

# Descriptive Metadata for Audio-Oriented Digital Collections

by John Huck, MLIS

<https://orcid.org/0000-0003-1150-5402>

This report—originally in the form of a website<sup>1</sup>—was created by John Huck in March of 2010 as a capping exercise in partial fulfillment of the requirements for the Master of Library and Information Studies degree at the University of Alberta, and was based on research completed as part of a directed study supervised by Dr. Ali Shiri undertaken in the spring and summer of 2009. The research was presented at the *Forum for Information Professionals* held by the Library and Information Studies Students' Association on February 5, 2010, for which slides are available.

The report was reformatted in May of 2020 for deposit into the University of Alberta's Education and Research Archive (ERA). Some minor format changes were made and an error in the column headings of *Table 4* was corrected at that time.

## Abstract

This research project asks whether audio material in online digital collections might have specialized descriptive metadata requirements, and, if so, what they might be. To answer these questions, an overview of current practice was undertaken. Eighteen audio-oriented online collections were identified and sample records from each were examined to discover the character of the metadata elements used in each collection. The results were interpreted both in the context of individual collections and on an aggregate level. The findings provide a new perspective on the interdisciplinary nature of online audio-oriented collections and their metadata requirements.

## Introduction

In recent years, the online environment has had a profound impact on music and the recording industry. In fact, some accounts of the internet boom around the turn of the last century have suggested that file sharing of music had a direct impact on the swift adoption of broadband internet access (Hartley, 2009). Since that time, file sharing has become commonplace, shifting from Napster to torrent technology, while legitimate online sellers like iTunes have entered the marketplace, and database services like Naxos Music Online have begun selling access to

---

<sup>1</sup> previously available at <http://capping.slis.ualberta.ca/cap10/JohnHuck/>

streaming audio. While these online entities account for a vast proportion of the public's current procurement of music, they are primarily mechanisms for the distribution of commercial recordings.

Non-commercial or unpublished recordings have traditionally been the concern of sound archives. The variety in these types of recordings is broad, including field recordings of various kinds, broadcast recordings, oral histories, and early recordings on fragile formats like 78s, wax cylinders and the like. The institutions that collect them often fall somewhere between the traditional worlds of archives, libraries and museums. For these institutions, the emergence of a networked world has meant new channels for distribution through the development of digital libraries and repositories. But it has also meant that the already complex aspects of descriptive cataloguing of this material have been transposed into the equally complex world of metadata. Zeng and Qin (2008) recommend that choices about metadata schema be based on an analysis of functional requirements. These types of online collections run the gamut from digital libraries, to archival catalogues with sound files, to open platforms for hosting user generated content. Despite this diversity, it seemed natural to ask whether sound recordings have specific metadata requirements that need to be taken into account in online audio collections.

Analysis of metadata requirements typically leads to the selection of one of the many formal metadata schemas that define metadata element sets. Schema comparisons and development of crosswalks between them constitute one approach to metadata research (Corthaut, Govaerts, Verbert, & Duval, 2008). While this approach is useful from a design point of view for digital libraries, it does not examine metadata as actually implemented. Many schemas, such as Dublin Core, may be modified, qualified, or used and interpreted in different ways (Cole & Shreeves, 2004); furthermore, not all elements in a given set are mandatory, and some may be rarely used. Therefore, it was decided that the best approach for this research project was to study metadata implementations in real collections, and to compare metadata elements rather than schemas. This approach would allow an appraisal of requirements for audio content that reflected its unique characteristics, independent of schema considerations. There was a second practical consideration for limiting the scope to elements instead of schemas, which was that detailed documentation about the schemas in use was available for some collections, but non-existent for others. It was also decided that the scope would be limited to elements in metadata records as publicly displayed, which meant focusing on descriptive metadata, since structural and technical metadata is not usually intended for human reading, and therefore not always provided. The benefit of this empirical method, though, was that comparisons could be made between collections that were quite disparate in nature.

The primary research question of this study is whether the nature of audio material in online digital collections leads to specialized descriptive metadata requirements, and if so, what they might be. A secondary question is what lessons we might learn from the practices of current collections.

The question of what to call the entities in this study requires some explanation. The possibilities included digital libraries, online collections, digital archives and digital repositories.

The choice was not easy, since, as shown in *Table 1*, the sponsoring institutions were of many kinds. Borgman identifies two perspectives on the term digital library: one that sees them as "content collected on behalf of user communities" and the other that sees them as "institutions or services" (Borgman, 2001, p. 35). Both perspectives describe my selected sites, but the word library seemed imperfect to describe material clearly archival in nature. "Digital archives" was ruled out by analogy and "repository" was rejected because it was thought to imply an institutional process rather than a curated presentation. As a compromise, the term "online collection" offered a solution. It is true that the archival community distinguishes between holdings and collections, since "collection" implies an artificially created grouping, rather than an organically derived fonds, but, since archives make selections from those holdings when they put audio online, the term seemed appropriate in this case. The phrase "audio-oriented" was adapted from Michael Lesk's term "text-oriented" as a way of characterizing collections whose main focus is on audio recordings, but that might contain supporting material in other formats, such as photographs or transcripts.

## Literature Review

### Traditional Approaches to Cataloguing in Sound Archives

Metadata requirements for audio recordings overlay the challenges of description that sound archives have always faced. Understanding the landscape of descriptive practice in cataloguing for sound recordings helps contextualize the question of requirements by highlighting the diversity of disciplinary practice.

Standard bibliographic and archival rules for description (AACR2, RAD, DACS) include rules for describing recordings, and these are used by most general libraries and archives. Additionally, specialized rules for recordings have been developed to supplement the main standards. AACR2 has proved a valuable touchstone, providing a set of rules that other standards have built upon. The Canadian Rules for Archival Description (RAD) and the U.S. equivalent, Describing Archives: A Content Standard (DACS), both relate to AACR2 in this way; RAD mimics the chapter structure of AACR2. Specific cataloguing rules for archival sound recordings have been issued by the Association for Recorded Sound Collections (Association for Recorded Sound Collections [ARSC], 1995) and the International Association of Sound and Audiovisual Archives (International Association of Sound and Audiovisual Archives [IASA], 1999) and both of these also incorporate references to the AACR2 chapter structure. The IASA rules, though, reflect a more international perspective, making reference to ISBD rules, and drawing on other frameworks, including FRBR, RAD, the ARSC rules, and the SAA's Oral History Cataloging Manual, amongst others.

While the notion of 'sound archives' suggests a network of collections, united by a common format, in fact the term amalgamates a diversity of disciplines. An IASA guide to sound archives (Lance, 1983) identifies several categories of sound archives, including: broadcasting, commercial recordings, dialect and linguistic collections, ethnomusicology, folklore, oral history

and natural history. The extensive set of examples in appendix B of the IASA cataloguing rules (IASA, 1999) also bears out this diversity. The recognition of this unity in diversity has not always been the case. For instance, the preface of the ARSC rules indicates that, when they were developed by the Associated Audio Archives, that organization was "at this time most interested in cataloging 78 rpm and cylinder recordings" (ARSC, 1995, p. ix). Public Archives Canada published a guide to procedures for sound archives in 1979 (Public Archives Canada [PAC], 1979) but it was primarily a description of how that institution managed its holdings. The guide distinguished PAC in its function of collecting spoken word recordings of debates, and so forth, from the National Library, which collected published recordings, and the Museum of Man, which collected recordings that "document the folk music and folk culture of Canada" (PAC, 1979, p. 2). The guide predates RAD by eight years, and so does not make reference to AACR2, but the narrow focus it takes is clearly related to the institutional structure from which it originates. The boundaries between different kinds of sound archives and institutions remain difficult to pin down today.

Some institutions have contended with collections of mixed commercial and archival recordings from the beginning. As one of the first ethnomusicology archives in North America, the Indiana Archive for Traditional Music, first founded in 1936 at Columbia University by George Herzog, followed the cataloguing model in use at the Berlin Phonogramm-Archiv, where Herzog had worked under Erich von Hornbostel (Archives of Traditional Music [ATM], 1975). The collection focus on oral tradition did not discriminate between commercial recordings, field recordings and broadcast transcriptions and the catalogue method therefore had to accommodate all three. The unit of description was the fonds and specific rules for constructing a title were specified. The catalogue was arranged by accession number, but supplemented with indexes to geographical areas, culture groups, subjects, collectors, performers, etc., and recording companies. With the exception of subjects, all of these elements were included in the constructed title, with the first element being geographical areas. In this way, all three types of recordings could be captured by a single cataloguing system.

## Digital Sound Archives

If traditional descriptive practice in archival audio collections was usually determined by individual approaches, based on local needs, the transition to the digital environment has brought with it new challenges overtop of the old. Bradley has observed that "much of the effort devoted to metadata in the heritage sector has focused on descriptive metadata as an offshoot of traditional cataloguing" (Bradley, 2009, p. 21), and warns against the dangers of approaching the problem of metadata as a traditional cataloguing problem that can be solved with a greater level of detail. Sound archives need to address the issue of metadata, however, because digital archives are only going to grow in importance for preservation and access in the coming years. A survey of the state of audio collections in academic libraries (Smith, Allen, & Allen, 2004) found that unique and fragile collections are at risk, because of a critical need for both preservation and access through cataloguing. Because there are no new physical recording formats being developed for the archival community, "there is little choice for sound

preservation except digital storage approaches" (Bradley, 2009, p. 4), so archives are looking to digital archives for both preservation and access purposes.

## Metadata

Metadata is important for essential processes in the digital environment beyond mere description. A popular definition calls metadata "data associated with either an information system or an information object for purposes of description, administration, legal requirements, technical functionality, use and usage, and preservation" (Baca, 1998). Relationship metadata, provenance and content ratings have also been cited as relevant categories (Lagoze, Lynch, & Daniel, 1996). Metadata implementation is different than cataloguing, in that the abundance of schemas means choices must be made even before description can begin. In making choices, Zeng and Qin recommend considering "the nature of collection objects; anticipated user needs; and constraints upon metadata creation, implementation and quality control" (Zeng & Qin, 2008, p. 88). Zeng and Qin also remind metadata designers to consider the needs of both end-users and system users, as metadata must act "both as inventory and user access tool" (Zeng & Qin, 2008, p. 93). A thorough understanding will support schema selection and the identification of desired elements in the schema.

## Multimedia Metadata

Choosing a metadata schema for multimedia is even more challenging, because "the multimedia domain is far too wide for any single standard" (Smith & Schirling, 2006, p. 86). Even amongst music schemas (e.g. ID3, MPEG-7, etc.) optimal choice can vary greatly depending on intended function (Corthaut, Govaerts, Verbert, & Duval, 2008). The CUIDADO project found when managing large music collections that a combination of automated metadata and bottom up descriptions was useful (Vinet, Herrera, & Pachet, 2002). Another approach, pioneered by the well-regarded Variations2 Indiana University Digital Music Library, has been to incorporate a work-based approach, following the FRBR model, as a way to resolve ambiguities that arise from music content specifically (Notess & Dunn, 2004).

Nevertheless, the need to harmonize approaches across the community of practitioners has been recognized (Bradley, 2009; Lai, et al., 2007; Smith & Schirling, 2006). Lai, et al., found that the heterogeneity of schemas and systems in use amongst three sound repositories hindered the operations of a federated search protocol. Bradley argues for versatility and extensibility as important principles for sound-related metadata, amongst others, so that schema can be combined as needed in application profiles, customized for particular needs, but still be capable of interoperability. One emerging area of interoperability is the use of RDF for semantic web queries. A combination of conventional metadata, enhanced with contextual user supplied information is considered as a possible path forward, especially for music related resources (Moutselakis & Karakos, 2009).

The need to address music and sound retrieval in the archival and folklore areas as a multi-faceted problem has recently been recognized. The EASAIER project (Enabling Access to Sound Archives Through Integration, Enrichment and Retrieval) has explored ways to use multiple techniques for idiomatic retrieval of music recordings, speech recordings and 'cross-media' materials (like printed scores), respectively (Damnjanovic, Barry, & Reiss, 2008). The heterogeneity of folklore collections that include sound recordings has been identified as a challenge to using pre-existing metadata schemas. Development of application profiles that integrate multiple schemas has been attempted as a possible solution (Lourdi, & Papatheodorou, 2004).

## Methodology

### Overview

The research question called for an empirical method to examine current practice in the field of online audio-oriented collections. The process included the following steps:

- identifying candidate collections for the study
- establishing selection criteria
- selecting 18 collections for inclusion in the study
- collecting sample records and documentation from each collection
- analyzing the data in the collection information and element sets

### Identifying Candidate Collections

In this study the selection criteria were not established first, because it was not clear at the outset what kinds of collections were actually online. It would have been difficult to establish an a priori definition of online audio-oriented collections that included all of the examples ultimately selected. The discovery of resources was, in fact, a significant aspect of the project.

The process of identifying candidates began with the examination of member lists of the following associations:

- Association of Recorded Sound Collections (ARSC)  
<http://www.arsc-audio.org/>
- International Association of Sound and Audiovisual Archives (IASA)  
<http://www.iasa-web.org/>
- Music Library Association (MLA)  
<http://www.musiclibraryassoc.org/>
- Society for Ethnomusicology (SEM): Guide to Programs in Ethnomusicology  
<http://webdb.iu.edu/sem/scripts/guidetoprograms/guidelist.cfm>

Next, online resource lists were explored, including the following:

- British Columbia Digital Library: Directories, Guides, Portals and Search Engines to Digital Libraries  
<http://bcdlib.tc.ca/guides.html>
- New York Public Library: Best of the Web: Recorded Sound and Moving Image Collections  
<http://www.nypl.org/weblinks/2723>
- Rutgers University Libraries: Subject Research Guides: Jazz  
[http://www.libraries.rutgers.edu/rul/rr\\_gateway/research\\_guides/jazz/jazzresearch.shtml](http://www.libraries.rutgers.edu/rul/rr_gateway/research_guides/jazz/jazzresearch.shtml)

## Selection Criteria

The discovery process resulted a list of over thirty candidate collections, which ranged from national archives to digital library on specialized topics and everything in between. After evaluating the characteristics of these collections, criteria were developed to make a selection of sample collections that would be easy to compare with each other. The selected collections all had the following characteristics:

- Content focus on recorded sound
- Search interface
- Availability of a significant number of recordings for streaming or downloading, either full recordings or short excerpts
- Publicly accessible and free to use
- Relatively stable and permanent set of content
- Operating with the law

The criteria excluded entities like torrent sites for mp3s, commercial services like iTunes or Naxos Music Online, sites with changing content like YouTube or online sellers like Beatport.com, online catalogues without sound files like Indiana University's Archives of Traditional Music, password protected sites like the Variations3 digital library, and websites that provided only a limited number of sample recordings. Where full recordings were not available, the extent of available samples was considered when deciding whether to include a collection. In fact, the provision of excerpted content was not uncommon, and usually reflected a practical approach to the challenge of offering access to copyright material. The remaining collections shared a common set of characteristics, but remained quite diverse in content. Further purposeful sampling was conducted to arrive at a manageable sample size that maintained a diversity of content. Eighteen online collections were included in the final sample selection.

Note that the collections will be referred to throughout this study with acronyms, which may be found in the list of selected collections in *Appendix A*.

## Collecting Data

Data collecting took place between May and November, 2009, and consisted of two parts. First, basic information was gathered about each collection, including information about its host or sponsor, the types of content it included, significant collection features, as well as any available documentation. Next, sample records were taken from each collection. Wherever possible, a variety of search approaches were employed to find the sample records. For most collections, five records were collected. In three cases, more than five records were collected, because of the variety of types of recordings offered in the collections. These collections were ASR (8 records), CHARM (8 records) and DLA (6 records). The process for capturing each record involved displaying the record, copying the text on the page to a plaintext file and taking one or more screenshots to aid in later interpretation.

## Analyzing Data

Data was analyzed in three phases. The first phase compiled into tables the information about collection hosts (*Table 1*), content types (*Table 2*), and collection features (*Table 3*). The second phase analyzed the metadata elements on a collection level, which allowed rough comparisons to be made between collections, in terms of the level of detail in the metadata and the consistency of the cataloguing. The results of this process are recorded in *Table 4*. The third phase analyzed the metadata elements on an aggregate level to look for commonalities and general trends within the sample set of collections. The results of this process are recorded in *Table 6* and summarized in *Table 5*. The findings recorded in these tables will be discussed in detail in the following section, but a few words should be said about the procedures used in the second and third phases of analysis.

In phase two, elements in each sample record were counted, and the mean number of elements per record for each collection was calculated. The sets of elements and their values from all the sample records for a given collection were then compiled into a master list of all elements observed in that collection. Duplicates were eliminated, and each element was marked with a number to indicate the number of sample records it appeared in. Special note was made of elements that occurred in more than 75% of the samples for a given collection.

In phase three, all of the elements that had occurred in at least 75% of the samples for a given collection (usually meaning four or five occurrences) were compiled into a master list of 223 elements. A reference to the collection each element came from was maintained. The list was then manually sorted into 17 groupings. Using a kind of affinity grouping, the groupings emerged from the data set, and then names were assigned to them. The number of elements included in each grouping, as well as the number of collections represented in each grouping were recorded. Finally, the groupings themselves, were divided among seven metadata categories identified by Lagoze, Lynch, & Daniel (1996), becoming de facto sub-categories of those categories. These authors did not intend for their categories be considered definitive and complete, since they were merely illustrative rather than exclusive. However, the categories have proved useful over time and have been used in this way by others (Greenberg, 2005).

## Considerations in Determining Affinity Groupings (Phase Three)

Greenberg (2005) points out that some kinds of metadata serve dual purposes, and this was found to be true when considering identification numbers and format descriptions. It was decided that the grouping "Identification Numbers" would include retrieval numbers, shelf numbers, and the like, but also issue numbers and matrix numbers. The logic behind the idea of combining these numbers into a single category was that that, while issue and matrix numbers are important for describing published recordings, equivalent numbers for archival recordings, such as reel numbers and accession numbers, are less relevant for description and play a role as retrieval numbers. Ultimately, though, the value of these numbers is that they act as unique identifiers for the recordings, regardless of whether for discographic or retrieval purposes, and it seemed better to create a single grouping for them on the basis of this commonality, and to consider them as a form of administrative metadata. This does go against the spirit of the ARSC rules, which identifies label and issue or matrix numbers as vital descriptive elements, but those rules were primarily created to catalogue commercial recordings, so it seems reasonable to make some allowances.

Elements that dealt with the formats of the original sound carriers were combined with elements indicating file components and assigned to the structural metadata category. While this category usually refers to the sequence and relations of segments of electronic files, it was interpreted more freely in this case. The use of the seven categories of Lagoze, Lynch, & Daniel (1996) was, in any event, an adaptation of sorts, since this study was only concerned with the publicly displayed metadata, which tended to be mostly descriptive in nature.

Finally, the grouping of Contributor/Author was used to group all references to people who played a role in the creation of the audio file, without distinguishing between the major and minor contributors. This included performers, recordists, composers, lyricists, transcribers, and uploaders. Some writers have pointed to a certain tension between traditional notions of authorship that AACR2 has supported, especially through its determination of main and added entries, and the distributed authorship that exists amongst a whole team of originators with digital content (Cwiok, 2005). This profusion of roles is common in the production of sound recordings as well. However, all of these contributors could be accounted for in an AACR2 description, in one way or another. So for the purposes of this study, the use of a general grouping for all authors and contributors was not seen as problematic.

## Other Challenges

Some challenges were encountered in working with the metadata elements. In some cases, a decision had to be made as to whether information about an object was metadata or whether it was dynamically generated contextual information, of the 'viewers who looked at this also looked at' variety. Usually, this kind of information was determined not to be metadata, and so was not considered an element for the purposes analysis. The presence of these kinds of features was recorded instead as a collection feature.

Another challenge resulted from the fact that some collections, such as DEKKMMA (Digitalisatie van het Etnomusicologisch Klankarchief van het Koninklijk Museum voor Midden-Afrika / Digitization of the Ethnomusicological Sound Archive of the Royal Museum for Central Africa), display all fields all the time, inserting null values as needed. This was taken as a sign that a database rather than digital library technology was being used. The empty fields were taken as an indication of a desired level of description, but for the purposes of this study, the null values were considered non-elements when calculating the mean number of elements. Likewise, repeating fields, such as subject fields, were counted only once, since in other cases, a string of subject headings might be contained within a single element field. For archival-style hierarchical records, where information about an item was contained in a linked pair parent/child records (usually a fonds/collection level record and an item level record), elements from both records were combined into a single record, and duplicate fields eliminated.

## Findings

The eighteen online audio-oriented digital collections included in this study are listed in *Appendix A*. They are listed with names of host institutions, URLs and acronyms that will be used throughout the discussion that follows. The findings are discussed in three sections that correspond to three phases of analysis:

- Collection profiles: hosts (*Table 1*), types of content (*Table 2*), collection features (*Table 3*)
- Metadata elements on a collection level (*Table 4*)
- Metadata elements on an aggregate level (*Table 5* and *Table 6*)

### Collection Hosts

Collection hosts were found to fall into five non-exclusive categories: university libraries, university departments or institutes, national libraries or archives, museums or archives, and consortia. The collections were often supported by several bodies working in collaboration, and four were supported by organizations that fell into more than one category. Consortia were distinguished from collaborations, and the category was used in cases where the material in the collection came from more than one institution. Less than half were hosted by university libraries, though two thirds were hosted by a university entity of some kind. National libraries, archives and museums hosted seven collections, and three consortia were found.

#### Types of Host Institutions

University Library	7
University Department or Institute	5
Museum or Archive	4
National Library or Archive	3
Consortium	3

Collections in Multiple Categories	4
------------------------------------	---

The Dismarc collection bears special mention, because it is the intended sound and music component of the Europeana joint digital library. Dismarc, which stands for Discovering Music Archives, holds collection and item level descriptions for over twenty European sound archives. In many cases, the collections were intimately connected to the activities of their hosts: some, like the Australian National Film and Sound Archive (NFSA) and the Vincent Voice Library (VVL), were an online presence for a long-standing collection, and others, like CHARM (AHRC Research Centre for the History and Analysis of Recorded Music), drew their material from off-line holdings.

The locations of the host institution was revealing. While almost all were based in North America, the United Kingdom and Europe, the content several of those collections had a specific geographic focus elsewhere. This likely reflects the fact that academics in Western countries often maintain research programs that study the cultures of other people and places.

#### Location of Host Institutions

Canada	1
United States	7
United Kingdom	3
Europe	5
South Africa	1
Australia	1

#### Geographic Focus of Content

Global*	7
The Americas	1
Canada*	1
US	2
UK	1
Europe	1
Middle East	1
Africa	3
Australia	1

\*CHARM, CPDP & VG were also considered to have a 'format focus'

#### Types of Content

The character of the content was similarly diverse. Music was the predominant content focus, but almost as many collections included the spoken word. Four were considered to have material scientific or environmental in nature, which included soundscapes, animals, or, in the case of SemArch (Semitisches Tonarchiv), the spoken word collected for linguistic analysis. It is interesting to note that three of the scientific collections accepted user submitted content, and this point of difference with the other collections could reflect a difference in purpose, wherein

the online collections are intended to consolidate current research data worldwide. This is certainly the case with SemArch and Xeno-Canto America (XCA).

Almost all the collections included field recordings, whether of music, the spoken word, or environmental sounds. The three that didn't – CHARM (AHRC Research Centre for the History and Analysis of Recorded Music), the Cylinder Preservation and Digitization Project (CPDP), and the Virtual Gramophone (VG) – were specifically focused on commercial 78s or cylinders. Interminglings of music, spoken word, field recordings and broadcast recordings were very common. This could be, in part, because the traditional approach to folklore research (Bartis, 1979) sees the spoken word and music as merely different manifestations of the culture of a people. This ecumenism is even reflected in the name of the American Folklife card catalog: Traditional Music and Spoken Word (TMSWC). Broadcast material was often found in collections with field recordings, which might be explained by the fact that these categories share the characteristic of being non-commercial recordings.

#### Recording Types

Music	13
Spoken Word	11
Scientific & Environmental	4
Multiple Categories	8

#### Recording Sources

78s & Cylinders	7
Field Recordings	15
Broadcast	7
User Submitted	3
Multiple Categories	7

#### Sound vs. Other Media

Sound Only	9
Sound and Other Media	9

The research data indicates that half of the collections contained additional media besides sound. Sometimes this meant video, as in the cases of Dismarc, NFSA, and Spoken Word Services (SWS), sometimes it meant transcriptions or field notes related to the recordings, as in the cases of the Henry Reed Collection (HRC), the James Koetting Ghana Field Recording Collection (JKC), and the Milman Parry Collection (MPC), and sometimes it meant that extended profiles on selected topics (e.g. performers or instruments) with text and photos were made available as value added content (DEKKMMA and VG). This admixture of media did help explain why the category of audio-oriented collections was difficult to define in the first place, since it couldn't always be defined as databases of sound recordings.

In some cases, the other media were given item level descriptions like the recordings (Dismarc, NFSA, and SWS), while in others the additional media was accessible through a separate search

path, or through a link from the recording they related to (DEKKMMA, MPC, and VG). It is likely that this trend towards audio-oriented collections filled with supporting material in other media will only continue, especially with the changing practices of constituencies like oral historians, who are doing more of their work with video (Sipe, 1991). In many cases, the additional media content supported the audio content by providing more context, especially in the case of field recordings. For instance, transcripts were available for some spoken word recordings.

## Collections Features

The presence of maps as a feature in the search and discovery interfaces was an especially interesting finding. While not found in overwhelming numbers, they were observed in recently constructed sites, like Archival Sound Recordings (ASR), and in sites that featured user submitted content, like the Freesound Project (FSP), SemArch, and XCA. Three of the five examples used embedded Google maps (ASR, FSP, and XCA), one used a clickable map of Africa as a search tool (DEKKMMA), while the fifth used maps for information only (SemArch), in a navigable browsing hierarchy of locations. The Google maps could be used to visualize the location of content or as a search mechanism. Sometimes a location indicated a precise longitude and latitude (FSP and XCA), but sometimes it indicated a general region of origin (ASR).

### Collection Features

Maps	5
Tagging & User Feedback	9
Metadata Schema	11
Controlled Vocabulary	5

The detailed investigation of interactive user features like commenting, tagging, designating 'favorites', and recommender systems was not part of the research question, but information on these aspects was recorded nonetheless to further contextualize the collections. The ability to create a list of favourites was the most commonly offered feature. Tagging and commenting are more directly related to metadata, but these features were only seen in a few cases (ASR, FSP, and XCA). FSP gives users the ability to tag, add multiple descriptions, and leave comments, which together form the core elements of FSP's descriptive metadata. Whereas in the case of ASR, tagging supplements, but does not replace the official metadata.

Only partial information about metadata schema and descriptive standards employed was collected, since the information was not always available. Nevertheless, data for some collections was found. The standards identified included MARC, MODS, METS, EAD and DC. Various controlled vocabularies were employed, such as LCSH, BBC keywords, and collection defined vocabularies (Dismarc). This is an area where further research is needed.

## Metadata Elements: Collection Level

Counting and comparing the number of metadata elements used in records from the various collections is a method that reveals some characteristics of those collections and allows comparisons to be made between collections. The average number of elements per record ranges from 9.6 for Dismarc samples, up to 24 for XCA samples, for a mean value of 14.77. The number of elements each set of five samples held in common ranges from 2 to 24, with a mean value of 11.

#### Summary of Collection Level Analysis

- *Column A: Average number of elements per record*
- *Column B: Number of elements common to all sample records*
- *Column C: Total number of unique elements found in all sample records*
- *Column D: Difference between column A and column B*
- *Column E: Difference between column C and column B*

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
Average Values For All Collections	14.77	11	19.06	3.77	4.29
<b>3 Examples</b>					
James Koetting Collection [JKC]	12	12	12	0	0
BL Archival Sound Recs. [ASR]	10.75	5	25	5.75	14.25
Spoken Word Services [SWS]	10.2	7	14	3.2	3.8

Some basic characteristics may be observed. First of all, the average number of elements per record (column A above) gives an indication of degree of description or cataloguing level of each collection, and in general, this is found to be beyond a merely basic level. This indicates that descriptive metadata is significant for the collections. In fact, the average number of elements per record across all collections is close in number to the fifteen elements in the Dublin Core metadata set. It's hard to know whether this says anything about Dublin Core, or whether it is just coincidence.

Second, the number of elements common to all the samples from a collection (column B), combined with the total number of unique elements seen in those samples (column C) can be compared to the average number of elements to give an indication of the consistency of application of a collection's metadata scheme, whatever it might be. In particular, the difference between columns B and C reveals a picture of variation from collection to collection in the consistency of application. On the one hand, collections like JKC were found to have a very consistent set of elements, while collections like ASR had a much lower number of common elements and a high number of unique observed elements – much higher than the average number of elements per record, in fact. This variation in consistency may be the result of inconsistent cataloguing, but, more likely, it is caused by heterogeneity in the content.

In fact, the collections where one finds the greatest difference between the total elements observed and the average number per record, or between the average number and the number held in common are ASR, Dismarc, NFSA and the Digital Library of Appalachia (DLA). These collections are either composed of many individual collections (Dismarc and DLA) or are processed by a single institution but display highly diverse content (ASR, NFSA). ASR has dozens

of separate collections, each consistent in character with itself, but quite different each to the other.

One would expect to find a similar situation on sites with user-submitted content like the Freesound Project, where the content is limited only by the imagination of its users, but, in fact, the deployment of elements in that collection remains quite consistent. This might be explained by the fact that, unlike archival holdings or important digitization projects, the recordings are being added progressively by users, one at a time, and their character is unknown by the system designers ahead of time. Therefore, there has been no attempt to accommodate specialized fields or radically heterogeneous sets of pre-existing metadata.

In a similar way, specialized collections that focus on a very particular type of material, such as one person's field recordings, can be quite consistent in their use of metadata elements. The ILAM (International Library of African Music) Digital Sound Archive, JKC, SemArch, and XCA collections all displayed no difference between the average number, the number in common and the total number observed. HRC is also close to this mark, considering the high number of elements it uses. Use of a database structure was observed to have some effect on consistency, in some cases. For instance, SemArch is apparently run with an Oracle database. However, singularity of purpose or common source of content were the more common explanations for element consistency. For instance, the content in HRC, ILAM, and JKC comes from a single ethnographer, in each case, and the XCA collection is very specific in its purpose of collecting bird songs. The nature of the birdsongs itself, lends itself to relatively straightforward (though potentially very specific) description, as there is no need to assign topical subjects to songs, or list performers, instrument names, and so forth.

In addition to more consistently used fields, specialized collections tended to use a higher number of elements than general collections. *Table 4* arranges the collections in ascending order of average number of elements per record. At the low end of the range (between 9.6 and 10.8 average elements), we find collections that contain commercial 78s and cylinders (ASR, CHARM, and CPDP), broadcast material (SWS), as well as aggregate style collections (ASR, Dismarc). In the middle section (12-16.2 elements), we find more focused collections, including three that have a consistent set of elements. The top end of the range consists primarily of specialized collections.

One reason that specialized collections had more elements, on average, per record, could be due to the fact that, in a collection of recordings all produced as part of a comprehensive research program (DEKKMMA), or by a single collector (HRC and MPC), the meaningful difference between records is found in the details. Collections of homogenous material also lend themselves well to comprehensive musicological analysis of forms (HRC) or scientific analysis of sound characteristics (DEKKMMA and XCA), which would likely increase the number of elements in use. It should be noted that NFSA and VVL make use of finding aid style parent/child records for collections and items in collections. As mentioned above, since the unit of study was the item, elements from the parent and child records were combined, which may mean that they had a greater number of elements than an equivalent bibliographic style

record. Also, the null fields in DEKKMMA, including a few that were not every used in the same records (e.g. voice type and instrument name) likely explains to some extent the high number of total fields, as well as the greater difference between the total number of fields and the average number of fields, and between the average and the number of common fields.

### Metadata Elements: Aggregate Level

After examining the characteristics of metadata on a collection level, the final step in the analysis was to derive a comprehensive view of trends and commonalities across the collections. While they were distinguished from each other by content type the collections were united in being 'oriented' towards audio, and the hope was that some indication of more specific characteristics they shared would emerge from a study of the most common metadata elements observed.

The principle behind choosing to include elements that had appeared in at least 75% of the sample records was that commonalities would be emphasized and general trends would be easy to see. A limitation of the method was that it did not allow nuances and details to be taken into account. But the number of samples taken from each database is small enough that the kinds of conclusions that are appropriate to draw from the results are general ones. In the discussion that follows, statements such as 'two thirds of collections had title metadata' should be taken to mean 'two thirds of collections had title metadata as a consistent element in the sample records examined.'

### Summary of Aggregate Level Analysis

- *This table summarizes Table 5, which summarizes the complete list of elements found in Table 6.*
- *Elements that appeared in 75% of the sample records from a given collection were included in the compiled list.*
- *Elements were organized into groupings through an intuitive affinity process.*
- *The groupings were given names and organized as sub-categories of seven metadata categories identified by Lagoze, Lynch, & Daniel (1996).*

<b>Metadata Categories</b>	<b>Groupings</b>	<b>Number of Metadata Elements in Groupings</b>	<b>Percentage of Total Metadata Elements</b>
1.0 Descriptive	8	149	66.80 %
2.0 Administrative	1	23	10.30 %
3.0 Terms & Conditions	1	4	1.80 %
4.0 Content Ratings	1	2	0.90 %
5.0 Provenance	2	15	6.70 %
6.0 Linkage	1	5	2.20 %
7.0 Structural	3	26	11.70 %

Of the 223 elements included in the comprehensive list, 149, or roughly two thirds, were considered to be descriptive metadata. Seven groupings included elements from at least two thirds of the collections: title (17 collections), author (16), description (16), identification numbers (14), keywords (13), date (13), and place (12). Of these, only identification numbers

was not considered a descriptive element (see the *Methodology* section for a discussion of this grouping). Also, it should be said that if all the groupings under the category structural (7.0) had been combined into a single grouping, it would have included 14 collections. If one adds the categories of administrative (2.0) (which contains the grouping identification numbers) and structural (7.0), they represent approximately 23% of the total number of elements.

One would expect elements like author and title to be common, and identification numbers were emphasized in the ARSC rules (ARSC, 1995), but it was somewhat surprising to find that date and place represented such a significant percentage of the total elements in category 1.0. There were almost twice as many place elements as there were collections that they came from, on account of the fact that 12 elements came from just 3 collections. Elements that recorded specialized musical details, such as key or range were not in widespread use, but the fact that such a high number (21) occurred consistently in four collections is interesting to note and probably due to a tight collection focus.

Other key findings include the fact that there were almost twice as many elements as there were collections in the author grouping, which confirms the intuitive understanding that recordings have multiple contributors. Almost all collections used a description or notes field. In fact, slightly more collections were represented in the description/notes grouping (1.3) than in the keywords grouping (1.4). When taken with the prevalence of date and place elements, this suggests that audio content tends to require explanation or contextual information to be meaningful.

#### Groupings Within the Category of Descriptive Metadata (1.0)

Grouping Name	Number of Collections Represented	Number of Elements	Percentage of All Elements in Category 1.0
1.1 Title	17	22	14.80 %
1.2 Contributor/Author	16	29	19.50 %
1.3 Description/Notes	16	18	12.10 %
1.4 Keywords	13	16	10.70 %
1.5 Date	13	16	10.70 %
1.6 Place	12	21	14.10 %
1.7 Language	5	6	4.00 %
1.8 Music Specific Details	4	21	14.10 %

## Discussion

What do the findings tell us about metadata requirements for audio-oriented digital libraries, collections and archives? The results found in this study suggest that metadata for these collections is not a case of one-size-fits-all. Collections might be created in the context of any number of fields of study, not just music, but they are likely to contain music or spoken word, and often together. They may be structured hierarchically like archival fonds, but are more

likely to be structured with an item level granularity. Many different disciplines create sound recordings for different purposes, and so the collections contained a diversity of broadcast recordings, field recordings, historic commercial recordings and archival recordings. The collections were also likely to have related materials in media other than sound.

### Heterogeneous and Homogeneous Content

Collections may bring together an array of recordings from disparate sources, or they may focus on a specialized topic that leads to the inclusion of other kinds of media and material besides sound. In the first case, there tends to be a low number of commonly recurring metadata elements from record to record, and in the second, there is a moderate to high number of elements, which do tend to recur from record to record. These factors suggest that the nature of the content plays a role in determining the metadata requirements, and moreover, the disciplinary context of content creation will determine the character of the content and the kind of pre-existing metadata that may already have been created. Heterogeneous content suggests the need for a flexible approach to metadata, while homogeneous content seems to demand a more customized, but therefore less nimble, approach to bring meaningful differences to the fore. In general, it was found that the number and consistency of metadata elements was related to the scope and consistency of content.

### Multiple Contributors

In terms of the descriptive elements, some general conclusions may be drawn from the set of most common elements. First, is the wide variety of roles that are involved in authoring a sound recording. A book too has many hands touch it as it is created, but editors and printing press operators are generally excluded from the formal descriptions entered in catalogues. In contrast, many of the roles in recordings are deemed relevant, whether or not they actually make any sound (viz. conductors and recordists). Audio metadata needs to accommodate this characteristic. At the same time, the proliferation of roles presents challenges for semantic interoperability (Cwiok, 2005), as adding qualifiers become more essential to make the contributor names meaningful.

### Time and Place

The prevalence of time and place metadata, coupled with the emerging trend of visualizing that data with dynamic maps speaks to another characteristic sound. Namely, that it often represents a specific time and place, which form an integral part of its meaning. The importance of the specificity appears to matter more, the more specialized the collection is. Time and place were found to be at least as important as descriptions and notes; there were 36 elements found for time and place and 18 for descriptions and notes. One could go further and argue that matrix and 'take' numbers that play such a crucial role in commercial recordings are mostly valuable because they tie a recording to a specific time and place, as well as to specific performers.

Because recordings present a particular moment, it seems that they require a lot of contextualization to bring out their full meaning. Description in sentence or paragraph form is also useful, but information of time and place, can be presented very precisely and concisely, and is therefore easier to structure as data. Once recorded in structured form, these elements become facets that offer great potential for innovative search and discovery mechanisms. Time and place as metadata are only useful, however, when the information was recorded in the first place. If the information is not there, it is impossible to recreate it. They also lose their usefulness in cases where recordings are heavily edited and come to represent a montage of different recordings. In general, though, this is an issue that has less bearing on field recordings and historical recordings, where the recording practice was very simple. The significance of time and place in a broader sense, is that they allow content to be located and integrated into world history, and this makes them universal, in a sense, because everyone can relate them to their own experience of time and knowledge of history, regardless of language or culture. Sound recordings do not have a monopoly on time and space by any means, yet recorded sound is 'about' time on many levels, and can't not be about time and place, it seems.

### Description and Notes

Finally, the consistent presence of a notes field in most of the collections suggests that sound recordings require summarization. In monographs and other text-based information packages, the title carries a lot of information and so receives a lot of attention in standard bibliographic descriptions. Recordings may relate to a musical work and therefore carry a title, but when they don't, even a constructed title may not suffice. One could argue that a constructed title is an abbreviated form of a note. The question is not so much aboutness as it is explanation. In this way, descriptive notes for recordings play a similar role to the scope and content notes used in descriptions of archival fonds.

### Interoperability

As mentioned above, there is a tension between the need for specific types of metadata and the desire for interoperability. As long as one is using a recognized schema, crosswalks are not a technological impossibility, but it is perhaps worth asking what type of material one wishes to be interoperable with. The Dismarc project represents a significant achievement in its integration of the catalogues of so many audio archives, but it was probably made easier by the fact that the collections were all audio-oriented. Even so, the commonality of elements was not high and consistency of application was not great. For example, in the sample records gathered from Dismarc, the information value related to the sound carrier 'CD' appeared variously in the extent, medium, and format fields. In the latter, it was defined by a controlled vocabulary. The records described holdings from five different archives. It would be useful to develop a commonly accepted theoretical understanding about audio-oriented metadata (and video-oriented metadata, for that matter), such that, at least audio-oriented collections could align

their approach to metadata with each other and improve the chances of successful interoperability. This would acknowledge the fact that sound archives will tend to want to share with other like-minded collections. In fact, this is exactly the principle behind the Dismarc project.

## Conclusion

What do the findings tell us about best practices for audio oriented digital libraries and online collections? This study examined metadata records, but did not evaluate whether that metadata was successful in supporting the desired functionality. Given that there was such a wide array of implementations, it is not possible to give a single model to emulate. However, some specific recommendations for audio-oriented metadata have emerged from this study. When developing metadata requirements for a pre-existing collection of content, in addition to evaluating user needs and modes of expected, the characteristics and consistency of the content should be assessed to determine the degree to which it is more or less homogeneous, as well as the extent of any pre-existing data that should be accommodated. Furthermore, in general, metadata schemas for audio-oriented content should accommodate:

- multiple contributor roles
- data about time and place of recording
- identification and issue numbers
- explanatory or descriptive notes

To some extent, these findings confirm basic understandings about sound recordings, previously identified in earlier cataloguing rules, such as the ARSC and IASA rules. The challenge of incorporating these recommendations into practice will be in striking a satisfactory balance between the conflicting imperatives of specificity and flexibility: specificity as required by the nature of the content and flexibility in achieving interoperability. Working within existing standards is recommended.

## Future Research

This study was exploratory in nature, and further research in the area of audio-oriented online collections would be welcome. Further work with the group of archives and collections in this study could conduct further sampling to improve the reliability of the results, or take a closer look at the metadata schemas in use. The pool of collections studied could be broadened to include additional collections subsequently identified by the author, such as the Densho digital archive<sup>2</sup> and the Maria Rogers Oral History Program digital archive<sup>3</sup>. As mentioned earlier, the measurement taken in this study was the metadata elements in use in the online collections, but the fact that particular elements are in use does not allow any conclusions to be drawn about their effectiveness in meeting the actual needs of users. With audio recordings in

---

<sup>2</sup> <http://www.densho.org/>

<sup>3</sup> <http://boulderlibrary.org/oralhistory/>

particular, there is a tendency, in some cases, to want to create highly detailed and descriptive records, without any sense of whether the added detail is meeting the needs of the users. The analysis of user tagging and descriptions, as noted in the Freesound project, would provide one way to evaluate user perspectives on metadata requirements.

## Acknowledgements

The author wishes to thank Dr. Ali Shiri for his guidance and support over the course of this research project. The author also wishes to thank Dr. Dinesh Rathi for his assistance in adapting the research paper for the web.

## References

- Baca, M. (Ed.). (1998). *Introduction to metadata: Pathways to digital information*. Los Angeles: Getty Information Institute.
- Bartis, P. (1979). *Folklife & Fieldwork: A layman's introduction to field techniques*. Washington, D.C.: American Folklife Center.
- Borgman, C. L. (2001). *From Gutenberg to the global information infrastructure: Access to the information in the networked world*. Cambridge, MA: MIT Press.
- Bradley, K. (Ed.). (2009). *Guidelines on the production and preservation of digital audio objects: IASA-TC04* (2nd ed.). Auckland Park, South Africa: International Association of Sound and Audiovisual Archives.
- A catalog of phonorecordings of music and oral data held by the Archives of Traditional Music*. (1975). Bloomington: Archives of Traditional Music, Indiana University.
- Cole, T. W., & Shreeves, S. L. (2004). Lessons learned from the Illinois OAI metadata harvesting project. In D. I. Hillmann & E. L. Westbrook (Eds.), *Metadata in practice* (pp. 174-190). Chicago: American Library Association.
- Corthaut, N., Govaerts, S., Verbert, K., & Duval, E. (2008). Connecting the dots: Music metadata generation, schemas and applications. *Proceedings of the 9th International Conference on Music Information Retrieval, Philadelphia, USA, September 14-18* (pp. 249-254).
- Cwiok, J. (2005). The Defining element: A discussion of the creator element Within metadata schemas. *Cataloging & Classification Quarterly*, 40(3), 103-133.
- Damnjanovic, I., Reiss, J., & Barry, D. (2008). Enabling access to sound archives through integration, enrichment and retrieval. *Proceedings of the 2008 IEEE International Conference on Multimedia and Expo, June 23-26 2008, Hannover, Germany* (pp. 1597-1598).

Greenberg, J. (2005). Understanding metadata and metadata schemes. *Cataloging & Classification Quarterly*, 40(3/4), 17-36.

Hartley, M. (2009, September 21). Download decade - Part 1: Thank you, Napster. *Globe and Mail* [online]. Retrieved March 29, 2010, from <http://www.theglobeandmail.com/news/technology/download-decade/thank-you-napster/article1014979/>

How to describe samples [online discussion thread]. (2005). Freesound Project. Accessed August 23, 2009 at <http://www.freesound.org/forum/viewtopic.php?t=36>

*IASA Cataloguing Rules*. (1999). Stockholm: International Association of Sound and Audiovisual Archives. Available at <http://www.iasa-web.org/content/iasa-cataloguing-rules>

Lagoze, C., Lynch, C. A., & Daniel, R. (1996). The Warwick Framework: A container architecture for aggregating sets of metadata. Available at <http://hdl.handle.net/1813/7248>

Lai, C., Fujinaga, I., Descheneau, D., Frishkopf, M., Riley, J., Hafner, J., et al. (2007). Metadata infrastructure for sound recordings. *Proceedings of the 8th International Conference on Music Information Retrieval, Vienna, Austria, September 23-27* (pp. 157-158). Vienna: Österreichische Computer Gesellschaft.

Lance, D. (Ed.). (1983). *Sound archives: A guide to their establishment and development*. Milton Keynes, England: International Association of Sound Archives.

Lesk, M. (2005). *Understanding digital libraries* (2nd ed.). San Francisco: Morgan Kaufmann.

Lourdi, I., & Papatheodorou, C. (2004). A metadata application profile for collection-level description of digital folklore resources. *Proceedings. 15th International Workshop on Database and Expert Systems Applications, 2004* (pp. 90-94).

Moutselakis, E. V., & Karakos, A. S. (2009). Semantic web multimedia metadata retrieval: A music approach. *PCI, 2009: 13th Panhellenic Conference on Informatics, 2009* (pp.43-47).

Notess, M., & Dunn, J. (2004). Variations2: improving music findability in a digital library through work-centric metadata. *JCDL: Fourth ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 422-422).

*Rules for archival cataloging of sound recordings* (Rev. ed.). (1995). Annapolis, MD: Association for Recorded Sound Collections.

Sipe, D. (1991). The future of oral history and moving images. *Oral History Review*, 19(1/2), 75-87. Reprinted in R. Perks, & A. Thomson (Eds.). (2006). *The oral history reader* (2nd ed.) (pp. 406-415). New York: Routledge.

Smith, A., Allen, D. R., & Allen, K. (2004). *Survey of the state of audio collections in academic libraries*. Washington, D.C.: Council on Library and Information Resources.

Smith, J. R., & Schirling, P. (2006). Metadata standards roundup. *IEEE MultiMedia*, 13(2), 84-88. [doi:10.1109/MMUL.2006.34](https://doi.org/10.1109/MMUL.2006.34)

*Sound archives: Guide to procedures*. (1979). Ottawa: Public Archives Canada, National Film Archives, Sound Archives Section.

Vinet, H., Herrera, P., Pachet, F. (2002). The CUIDADO project. *Third International Conference on Music Information Retrieval, Ircam - Centre Pompidou, Paris, France, October 13-17, 2002* (pp. 197-203). Paris: Ircam. ISMIR 2002, Paris.

Zeng, M. L., & Qin, J. (2008). *Metadata*. New York: Neal-Schuman Publishers, Inc.

## Appendix A: Audio-Oriented Digital Collections

The following collections were included in this study according to the selection criteria described in the *Methodology* section. An alphabetical table of collection name acronyms is provided below.

### **Archival Sound Recordings [ASR]**

British Library

<http://sounds.bl.uk>

### **Spoken Word Services [SWS]**

Glasgow Caledonian University

<http://www.spokenword.ac.uk>

### **CHARM (AHRC Research Centre for the History and Analysis of Recorded Music) [CHARM]**

Arts and Humanities Research Council

Royal Holloway University of London

King's College London

University of Sheffield

<http://www.charm.rhul.ac.uk/>

### **The Virtual Gramophone: Canadian Historical Sound Recordings [VG]**

Library and Archives Canada

<http://collectionscanada.ca/gramophone/index-e.html>

### **National Film & Sound Archive, Australia [NFSA]**

<http://www.screensound.gov.au>

### **Dismarc (Discovering Music Archives) [Dismarc]**

Europeana Digital Library (music component)

<http://www.dismarc.org>

### **DEKKMMA (Digitalisatie van het Etnomusicologisch Klankarchief van het Koninklijk Museum voor Midden-Afrika / Digitization of the Ethnomusicological Sound Archive of the Royal Museum for Central Africa) [DEKKMMA]**

Royal Museum for Central Africa (RMCA)

Université Libre de Bruxelles

Ghent University

<http://music.africamuseum.be/>

### **Milman Parry Collection [MPC]**

Harvard University

<http://chs119.harvard.edu/mpc/>

**Digital Library of Appalachia [DLA]**

Appalachian College Association Central Library  
<http://www.aca-dla.org/dlamusic/dlamusic.html>

**The James Koetting Ghana Field Recording Collection [JKC]**

Brown University  
<http://dl.lib.brown.edu/koetting/>

**Cylinder Preservation and Digitization Project [CPDP]**

University of California, Santa Barbara  
<http://cylinders.library.ucsb.edu/overview.php>

**Traditional Music and Spoken Word Catalog [TMSWC]**

American Folklife Centre  
<http://memory.loc.gov/diglib/ihas/html/afccards/afccards-home.html>

**Fiddle Tunes of the Old Frontier: The Henry Reed Collection [HRC]**

Library of Congress, American Memory  
<http://memory.loc.gov/ammem/collections/reed/>

**Xeno-Canto America: Bird sounds from the Americas [XCA]**

Xeno-Canto Foundation  
Naturalis: National Museum of Natural History (Holland)  
<http://www.xeno-canto.org/america>

**The Freesound Project [FSP]**

Music Technology Group (MTG) of the Universitat Pompeu Fabra in Barcelona)  
<http://www.freesound.org>

**ILAM (International Library of African Music) Digital Sound Archive [ILAM]**

International Library of African Music  
Rhodes University, Grahamstown, South Africa <http://ilam.ru.ac.za/>  
<http://greenstone.ilam.ru.ac.za>

**Vincent Voice Library (G. Robert Vincent Voice Library) [VVL]**

Michigan State University Libraries  
<http://vvl.lib.msu.edu/index.cfm>

**SemArch (Semitisches Tonarchiv) [SemArch]**

Ruprecht-Karls-Universität Heidelberg, Seminar für Sprachen und Kulturen des Vorderen Orients  
Semitistik  
<http://www.semarch.uni-hd.de/index.php4?&lang=en>

## Table of Collection Name Acronyms

ASR	Archival Sound Recordings
CPDP	Cylinder Preservation and Digitization Project
CHARM	CHARM (AHRC Research Centre for the History and Analysis of Recorded Music)
DEKKMMA	DEKKMMA (Digitalisatie van het Etnomusicologisch Klankarchief van het Koninklijk Museum voor Midden-Afrika / Digitization of the Ethnomusicological Sound Archive of the Royal Museum for Central Africa)
Dismarc	Dismarc (Discovering Music Archives)
DLA	Digital Library of Appalachia
FSP	The Freesound Project
HRC	Fiddle Tunes of the Old Frontier: The Henry Reed Collection
ILAM	ILAM (International Library of African Music) Digital Sound Archive
JKC	The James Koetting Ghana Field Recording Collection
MPC	Milman Parry Collection
NFSA	National Film & Sound Archive, Australia
SemArch	SemArch (Semitisches Tonarchiv)
SWS	Spoken Word Services
TMSWC	Traditional Music and Spoken Word Catalog
VG	The Virtual Gramophone: Canadian Historical Sound Recordings
VVL	Vincent Voice Library (G. Robert Vincent Voice Library)
XCA	Xeno-Canto America: Bird sounds from the Americas

## Appendix B: Data Tables

*N.B. An alphabetical table of collection name acronyms is provided in Appendix A.*

Table 1: Collection Hosts

Collection Name	University Library	University Department	National Library or Archive	Museum or Archive	Consortium
ASR			x		
SWS	x				
CHARM	x				x
VG			x		
NFSA			x		
Dismarc					x
DEKKMMA		x		x	
MPC	x				
DLA	x				x
JKC	x				
CPDP	x				
TMSWC				x	
HRC				x	
XCA		x		x	
FSP		x			
ILAM		x			
VVL	x				
SemArch		x			

Table 2: Content Types

Collection Name	Music	Spoken Word	Scientific	78s & Cylinders	Field Recordings	Broadcast	User Submitted
ASR	x	x	x	x	x	x	
SWS		x			x	x	
CHARM	x			x			
VG	x			x			
NFSA	x	x		x	x	x	
Dismarc	x	x		x	x	x	
DEKKMMA	x				x		
MPC		x			x		
DLA	x	x			x	x	
JKC	x				x		
CPDP	x	x		x			
TMSWC	x	x			x		
HRC	x				x		
XCA			x		x		x
FSP	x	x	x		x		x
ILAM	x			x	x	x	
VVL		x			x	x	
SemArch		x	x		x		x

Table 3: Collection Features

Collection Name	Contains Sound Only	Contains Sound & Other Media	Maps	User Feedback & Tagging	Schemas & Standards	Controlled Vocabulary
ASR	x		x	Tags, Favourites, Comments, Recommender	METS, DC compliant	
SWS		x		Linking, Citation		LCSH, BBC Keywords
CHARM	x				METS	
VG		x			AACR2, custom descriptive manual	LCSH
NFSA		x		Favourites		
Dismarc		x		Favourites	DC-Lib profile	LCSH, custom vocabulary
DEKKMMA		x	x			
MPC		x		Favourites	TED database, XML schema	
DLA		x		Favourites	DC	
JKC		x			Filemaker Database	
CPDP	x			RSS, Facebook Fans	AACR2, MARC	LCSH
TMSWC	x				MODS	
HRC		x				American Memory Genre Terms
XCA	x		x	React [i.e., comments]		
FSP	x		x	Recommender, Tagging, Description, Comments		
ILAM	x				DC?	
VVL	x				EAD	
SemArch	x		x			

Table 4: Observed Metadata Elements: Collection Level

- *Table rows are ordered by ascending values in column A*
- *Column A: Average number of elements per record*
- *Column B: Number of elements common to all sample records*
- *Column C: Total number of unique elements found in all sample records*
- *Column D: Difference between column A and column B*
- *Column E: Difference between column C and column B*

<b>Collection Name</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
Dismarc	9.6	2	21	7.6	11.4
SWS	10.2	7	14	3.2	3.8
CHARM	10.4	4	17	6.4	6.6
CPDP	10.6	9	13	1.6	2.4
ASR	10.75	5	25	5.75	14.25
FSP	10.8	10	12	0.8	1.2
JKC	12	12	12	0	0
SemArch	13	13	13	0	0
TMSWC	13.2	10	17	3.2	3.8
VVL	13.4	11	14	2.4	0.6
ILAM	14	14	14	0	0
DLA	15.5	8	23	7.5	7.5
NFSA	16.2	9	25	7.2	8.8
MPC	17.4	13	21	4.4	3.6
DEKKMMA	20.4	10	27	10.4	6.6
VG	20.4	16	24	4.4	3.6
HRC	24	21	27	3	3
XCA	24	24	24	0	0
<i>Average</i>	<i>14.77</i>	<i>11</i>	<i>19.06</i>	<i>3.77</i>	<i>4.29</i>

Table 5: Observed Metadata Elements: Aggregate Level (Summarized)

- *This table summarizes Table 6.*
- *Only elements that appeared in 75% of the sample records from a given collection were included in this list.*
- *Elements were organized into groupings through an intuitive affinity process.*
- *The groupings were given names and organized as sub-categories of seven metadata categories identified by Lagoze, Lynch, & Daniel (1996).*

Categories & Sub-Categories	Number of Observed Elements	Number of Collections Represented	Notes
1.0 Descriptive			
1.1 Title	22	17	
1.2 Contributor/Author	29	16	
1.3 Keywords	16	13	
1.4 Description/Notes	18	16	
1.5 Music Specific Details	21	4	1 element from JKC, with the remainder from DEKMMMA, HRC, & XCA.
1.6 Date	16	13	
1.7 Place	21	12	12 elements from DEKMMMA, SemArch & XCA. 1 element combines date & time.
1.8 Language	6	5	
2.0 Administrative			
2.1 Identification Numbers (e.g. Issue or Archival Reference Numbers)	23	14	
3.0 Terms & Conditions			
3.1 Legal/Rights	4	4	
4.0 Content Ratings			
4.1 Ratings and Download Statistics	2	1	
5.0 Provenance			
5.1 Publisher	8	3	5 elements from VG
5.2 Collection Source	7	7	
6.0 Linkage			
6.1 Relation	5	5	
7.0 Structural			
7.1 Format/Physical Description	15	9	
7.2 Duration	6	6	
7.3 Filesize	5	3	
<b>Total</b>	<b>223</b>		

Table 6: Observed Metadata Elements: Aggregate Level (Complete List)

- *This table is summarized by Table 5.*
- *Only elements that appeared in 75% of the sample records from a given collection were included in this list.*
- *Elements were organized into groupings through an intuitive affinity process.*
- *The groupings were given names and organized as sub-categories of seven metadata categories identified by Lagoze, Lynch, & Daniel (1996).*
- *A list of collection name acronyms is found in Appendix A.*
- *The right hand column gives the number of sample records from the collection that included the element.*
- *Where translations and elaborations have been added, they are shown in brackets.*

<b>Metadata Elements (categories &amp; sub-categories in bold)</b>	<b>Collection Name</b>	
<b>1.0 Descriptive</b>		
<b>1.1 Title</b>		
Title	TMSWC	5
Title	DLA	6
Title	ASR	8
Title/Work	CHARM	8
Title	CPDP	5
Title in Original Language	DEKKMMA	4
Title	Dismarc	5
[Sample Name]	FSP	5
[Title]	SWS	5
Title	ILAM	5
[title]	JKC	5
Title	MPC	5
Translated Title	MPC	5
Alternate Title(s)	HRC	4
Item Title	HRC	5
Title	NFSA	5
[Title]	SemArch	5
Unit Title	VVL	5
[Unit Title]	VVL	5
Title	VG	5
English [Bird Name]	XCA	5
Scientific [Bird Name]	XCA	5
<b>1.2 Contributor/Author</b>		
Recording Engineers	TMSWC	5
Primary Performer / Group	DLA	5
Recordists	ASR	6

Composer	CHARM	8
Performer	CHARM	8
Performer(s)	CPDP	5
Personal Name	CPDP	5
People	DEKKMMA	4
Performers	DEKKMMA	5
Collector	DEKKMMA	4
Contributor	Dismarc	5
File Added By [users] On [date]	FSP	5
Collector	SWS	5
Reporters/Presenters	SWS	4
Artist [Composer, Performer, Creator]	ILAM	5
Contributor	ILAM	5
Creators and Contributors	JKC	5
Transcriber	MPC	5
Name [of Singer]	MPC	5
Author/Creator	HRC	5
Performer	HRC	5
Collector	HRC	5
Credits	NFSA	4
Bearbeiter [Recordist]	SemArch	5
Sprecherin [Speaker]	SemArch	5
[Speaker]	VVL	4
Performer Heading	VG	5
Performer	VG	5
Recordist	XCA	5
<b>1.3 Keywords</b>		
Subject	TMSWC	5
Subject	DLA	6
DLA Category	DLA	5
Subject	CPDP	5
Tags	FSP	5
BBC keywords	SWS	4
Genre	ILAM	5
Keywords [Instrument Names]	ILAM	5
Genre(s)	JKC	5
Subject	HRC	5
Category [Subject]	NFSA	5
Produced as	NFSA	5
Thema [Subject]	SemArch	5
Subject [LOC]	VVL	5
Genre	VG	5
Type [of Call]	XCA	5

<b>1.4 Description/Notes</b>		
Description	ASR	8
Notes	TMSWC	5
Description	DLA	6
Note/Additional/Other Data	CHARM	6
Note	CPDP	5
Remarks	DEKKMMA	5
Description by [User]	FSP	5
Full Description	SWS	5
Description	ILAM	5
Abstract	JKC	5
Other Notes	JKC	5
Note	MPC	4
Notes	HRC	5
[Notes]	HRC	5
Summary	NFSA	4
Beschreibung des Tondokuments [Description]	SemArch	5
[Description]	VVL	5
Remarks by Recordist	XCA	5
<b>1.5 Music Specific Details</b>		
All Pitch Candidates	DEKKMMA	5
Tempo	DEKKMMA	4
Tessitura	DEKKMMA	5
Melodic Progress (Fragment)	DEKKMMA	4
Octave Reduced Tone Scale	DEKKMMA	5
type [Group or Solo Performance]	DEKKMMA	5
Instrumentation	JKC	5
Compass [Note Range]	HRC	5
Strains	HRC	5
Rendition	HRC	5
Meter	HRC	5
Phrase Structure	HRC	5
Key	HRC	5
Background	XCA	5
Variable	XCA	5
Volume	XCA	5
Sound Characteristics	XCA	5
Speed	XCA	5
Pitch	XCA	5
Number of Notes	XCA	5
Length	XCA	5

<b>1.6 Date</b>		
Date	CHARM	6
Date Recorded	TMSWC	5
Updated	TMSWC	5
Recording Date	ASR	7
Year of Release	CPDP	5
Date	DEKMMMA	5
Broadcast [Date]	SWS	5
Date	ILAM	5
Single Date	MPC	5
Recording Chronology	HRC	5
Broadcast Date	NFSA	4
Aufgenommen am [Recording Date]	SemArch	5
In Datenbank seit [Date Added to Database]	SemArch	5
Recorded [Date]	VG	5
Released [Date]	VG	5
Date	XCA	5
<b>1.7 Place</b>		
Place of Publication/Creation	TMSWC	5
Place	DLA	6
Recording Location	ASR	7
Country	DEKMMMA	5
Region	DEKMMMA	4
Village	DEKMMMA	4
Created/Published [Date/Place]	HRC	5
Coverage [Country]	ILAM	5
Place of Recording	JKC	5
Location	MPC	5
Country of Origin	NFSA	5
Land [Country]	SemArch	5
Großraum [Large Area]	SemArch	5
Ort [Place]	SemArch	5
Region [Region]	SemArch	5
Location	VG	5
Country	XCA	5
Location	XCA	5
Elevation	XCA	5
Latitude	XCA	5
Longitude	XCA	5
<b>1.8 Language</b>		
Language	HRC	5
Language	TMSWC	4

Original Language	DEKMMMA	5
Language	ILAM	5
Sprachtyp [Language Type]	SemArch	5
Dialekttyp [Dialect]	SemArch	5
<b>2.0 Administrative</b>		
<b>2.1 Identification Numbers</b>		
AFS Number	TMSWC	5
Identifier	DLA	5
Num [Matrix]	CHARM	6
Audio [Item Number]	CPDP	5
Issue Number/Label	CPDP	5
Tape Number	DEKMMMA	5
BBC Programme Number	SWS	5
Identifier	ILAM	5
Source [Matrix & Retrieval Numbers]	ILAM	5
Reel#	JKC	5
Item Number	MPC	5
Recording Number	MPC	5
Record Number	MPC	5
Call Number	HRC	5
Digital ID	HRC	5
Title No.	NFSA	5
Alt. VVL Number	VVL	5
VVL Call No.	VVL	5
Amicus No.	VG	5
Issue No.	VG	5
Matrix No.	VG	4
Take No.	VG	4
Cat. Nr.	XCA	5
<b>3.0 Terms &amp; Conditions</b>		
<b>3.1 Legal/Rights</b>		
Copyright Status	VVL	5
Rights	DLA	5
Legal	ASR	8
Rights	SWS	5
<b>4.0 Content Ratings</b>		
<b>4.1 Ratings and Download Statistics</b>		
Rating	FSP	5
Downloads [Number]	FSP	5
<b>5.0 Provenence</b>		

<b>5.1 Publisher</b>		
Source	JKC	5
Publication/Creation	JKC	5
Recording Source	VVL	5
Manufacturer	VG	5
Distributor	VG	5
Transcribed Label	VG	5
Generic Label	VG	5
Issue Type	VG	5
<b>5.2 Collection Source</b>		
Catalogue	CHARM	8
Repository	TMSWC	5
Holding Library	DLA	6
Original Item Location	CPDP	5
Host Collection	JKC	5
Collection	MPC	5
Publisher	ILAM	5
<b>6.0 Linkage</b>		
<b>6.1 Relation</b>		
Relation	DLA	6
Relation	ILAM	5
Available Text	MPC	5
Quantity [Related Files]	VVL	5
Discographical Reference	VG	5
<b>7.0 Structural</b>		
<b>7.1 Format/Physical Description</b>		
Form [i.e. Sound Recording]	TMSWC	5
Type	DLA	6
Format	DLA	6
Format [Source Format and Specification of All Audio Components Used For Transfer]	ILAM	5
Physical Description	JKC	5
Collection Method [i.e. Recorded]	MPC	5
Physical Medium	MPC	5
Object Type	HRC	5
Medium	HRC	5
Media	NFSA	5
Type [Type of Disc]	NFSA	5
Item Category [Access Copy]	NFSA	5
Format	NFSA	5
Formats	VVL	5

Physical Description	VG	5
<b>7.2 Duration</b>		
Duration	ASR	8
Duration	FSP	5
Duration	SWS	5
Duration	HRC	5
Running Time	VVL	5
Length [MP3 File]	XCA	5
<b>7.3 Filesize</b>		
File Size	DLA	5
Filesize	FSP	5
Type [File Format, Duration]	FSP	5
Bitrate of MP3	XCA	5
Sampling Rate	XCA	5