

# OPEN DATA MANAGEMENT & CLOUD EXAM PROJECT

AUDIO MUSIC FILE ARCHIVING

PATRICK INDRI

MAY 6, 2020

# STRUCTURE OF THE PROJECT

## Aim of the project

Investigation of audio file archiving for music.

In particular:

- ▶ UML metadata model;
- ▶ XSD implementation and XML sample document;
- ▶ discussion of data discovery/access and interoperability;
- ▶ discussion of long term archiving and data preservation.

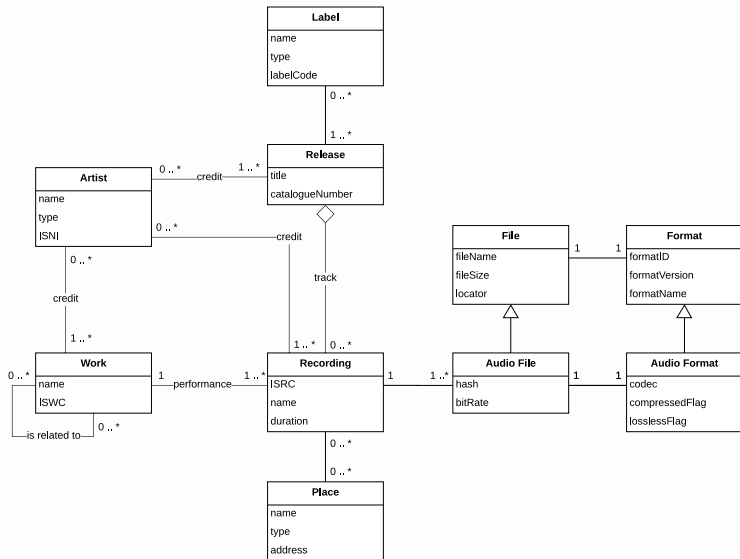
**Data resource:** not an actual dataset but music files in general.

There is no widely used and standardised metadata model for music audio files.

- ▶ Dublin Core: simple (15 terms), focus on descriptive metadata;
- ▶ EbuCore: detailed DC extension, fine grain technical and administrative metadata for broadcasting;
- ▶ METS: handles the structural/hierarchical metadata of a digital library. Open flexibility (no vocabulary).

What should the data model represent?

- ▶ Songs and their different versions;
- ▶ Groups of songs (releases);
- ▶ Artists;
- ▶ Basic technical metadata;
- ▶ Relations between songs, releases and artists.



Choice of implementation:

- ▶ RDB: easy to enforce constraints (primary/foreign keys), widely used, easy to model relationships, rigid structure;
- ▶ **XSD**: flexible, easily handle partial data, harder relationship handling.

The proposed XSD implementation should:

- ▶ Refine Dublin Core;
- ▶ Balance integrity constraints and partial data;
- ▶ Model relationships with detail.

The resulting XSD can be retrieved [here](#).

# XML EXAMPLE

Example of an XML document, valid against the proposed XSD.

```
<work>
  <ISWC id="ISWC_T-000.000.000-A"></ISWC>
  <title lang="en">
    <dc:title>Test Work</dc:title>
  </title>
  <hasArtist label="Will Wilson" description="Singer">
  </hasArtist>
  <hasPerformance label="Test Rec." description="Studio Ver.">
    <relationIdentifier>
      <ISRC idref="ISRC_AAAAA0000000"></ISRC>
    </relationIdentifier>
  </hasPerformance>
</work>
```



# DIFFICULTIES AND POSSIBLE EXPANSIONS

# DATA DISCOVERY: SEARCH/FILTER SERVICE

# DATA ANNOTATION

# STORAGE AND CLOUD SOLUTIONS

# DATA PRESERVATION

# INTEROPERABILITY

# AUDIO FILE FORMATS

# FINAL CONSIDERATIONS



# REFERENCES