Exploring Early Music Through Linked Data

Tim Crawford (Goldsmiths University of London), David Lewis, Kevin Page, David Weigl (Oxford e-Research Centre)

The music collections

Within the Transforming Musicology project, one research theme was music of the 16th Century. This theme built on the presence of two corpora: **Early Music Online (EMO)** and the **Electronic Corpus of Lute Music (ECOLM)**.

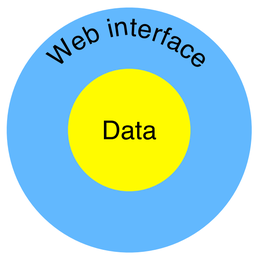
Early Music Online

EMO consists of British Library early modern printed music, with an enriched catalogue and digitally scanned microfilm images of 342 books, totalling about 8,500 musical works.

Electronic Corpus of Lute Music

ECOLM is a collection of about 2,400 transcriptions of pieces for lute and accompanying catalogue information. Many of the pieces date from the same period as the EMO collection – some are drawn from the books in EMO.

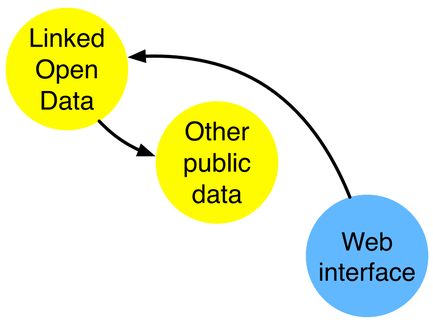
Data silos vs Linked Open Data



Traditional catalogues and online editions (including EMO and ECOLM) put structured data into a private database. Public access to the information is made possible by a bespoke user interface.

This interface acts as a gatekeeper for the data. It constrains the uses to which the data can be put.

If the software becomes outdated, the data may become inaccessible. There is no trivial way to copy or archive the contents of database without direct access.



As part of the Transforming Musicology project – and with help from the Semantic Media Network – we have translated a subset of the EMO/ECOLM data into Open Linked Data and published it online.

This means others can build interfaces of their own with our data.

We connect records of people and places with external resources such as dbpedia, VIAF and MusicBrainz. This allows us to use external information in our interfaces and makes it easier for others to use our data in a broader context.

Since the data is now published in a generic way, we can try various interface approaches, using bespoke or pre-existing tools.

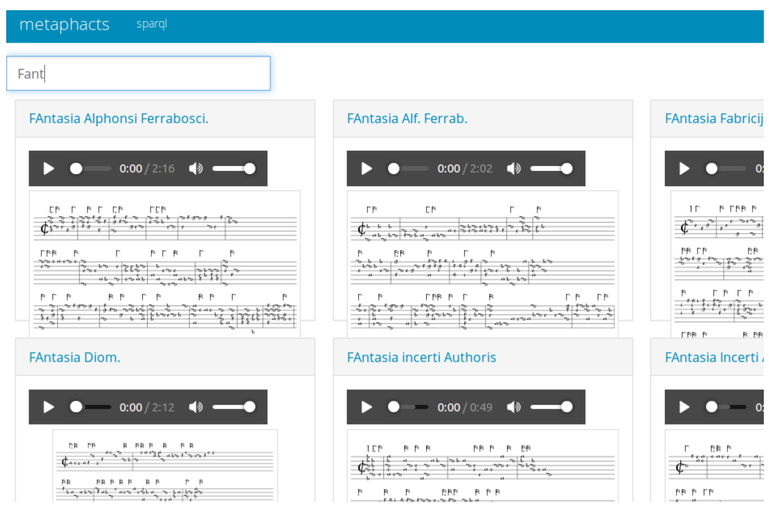
Interfaces to Linked Open Data

Funding from the **Semantic Media Network** allowed us, in collaboration with the BBC, to link to their **Early Music Show** records reflecting broadcasts of hundreds of episodes of their current weekend show, including 2,807 works.

An interface was built to pull information from the BBC resource, EMO, ECOLM and external sites such as Wikipedia to generate information pages around the broadcast schedules.

Since one key argument for publishing data directly, rather than just through a website, is the ability to apply other interfaces, we can test this by applying software devised for other cultural heritage contexts.

ResearchSpace, initiated and led by Dominic Oldman of the British Museum and built by Metaphacts, is a framework for exploring digitised collections and catalogues published as Linked Data.



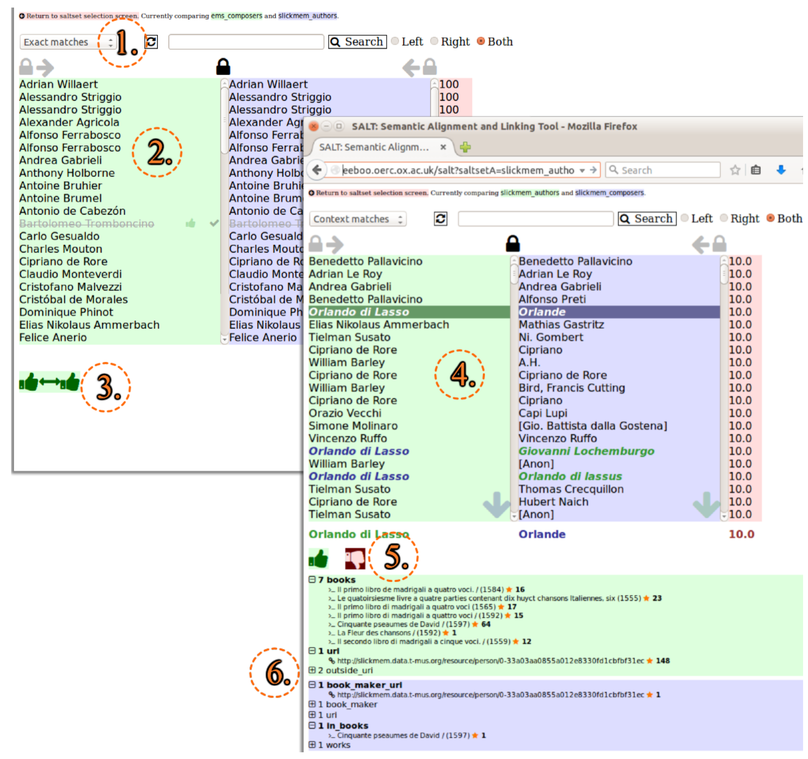
Using this framework on our data we can, for example, browse music of ECOLM, filtering the list, previewing the tablature and hearing audio samples – functionality not available on the ECOLM web interface.

Making links

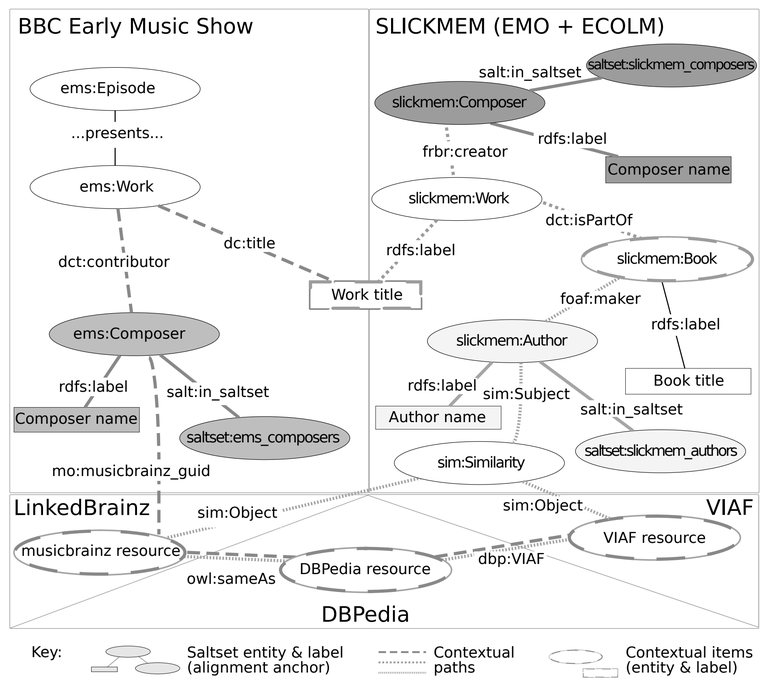
An important part of making a resource useful is context. Information sources become more powerful by combining them – a process that researchers undertake manually all the time. By linking to other resources, we hope to enable richer and more natural research questions to be asked.

Linking text fields in a database is laborious, but also susceptible to errors (e.g. misidentifying an author). Often, expert knowledge is needed for disambiguation, and sometimes the context needed to inform a decision is not available directly in the database.

The Semantic Alignment and Linking Tool (SALT) assists a cataloguer in making and annotating links between resources. The tool presents a web interface that can use its own primitive reasoning to suggest matches, but which also allows the cataloguer to use their own judgments.



Links made using this tool are published as Open Linked data, including the match, the reasoning behind the match and who made that connection.



References:

David M. Weigl, David Lewis, Tim Crawford, Ian Knöpke and Kevin Page (2017). "On providing semantic alignment and unified access to music-library metadata". *International Journal of Digital Libraries*. https://doi.org/10.1007/s00799-017-0223-9

Tim Crawford, Ben Fields, David Lewis and Kevin Page (2014). "Explorations in Linked Data practice for early music corpora". In *Proceedings of the Joint Conference on Digital Libraries (JCDL 2014)*, pp 309-12. DOI:10.1109/JCDL.2014.6970184

Richard J. Lewis, Tim Crawford and David Lewis (November 2015). "Exploring information retrieval, semantic technologies and workflows for music scholarship: the Transforming Musicology project". *Early Music* 43.4, pp. 635-47. https://doi.org/10.1093/em/cav073

OXFORD e-Research CENTRE
UNIVERSITY OF OXFORD
Arts & Humanities Research Council
Goldsmiths UNIVERSITY OF LONDON
Semantic Media Network a new paradigm for navigable content for the 21st Century
Transforming Musicology