A Data Model for Digital Musicology and its Current State – The Music Encoding Initiative

Kepper, Johannes

kepper@edirom.de
University of Paderborn

During the last 10 years, XML has gained general acceptance as a data model in the Digital Humanities. Actually, it even leveraged the success of digital projects in the humanities. Meanwhile TEI is the unchallenged standard for all projects in the fields of literature studies, epigraphy, linguistics, history sciences and so on. Many thoughts were invested to bring TEI and other related formats like EpiDoc to a level that suffices general scholarly needs.

At first sight, things went differently in the field of music encoding. Around the year 2000 a couple of XML-based encoding schemes for music notation emerged, and within just a few years MusicXML became the best-known and most widespread music encoding format. It was intended to serve as an interchange format between different music applications, and even today it is virtually indispensable for this very important task. At the same time, this orientation of MusicXML requires a certain "simplicity" that facilitates implementations in various applications.

The Music Encoding Initiative (MEI) went a different way. Not aiming at application support in the first place, an encoding model for *scholarly* purposes was developed over years. Strongly influenced by the concepts of the TEI, Perry Roland as initiator of the format tried to transfer these concepts to the field of music encoding. For large parts, this is quite easy: music notation is a kind of text, and many unspecific modules of TEI can be reused for music encoding with only small changes. But then again, music notation itself offers a much higher complexity than other texts. It is multi-dimensional not only because of its layout of multiple vertically aligned staves, but also because of its simultaneity of harmonic and melodic progression. In music notation, the text itself consists of overlapping hierarchies and therefore demands a quite sophisticated data model. Most often, it is virtually impossible to preserve all possible meanings (or better: interpretations) of a musical text with reasonable effort. The reason is that the written text is only a part of the complete information. Every notation serves a certain purpose, and each composer or copyist uses only as many symbols as he needs to be explicit to his contemporaries. Besides this, the "rules" of music notation changed significantly over time, even though these developments often seem to be very subtle.

All this leads to the problem that there is no absolutely fixed terminology in music notation. Some phenomena are still not completely understood or even defined, such as the problem of dots, strokes and hooks in scores from the classical period. The lack of a complete and well-defined terminology even for restricted repertoires makes the encoding of music notation on a scholarly level highly demanding, and, at the same time, the implementation and usage of such an encoding scheme is anything but trivial.

The Music Encoding Initiative has chosen this way, and currently it stands on an important turning point: In a one-year project funded by the NEH and DFG the original model was revised and has proven to meet all essential scholarly requirements for such a format. In the next years, it needs to be disseminated in the fields of musicology, music information retrieval, music philology and digital humanities in general. A first step in this direction is the TEI's Special Interest Group on music encoding, whose members were actively involved in the recent developments on MEI, and who seek to find ways to bring MEI and TEI closer together.

Due to the complexity of music notation – and thus music notation encoding too – application support for MEI is crucial to ensure its dissemination: Almost no traditional musicologist would be willing to work with a XML-editor like Oxygen. There are several projects currently working on such applications for MEI: The DiMusEd-Project, situated in Tübingen (Germany), uses SVG to render encodings of multiple sources of music notated with medieval neumes. Although this repertoire

uses a limited set of symbols, this project already shows the benefits of a dynamic rendering from an encoding instead of engraved scores. The Edirom project, (Detmold, Germany) aims to establish workflows for digital scholarly editions of music. In the application for preparing such editions it is already using MEI to store all structural information about the musical text as well as the containing documents. For moving from basically facsimile-based editions to completely digital editions it is planned to offer complete encodings of all relevant sources including the rendering-facilities already demonstrated by the DiMusEd-project. In order to achieve this goal Edirom closely collaborates with the most ambitious of all ongoing MEI-related projects: TextGrid. A sub-project of this major German initiative, which is also located in Detmold, seeks to develop a limited scorewriter for MEI offering a graphical user interface for musicologists. In this case „limited" means that the project neither intends to support MEI completely nor tries to keep up with the engraving quality of already existing scorewriters: the unambiguity of the output is more important than its beauty.

All these German projects collaborate closely with the ongoing efforts in the US to further improve the format itself and to provide interchange to other relevant formats such as Humdrum and MusicXML. Depending on further funding by NEH and DFG respectively it is intended to provide reasonable collections of MEI encodings to facilitate further usage of the format. Although MEI will not find the wide acceptance MusicXML already has, all these components will help to disseminate MEI in the academic world, to promote interchange of high-quality data and to explore new methods for digital representations of written music.

The talk will provide a short introduction to the current state of MEI – both the format itself and the projects and applications already working on and with it.