Music Encoding Conference, May 2020

MEI and Verovio for MIR: A Minimal Computing Approach

Natalia Ermolaev, Ph.D.
Associate Director
Center for Digital Humanities
Firestone Library Floor B
Princeton University
Princeton, NJ 08544
(609) 258-6243
nataliae@princeton.edu

Mark Saccomano, Ph.D.
(510) 759-9854
Columbia University
Department of Music
2960 Broadway
New York NY 10027
m.saccomano@columbia.edu

While the increase in digital editions, online corpora, and browsable databases of encoded music presents an extraordinary resource for contemporary music scholarship, using these databases for computational research remains a complex endeavor. Although norms and standards have begun to emerge, and interoperability among different formats is often possible, researchers must devote considerable time to discover, learn, and maintain the skill sets necessary to make use of these resources. This talk will discuss our efforts to minimize the burden of technological expertise required to disseminate and access encoded music through the use of open-access tools and their integration into an online search and display using the principles of minimal computing. Minimal computing is a design philosophy that seeks to maximize access to digital materials through reducing reliance on specific hardware and software requirements.[1] We will examine the advantages of using MEI to encode archival sources of notated music. MEI is a practical choice not only because of its adaptability and increasingly common use in digital musicological projects, but also due to the availability of Verovio, a convenient means for displaying the encoded files in a web browser. It will also compare our project's functionality with that of other sites and detail future directions for the utility and extensibility of the model.

The Serge Prokofiev Archive as Data project (https://mss2221.github.io/spademo/) began as an exploration into the possibilities for the representation and display of archival collections using available open source tools, programs, and platforms. The occasion for the

[1] Jentery Sayers, "Minimal Definitions," https://go-dh.github.io/mincomp/thoughts/2016/10/02/minimal-definitions/ (2016).

project was Columbia University's acquisition in 2014 of the Serge Prokofiev Archive, a vast collection of nearly 20,000 items—including original manuscripts, correspondence, published material, audio-visual items, and photographs—related to the years that the Russian composer Sergei Prokofiev (1891-1953) lived and worked in the United States and Europe. An early project goal was devising a way to display the notated music contained in handwritten notebooks Prokofiev used to record his musical ideas. The results were encouraging, and a prototype was developed using MEI, Verovio, and Jekyll that could display an engraved version of the score derived from a digitally encoded version of the manuscript, as well play the score in a browser using a simple interface. The prototype proved to be extensible to other collections of symbolic music data, making this suite of open source tools extremely useful for researchers and archivists who would like to disseminate their work to a larger audience. Our talk will elaborate on the process involved in bringing this prototype to fruition and the steps involved in porting the design to other corpora of notated music.

This project represents the first steps toward creating a template that integrates search, retrieval, and analysis of a simple repository of MEI files. It also introduces a simple workflow for creating the encoded files. Notation software was used to initially enter the music contained in the sketchbooks. Scores were then exported as MusicXML and transformed to MEI through a web-based tool offered by Verovio.[2] Verovio is especially useful because it can also transform MEI data into a MIDI stream that can be played in a browser. For archival materials such as unlabeled sketch scores, this can aid in identification of fragments, as well as increase access to the musical holdings beyond those with the skill to imagine what the music sounds like from notation alone. To allow for browsing of encoded materials, we chose a static site generator, Jekyll, with its own set of readily available templates. One of these templates, "Ed.", is specifically designed for minimal editions of literary works.[3] We were able to adapt Ed. to display and play MEI files as well, making this template especially useful for music with associated text and lyrics.

We would like our prototype to be a resource to scholars in music studies—an example of open data and code that will lessen the demand for technical expertise for both the researcher and the user, while demonstrating the functionality that can be added to a single site accessed through an ordinary web browser. Full realization of this project's potential, however, remains stifled due to the difficulty of devising appropriate technical solutions for MIDI playback. Workarounds had to be devised to accommodate how different browsers and operating systems handle audio output. We expect that improvements in web audio technology and browser compatibility will soon make such workarounds unnecessary.

While the raw data of much notated music may be ready to be downloaded for analysis, the high-level computing skills required to retrieve and analyze that data means that it remains out of reach to many. In order to make collections such as these more accessible, computing resources for encoding, retrieval, and display need to be made available to both

---

[2] https://www.verovio.org/musicxml.html

[3] "Ed. A Jekyll Theme for Minimal Editions," https://elotroalex.github.io/ed/

researchers and end users. Much like Stanford University's Josquin Research Project, the Serge Prokofiev Archive as Data represents an effort to consolidate the features of search, display, and analysis of encoded music scores, but through the use of a minimal computing model. It thus not only facilitates access to musical data, but also simplifies dissemination of work by individual researchers by providing a template that does not require knowledge of server operations or database design.