

Linked Data for Music Collections: A User-Centred Approach

Proposal for Digital Humanities 2013 Conference

1 November 2012

Jonathan Grimes, MPhil in Digital Humanities, Trinity College Dublin, Ireland

Professor Séamus Lawless, Trinity College Dublin, Ireland

The emergence of the Linked Data movement offers music collections with the potential to change how they structure and share their data^[1]. The ability to use Linked Data to combine music metadata with other information sources offers rich semantic possibilities for both access to and promotion of music, and provides considerable scope for music collections to remodel their data in a way which makes a significant contribution to the growing web of data^[2].

Previously, collections have focused on building and maintaining relational database models and have used, where possible, common metadata schemas to help build interoperability into their collections. While this has achieved a degree of interchange of structured data between collections, much of this approach has been limited by the type of data structure, the standards used, and above all the development of customised solutions such as APIs for sharing such data^[3]. The technologies associated with Linked Data hold the promise of enabling music collections to achieve a high level of semantic interchange with other information resources, and examples of such interchange might include the combining of music metadata with geospatial, historical or biographical data^[4]. The advantage of such an approach would allow music archives to focus on their core data collection (i.e. music metadata) and achieve a streamlining of their workflow. The publishing of algorithmic data related to musical form and content as linked data also offers the possibility of developing new applications for music information retrieval^[5], and allows exciting possibilities for music searching and recommendation.

This paper will investigate a number of approaches to the publication of music collections as Linked Open Data. It will explore the motivations for, and benefits of, taking this approach to encoding and exposing musical data. A critical appraisal of the state of the art technology will be outlined and its suitability with relation to music collections discussed. The paper will also explore some of the key changes in maintaining a digital collection and will attempt to address how the notion of a centralised collection has changed as a result of this emerging technology. This impending paradigm shift is one which challenges the notion of data ownership and places organisations and collections in a different position when it comes to managing data.

The paper uses the development of a database system for the Contemporary Music Centre (CMC), a national archive and resource centre for Irish composer's music, to illustrate the design and theoretical processes involved in enabling some of the Linked Data technologies. This database tracks the

metadata on Irish composers' compositions and associated materials and is currently being redesigned as part of a new web site which is under development. In addition to replacing the current structure and underlying technologies which power the database, the project also wishes to future-proof the system for the medium-term and take advantage of the emerging trends in digital collection management.

In deciding upon the particular approach to follow to deploy Linked Data in this system, a series of interviews examining the different user groups' attitudes and views on the linking of CMC's content to external data was carried out. The feedback received from these interviews, when combined with the in-depth review of the state of the art, was a key input into developing an approach for the CMC's content. The results of this user-based research also helped to inform the design of a road map of how an approach to Linked Data for a music data source might be taken. The research also illustrates some of the potential problems with the technology and suggests some ways in which these problems might be overcome.

The user research reveals a number of notable findings in relation to attitudes towards the linking of external data to music information. While many identify with the idea of a distributed music collection, users are cautious about exposing and linking music-related data and content to external non-music-related content. Key among users is the need for such a Linked Data-driven system to ensure accuracy of data and maintain the trust of its users. There was also a strong view that such external links needed to be curated by both CMC and the users, and that the Linked Data provided should balance with and enhance the core collection data rather than overshadow it.

The solution proposed as a result of this user research involves the restructuring and mapping of CMC's database to Music Ontology classes and concepts and presents a number of approaches involving the presentation of the remodelled data.

-
1. Raimond, Yves, Christopher Sutton, and Mark Sandler. "Interlinking Music-related Data on the Web." IEEE Computer Society (2009). ↩
 2. Raimond, Yves, and Mark Sandler. "A Web of Musical Information." Conference Proceedings on Music Information Retrieval. ↩
 3. Heath, Tom, and Christian Bizer. Linked Data: Evolving the Web into a Global Data Space. Ed. Frank Van Harmelen & James Hendler. Vol. 1. Morgan & Claypool, 201. ↩
 4. Kobilarov, Georgi, Tom Scott, and Yves Raimond. "Media Meets Semantic Web—how the BBC Uses DBpedia and Linked Data to Make Connections." The Semantic Web (2009). ↩
 5. Dixon, Simon et al. "Towards a Distributed Research Environment for Music Informatics and Computational Musicology." Journal of New Music Research 39.4 (2010). ↩