Metadata Infrastructure for Sound Recordings

Catherine Lai[1], Ichiro Fujinaga[1], David Descheneau[2], Michael Frishkopf[2], Jenn Riley[3], Joseph Hafner[4], and Brian McMillan[4]

Music Technology, Schulich School of Music, McGill University[1]; Department of Music, University of Alberta[2];

Digital Library Program, Indiana University[3]; McGill University Libraries, McGill University[4]

Introduction

This poster describes the first iteration of a working model for searching heterogeneous distributed metadata repositories for sound recording collections.

As an expanding number of digital sound recording collections emerge, the task of searching at individual repository sites becomes **impractical** for end users. For example, users need to:

- learn multiple interfaces located at different sites
- manually combine search results
- meet repeated authorization requirement

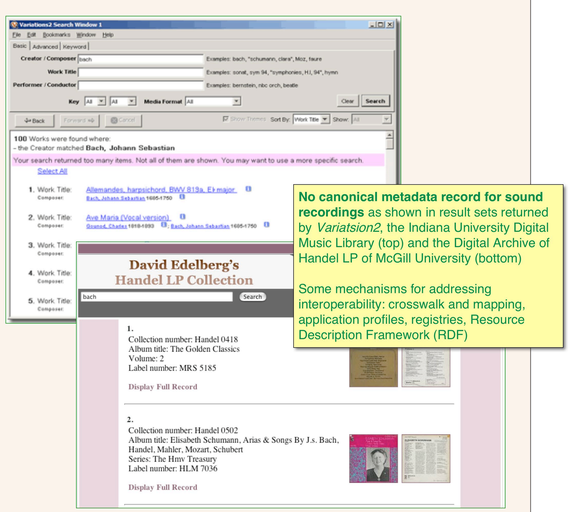
The **benefits** of a distributed search have increasingly become apparent for users because it:

- searches across multiple repositories
- hides the difference in the mechanics of interactions
- facilitates the flow of data between applications

Metadata Integration and Interoperability Issues

Institutions have adopted different metadata models for sound recordings in digital libraries to enhance their music systems. These differences lead to:

- multiple protocols
- multiple metadata schemes
- multiple formats of data

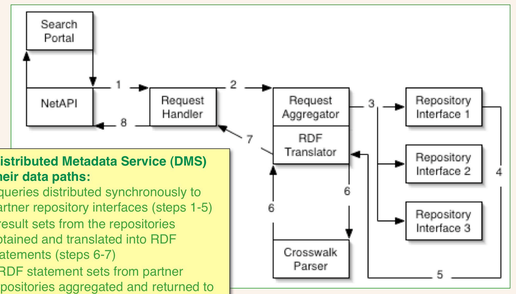


No canonical metadata record for sound recordings as shown in result sets returned by *Variatsion2*, the Indiana University Digital Music Library (top) and the Digital Archive of Handel LP of McGill University (bottom)

Some mechanisms for addressing interoperability: crosswalk and mapping, application profiles, registries, Resource Description Framework (RDF)

System Function

The core function of the metadata infrastructure for sound recordings supports the following processes:



The Distributed Metadata Service (DMS) and their data paths:

- queries distributed synchronously to partner repository interfaces (steps 1-5)
- result sets from the repositories obtained and translated into RDF statements (steps 6-7)
- RDF statement sets from partner repositories aggregated and returned to the search portal for display (steps 7-8)

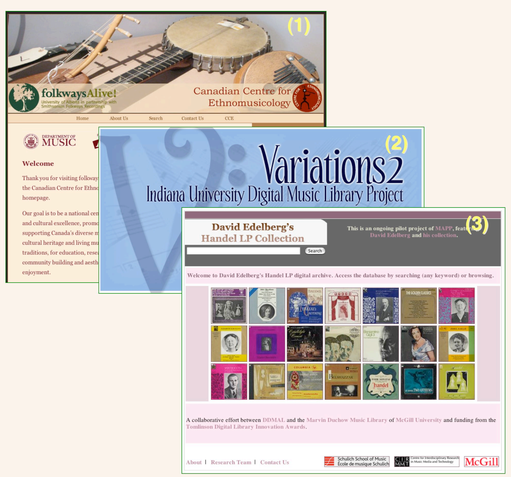
Implementation

Three universities hosting heterogeneous digital repositories of sound recordings participate in the prototype of cross-repository interoperability, each implementing a **variety of connection methods and search APIs**:

(1) *FolkwaysAlive!* at University of Alberta, with connection to an RDF server

(2) *Variations2* of Indiana University, with connection to a MySQL database

(3) Digital Archive of Handel LP at McGill University, with HTTP connection with custom XML query API



Distributed Metadata Services

The Distributed Metadata Services (DMS) perform real-time querying of existing metadata repository services, in the context of the federation:

(1) all requests intercepted and passed to the Request Handler

(2) query requests forwarded to the Request Aggregator

(3) query string synchronously passed to all partner repository interfaces, which are specified in a **crosswalk configuration file**:

```
<itst_crosswalk>

  <collection label="fw">
    <namespace uri=http://mudoc.org/meta/sfmeta# prefix="sfmeta" />
     <equiv dc=http://purl.org/dc/elements/1.1/title
     local="http://mudoc.org/meta/sfmeta#title" />
     <equiv dc="http://purl.org/dc/elements/1.1/creator"
     local="http://mudoc.org/meta/sfmeta#displayname" />
     <equiv dc="http://purl.org/dc/elements/1.1/creator"
     local="http://mudoc.org/meta/sfmeta#displayname2" />
     <equiv dc="http://purl.org/dc/elements/1.1/date"
     local="http://mudoc.org/meta/sfmeta#yearreleased" />
     <equiv dc="http://purl.org/dc/elements/1.1/publisher"
     local="http://mudoc.org/meta/sfmeta#recordlabel" />
  </collection>

  <collection label="variations2">
     <namespace uri="http://server1.variations2.indiana.edu/meta/elements#"
     prefix="variations2" />
      <equiv dc="http://purl.org/dc/elements/1.1/title"
     local="http://server1.variations2.indiana.edu/meta/elements#DisplayTitle" />
     <equiv dc="http://purl.org/dc/elements/1.1/date"
     local="http://server1.variations2.indiana.edu/meta/elements#PublicationDate" />
     <equiv dc="http://purl.org/dc/elements/1.1/pub
     local="http://server1.variations2.indiana.edu/m
  </collection>

  <collection label="handel">
     <namespace uri="http://coltrane.music.mcgill.ca/handel/lp/terms#" prefix="handel" />
     <equiv dc="http://purl.org/dc/elements/1.1/Identifier"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#id" />
     <equiv dc="http://purl.org/dc/elements/1.1/Title"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#albumSetTitle" />
     <equiv dc="http://purl.org/dc/elements/1.1/Title"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#albumSetAlternativeTitle" />
     <equiv dc="http://purl.org/dc/elements/1.1/Title"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#albumSetVariationsTitle" />
     <equiv dc="http://purl.org/dc/elements/1.1/Language"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#language" />
     <equiv dc="http://purl.org/dc/elements/1.1/Title"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#uniformTitle" />
     <equiv dc="http://purl.org/dc/elements/1.1/Creator"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#statementResponsibility" />
      <equiv dc="http://purl.org/dc/elements/1.1/Source"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#labelNumber" />
     <equiv dc="http://purl.org/dc/elements/1.1/Source"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#genericLabelName" />
     <equiv dc="http://purl.org/dc/elements/1.1/Source"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#transcribedLabelName" />
     <equiv dc="http://purl.org/dc/elements/1.1/Relation"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#seriesStatement" />
     <equiv dc="http://purl.org/dc/elements/1.1/Relation"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#editionStatement" />
     <equiv dc="http://purl.org/dc/elements/1.1/Identifier"
     local="http://coltrane.music.mcgill.ca/handel/lp/terms#collectionOnlineAccess" />
  </collection>

  <common label="Dublin Core 1.1">
     <namespace uri="http://purl.org/dc/elements/1.1/" prefix="dc" />
  </common>

</itst_crosswalk>
```

The crosswalk configuration file, which defines the repositories supported at the DMS site and specifies valid local-to-common metadata equivalences.

(4) crosswalk configuration read at runtime with each search

(5) metadata result sets transformed into RDF statements by the RDF Translator:

```
<rdf:Description rdf:about="http://coltrane.music.mcgill.ca/handel/lp/collection.php?0606">
<handel:albumSetTitle rdf:nodeID="node_0606_1"/></rdf:Description>
        <rdf:Description rdf:nodeID="node_0606_1">
        <handel:albumSetTitle>KATHLEEN FERRIER</handel:albumSetTitle >
        </rdf:Description>
<rdf:Description rdf:about="http://coltrane.music.mcg
<handel:language rdf:nodeID="node_0606_2"/></rdf:
        <rdf:Description rdf:nodeID="node_0606_2">
        <handel:language>English</handel:language>
<rdf:Description rdf:about="http://coltrane.music.mcgill.ca/handel/lp/collection.php?0606">
<handel:statementResponsibility rdf:nodeID="node_0606_3"/></rdf:Description>
        <rdf:Description rdf:nodeID="node_0606_3">
        <handel:statementResponsibility>ARIAS BY PURCELL AND HANDEL
        </handel:statementResponsibility ></rdf:Description>
```

RDF transformation example of an excerpt of McGill Handel LP Collection metadata. RDF preserves source metadata elements and their semantics and relationships.

(6) metadata statement predicates in the RDF converted to a common (i.e., Dublin Core) equivalents, according to the custom metadata-to-DC equivalences specified in the configuration file:

```
<rdf:Description rdf:about="http://coltrane.music.mcgill.ca/handel/lp/collection.php?0606">
        <dc:Title>KATHLEEN FERRIER</dc:Title>
        <dc:Language>English</dc:Language>
        <dc:Creator>ARIAS BY PURCELL AND HANDEL</dc:Creator>
</rdf:Description>
```

(7) metadata combined for all repositories and returned as search results for display

Challenges & Future Work

The absence of uniformity in the repository API and metadata implementation raised issues in areas of search and retrieval, system performance, and presentation. For example:

- different treatments of Boolean operators, wildcard operators, and query matching algorithm between the APIs affect the results of federated search operations
- custom configuration of performance metrics at the individual repositories, which is often implemented and optimized for local use, influences the overall performance
  - ➢ incorporate parallel/asynchronous access to partner metadata repositories
  - ➢ use of a central cache to cache queries and asset metadata in an RDF store
- mapping of custom metadata to different common elements is possible, thus jeopardizes the quality of presentation
  - ➢ provide visible trace of the query
  - ➢ enforce consistent and precise mappings from local to common schemes
  - ➢ implement authority control

Acknowledgment

This research is, in part, funded by the ITST Research Grant of SSHRC, Canada.