Instead of only being associated as keywords, linked data makes it possible to define the relationships between objects and concepts in a concrete way, allowing users to construct searches based on particular meanings of terms instead of just looking for strings. For example, a semantic search could ask for all pieces composed by Beethoven. This is typically achieved by storing data in RDF (Resource Description Framework) as subject-object-predicate "triples." For example, "Beethoven—composed—Symphony No. 3": not only is Beethoven linked to Symphony No. 3, but the nature of the relationship between the two is identified. Each part of a triple is expressed as a URI (Uniform Resource Identifier) that uniquely identifies it. Weigl et al. provide a good overview of other music resources using linked data, as well as guidance on the uses of linked data for "aggregated access across datasets" [26]. Many other music databases are also pursuing linked data approaches. For example, Weissenberger has incorporated linked data into the design for the Linked Irish Traditional Music (LITMUS) project for the Irish Traditional Music Archive [27]. Hankinson and Craig-McFeely detail uses for linked data in medieval music in the context of DIAMM (Digital Image Archive of Medieval Music) [14] [10].

The SIMSSA DB builds in linked-data compatibility by using external resources that assign URIs. VIAF combines multiple name authority files from different institutions into a single authority service, assigning URIs for linked data as well. For example, in VIAF, Beethoven's ID is 32182557, which has the associated unique URI: viaf.org/viaf/32182557/. Such identifiers connect back to all the different records from libraries from which VIAF has gathered data. Variant spellings, extra names, and different dates can all be united by VIAF URIs, and the data can also be queried. Linked data resources such as VIAF also provide good support for multilingual use, with URIs linking data in many different languages and scripts. The SIMSSA DB fills as many fields as possible with terms from resources that assign URIs (VIAF and the Library of Congress). There are also plans incorporate data from RISM and DIAMM in the future.

Linked data can also help track the provenance of metadata, making it possible to identify who made a particular contribution. Such provenance annotations will take the form of linked data quads. These expand the original linked data triple to include provenance. Thus, if a triple reads "This composer—*composed*—this mass", the quad expands it to: "This composer—*composed*—this mass—*according to*—this historian."

7 CONTRIBUTING TO COMMON RESOURCES

In the process of using services such as VIAF or LC vocabularies, users may discover gaps in the available information, especially when studying early music or other highly specialized repertoire. Therefore, work is being done on developing processes to make additions to common resources such as RISM or VIAF, or even suggest new terms to LC vocabularies. (There is a process for creating a proposal for a new term [20], facilitated by the SACO Music Funnel [15] for music-related terms [19].)

The existing relationship between RISM and VIAF provides a good example of this: as previously mentioned, RISM uses VIAF to import personal name authority files for Muscat. In addition to this, as libraries add new records to Muscat, RISM can send new information back to VIAF on a regular basis, improving VIAF for everyone in the process. We plan to incorporate RISM's data into our upload functionality in the future, cultivating a similar relationship. There is some precedent for research projects contributing back to RISM, such as the Moravian Music Foundation [18], and we are looking forward to collaborating with RISM both in terms of contributing data, such as inventories for chant manuscripts, as well as building a process that others can replicate.

Another issue that arises in building a website of free content is licensing and redistribution, including educating researchers on these topics. Most academics are acquainted with copyright laws allowing them to use materials in fairly permissive ways (e.g., Canadian copyright law has "fair dealing"). Sharing and redistributing files online, however, has different implications. For example, files from the Josquin Research Project are available in their GitHub repository [16] under a CC-BY-SA 4.0 license [17]. This license requires anyone using these files to specify attribution, link to the license, and highlight changes, and modified versions of the original files must be distributed under the same license. Provided these requirements are met, files can be shared freely, including copying and redistribution, and even adapted for commercial purposes. To address this, the SIMSSA DB includes a licensing field, and we also take care to use appropriate licenses for new datasets.

8 FUTURE WORK

Currently we are implementing changes from our first rounds of user testing, working to ensure that it is user-friendly for musicologists and other target users. We are aiming to release a first version in summer of 2020 and continue user testing with a broader audience. Additional plans include expanding the search to include polyphonic and harmonic tools, both under development as part of SIMSSA. David Garfinkle developed PatternFinder for polyphonic search [9], allowing users to search for a melody or chord progression across all the voices in a given file. SIMSSA researchers [3, 11] are also working on developing automated harmonic analysis, which will be adapted for harmonic search to find particular chords or progressions over many pieces. Ultimately, SIMSSA DB will be integrated into the SIMSSA Project's OMR workflow.

ACKNOWLEDGMENTS

This research is part of the SIMSSA Project, which is funded by a Partnership Grant from the Social Sciences and Humanities Research Council of Canada (SSHRC 895-2013-1012) and also funded by the Fonds de recherche du Québec-Société et culture (FRQSC).

[14]https://www.diamm.ac.uk/

[15]https://www.loc.gov/aba/pcc/saco/Music_Funnel.html

[16]https://github.com/josquin-research-project/jrp-scores

[17]https://creativecommons.org/licenses/by-sa/4.0/

REFERENCES

[1] Raymond Blum and Betsy Beyer. 2019. Achieving digital permanence. *Commun. ACM* 62, 5 (April 2019), 36–42. https://doi.org/10.1145/3303872

[2] Andrew W. Brinkman, Daniel Shanahan, and Craig Stuart Sapp. 2016. Musical Stylometry, Machine Learning, and Attribution Studies : A Semi-Supervised Approach to the Works of Josquin. In *Proceedings of the 14th International Conference on Music Perception and Cognition*. ICMPC, San Francisco, CA, United States, 91–97.

[3] Nat Condit-Schultz, Yaolong Ju, and Ichiro Fujinaga. 2018. A Flexible Approach to Automated Harmonic Analysis: Multiple Annotations of Chorales by Bach and Prætorius. In *Proceedings of the 19th International Society for Music Information Retrieval Conference*. ISMIR, Paris, France, 66–73.

[4] Julie E. Cumming and Cory McKay. 2018. Revisiting the origins of the Italian Madrigal. Medieval and Renaissance Music Conference, Maynooth, Ireland.

[5] Julie E. Cumming, Cory McKay, Jonathan Stuchbery, and Ichiro Fujinaga. 2018. Methodologies for creating symbolic corpora of Western music before 1600. In *Proceedings of the 19th International Society for Music Information Retrieval Conference*. ISMIR, Paris, France, 491–8.

[6] DDMAL. 2018. Welcome to SIMSSADB's documentation! — SIMSSADB 0.1 documentation. https://elvis-project.github.io/simssadb/html/index.html

[7] Lorcan Dempsey. 2015. A New Information Management Landscape: From Outside-in to Inside-out. In *New Roles for the Road Ahead: Essays commissioned for ACRL's 75th anniversary*. The Association of College & Research Libraries, a division of the American Library Association, Chicago, IL.

[8] Répertoire International des Sources Musicales (RISM). 2019. Home - RISM. http://www.rism.info/home.html

[9] David Garfinkle, Claire Arthur, Peter Schubert, Julie Cumming, and Ichiro Fujinaga. 2017. PatternFinder: Content-Based Music Retrieval with Music21. In *Proceedings of the International Workshop on Digital Libraries for Musicology*. ISMIR, Shanghai, China, 5–8.

[10] Andrew Hankinson and Julia Craig-McFeely. 2017. Building the new DIAMM: Linking and sharing data for medieval musicology. In *Conference Abstracts of Digital Humanities*. Alliance of Digital Humanities Organizations, Montreal, QC, 69–73.

[11] Yaolong Ju, Samuel Howes, Cory McKay, Nat Condit-Schultz, Jorge Calvo-Zaragoza, and Ichiro Fujinaga. 2019. An Interactive Workflow for Generating Chord Labels for Homorhythmic Music in Symbolic Formats. In *Proceedings of the 20th International Society for Music Information Retrieval Conference*. ISMIR, Delft, Netherlands, n.p.

[12] Classical Archives LLC. 2019. Classical Music on Classical Archives: Home. https://www.classicalarchives.com/

[13] Cory McKay, Julie Cumming, and Ichiro Fujinaga. 2017. Characterizing Composers using jSymbolic2 Features. 18th International Society for Music Information Retrieval Conference (Late-Breaking), Suzhou, China.

[14] Cory McKay, Julie Cumming, and Ichiro Fujinaga. 2018. jSymbolic 2.2: Extracting Features from Symbolic Music for use in Musicological and MIR Research. In *Proceedings of the 19th International Society for Music Information Retrieval Conference*. ISMIR, Paris, France, 348–54.

[15] Cory McKay, Andrew Hankinson, Julie Cumming, and Ichiro Fujinaga. 2017. A Database Model for Computational Music Research. DLfM '17 4th International Workshop on Digital Libraries for Musicology (poster), Shanghai, China.

[16] Cory McKay, Tristano Tenaglia, Julie Cumming, and Ichiro Fujinaga. 2017. Using statistical feature extraction to distinguish the styles of different composers. In *Abstracts of the Annual International Medieval and Renaissance Music Conference*. MedRen, Prague, Czech Republic, 111.

[17] Néstor Nápoles, Gabriel Vigliensoni, and Ichiro Fujinaga. 2018. Encoding Matters. In *DLfM '18: Proceedings of the 5th International Conference on Digital Libraries for Musicology*. Association for Computing Machinery, New York, NY, USA, 69–73.

[18] OCLC. 2013. Moravian Music Foundation. https://www.oclc.org/en/member-stories/moravian.html

[19] Library of Congress. 2010. SACO Music Funnel Project - SACO - Program for Cooperative Cataloging (Library of Congress). https://www.loc.gov/aba/pcc/saco/Music_Funnel.html

[20] Library of Congress. 2019. Process for Adding and Revising Library of Congress Subject Headings | Cataloging and Acquisitions (Library of Congress). https://www.loc.gov/aba/cataloging/subject/lcsh-process.html

[21] Emilia Parada-Cabaleiro, Anton Batliner, Alice Baird, and Björn W Schuller. 2017. The SEILS Dataset: Symbolically encoded scores in modern-early notation for computational musicology. In *Proceedings of the 18th International Society for Music Information Retrieval Conference*. ISMIR, Suzhou, China, 575–581.

[22] Pat Riva, Patrick Le Boeuf, and Maja Žumer. 2017. *IFLA Library Reference Model: A Conceptual Model for Bibliographic Information*. International Federation of Library Associations and Institutions, Den Haag, Netherlands.

[23] Craig Stuart Sapp. 2005. Online Database of Scores in the Humdrum File Format. In *Proceedings of the 6th International Society for Music Information Retrieval Conference*. ISMIR, London, UK, 664–665.

[24] Peter Schubert and Julie Cumming. 2015. Another Lesson from Lassus: using computers to analyse counterpoint. *Early Music* 43, 4 (Nov. 2015), 577–86.

[25] Dean Seeman and Lisa Goddard. 2015. Preparing the Way: Creating Future Compatible Cataloging Data in a Transitional Environment. *Cataloging & Classification Quarterly* 53, 3-4 (May 2015), 331–340. https://doi.org/10.1080/01639374.2014.946573

[26] David M. Weigl, David Lewis, Tim Crawford, Ian Knopke, and Kevin R. Page. 2019. On providing semantic alignment and unified access to music library metadata. *International Journal on Digital Libraries* 20, 1 (March 2019), 25–47. https://doi.org/10.1007/s00799-017-0223-9

[27] L.K. Weissenberger. 2018. Linked data in music and its potential for ITMA/traditional music web resources. *Brio* 55, 1 (2018), 52–57.