Digital libraries for music and voice

Polina Proutskova

Table of Contents

I. Digital libraries – a history

The beginnings of digital libraries are often traced back to J.C.R. Licklider's "Libraries of the Future" (Licklider, 1965). While his dream of the library becoming a computer has not come true, many of his ideas about information retrieval and user interfaces are still valid. Most importantly, in his book he imagined the world with a broad access to knowledge, as we know it today.

LexisNexis could be considered a predecessor of today's digital libraries: in the 1970s it pioneered the electronic accessibility of legal and journalistic documents (Berman, 2006). On 2nd of April, 1973, LEXIS launched publicly, offering full-text searching of all Ohio and New York cases. In 1980, LEXIS completed its hand-keyed electronic archive of all U.S. federal and state cases. The NEXIS service, added that same year, gave journalists a searchable database of news articles[1].

Another important predecessor of digital archives is the Oxford Text Archive (OTA) - an archive of electronic texts and other literary and language resources which have been created, collected and distributed for the purpose of research into literary and linguistic topics (Proud, 1991). The OTA was founded by Lou Burnard of Oxford University Computing Services in 1976, and is the oldest archive of academic textual resources in electronic form. From 1996 to 2008, the OTA was one of the centres of the Arts and Humanities Data Service (AHDS), and hosted AHDS Literature, Languages and Linguistics, a national centre for the support of digital research in literary and linguistic subject areas in the UK. AHDS Literature, Languages and Linguistics was funded to provide free advice to those from UK Higher Education in the best practice in digital resource creation. The OTA continues to accept deposits of primary-source academic electronic editions and

[1] See LexisNexis official homepage: http://www.lexisnexis.com/ and their Wikipedia page: http://en.wikipedia.org/wiki/LexisNexis

linguistic corpora, and is one of the key centres in the emerging European research infrastructure (CLARIN: Common Language Resources and Technology Infrastructure[2]).

Project Gutenberg was founded by Michael S. Hart, then a student of the University of Illinois, in 1971 (Gutenberg, 1971)[3]. Project Gutenberg is a volunteer effort to digitize and archive cultural works, to "encourage the creation and distribution of eBooks" (Hart, 2004). Hart was given unlimited access to a Xerox Sigma V computer, which was one of the 15 nodes on ARPANET, the computer network that would become internet. Hart's initial goal was to make the 10,000 most consulted books available to the public at little or no charge, and to do so by the end of the 20th century (Day, 2000). He began with a copy of the United States Declaration of Independence from his backpack, and this became the first Project Gutenberg e-text. Today the project is run by volunteers and is hosted by ibiblio at the University of North Carolina at Chapel Hill. As of November 2011, Project Gutenberg claimed over 38,000 items in its collection.

However, what we generally understand as a digital library first came into popularity in 1990s, when the internet became available to a large number of users and made instant, location independent access to information possible.

The years 1994-2004 are characterised by major programming funding from the US government. The National Digital Library Program was established that funded projects such as the Digital Library Initiative (DLI)[4] and its successor DLI-2[5]. Several organisations got together to push for an inclusive, international effort to pool resources, knowledge and experience to create digital libraries. This legitimized digital libraries as a subject for academic research and drew the attention of the public to it. Also The American Memory project at the Library of Congress was initiated, which today continues to provide free and open access through the Internet to written and spoken words, sound recordings, still and moving images, prints, maps, and sheet music that document the American experience[6]. At Carnegie Mellon University The Million Book Project was started in which american researchers collaborated with partners in India and China and scanned books in many languages[7]. The Internet Archive, a non-profit digital library with the stated mission of "the universal access to all knowledge"[8] was established in 1996[9]. It primarily archives the World Wide Web but also offers free access to a vast amounts of electronic materials from public domain.

While US government funding for DLI was discontinued in 2004, the European Commission became active with its i2010 programme in 2005-2009 which promised to digitize and make accessible millions of entities including various formats of cultural heritage[10]. i2010 was followed by Digital Agenda for Europe[11], currently running. While i2010 was concerned with digitization of analogue objects in the first place, the Digital Agenda is focussed on issues of licensing and interoperability to create a digital single market. For the first time digital libraries were present in a funding programme title in Frame Programme 7 (FP7) Information and Communication Technologies (ICT) Challenge 4: Digital Libraries and Content[12], which started in 2006 and is still running.

Google Books' predecessor projects Google Print and Google Book Search were launched in 2004,

---

2 CLARIN: http://www.clarin.eu/external/

3 Project Gutenberg official website: http://www.gutenberg.org/

4 DLI: http://dli.grainger.uiuc.edu/national.htm

5 DLI-2: http://www.nsf.gov/pubs/1998/nsf9863/nsf9863.htm

6 Library of Congress American Memory portal: http://memory.loc.gov/ammem/index.html

7 The Million Book Project: http://www.rr.cs.cmu.edu/mbdl.htm, its successor The Universal Digital Library: http://www.ulib.org/ . See also http://en.wikipedia.org/wiki/Million_Book_Project

8 See http://archive.org/details/SDForumBK

9 The Internet Archive: http://archive.org/

10 i2010: http://ec.europa.eu/information_society/eeurope/i2010/index_en.htm

11 Digital Agenda for Europe: http://ec.europa.eu/information_society/digital-agenda/index_en.htm

12 EU FP7 ICT Challenge 4: Digital Libraries and Content: http://cordis.europa.eu/fp7/ict/programme/challenge4_en.html

with the Google Books functionality as we know it today emerging by 2007[13]. The aim of the project is to digitize all published materials make them searchable online. Google's initiative has been hailed for its potential to offer unprecedented access to what may become the largest online corpus of human knowledge and promoting the democratization of knowledge, but it has also been criticized for potential copyright violations (Herwig, 2007). As of 2010, the number of scanned books was over 15 million. Google estimated in 2010 that there were about 130 million unique books in the world, and stated that it intended to scan all of them by the end of the decade[14].

Along with GoogleBooks and the Library of Congress, Europeana has become a major player among the world's institutional and enterprise digital libraries[15]. Europeana is a single access point to millions of books, paintings, films, museum objects and archival records that have been digitised throughout Europe[16]. Europeana currently provides access to about 10 million digital objects contributed by 20 000 European institutions. It is the EU Commission flagship project, heavily funded and publicised. Technically Europeana is an aggregated digital library: for each of its objects the central server only holds the metadata and the link to its original host(s).

Digital libraries are widely represented in academic research today. There are two peer-reviewed journals dedicated to digital library research: The International Journal on Digital Libraries (Springer)[17] and the Magazine of Digital Library Research (public domain)[18].

II Digital libraries of music and voice

Music is that something that everyone is familiar with but that turns out to be so elusive and intangible as soon as one tries to pin it down. Music is not defined by any of the artefacts that represent it, be it a score, a recording or a description. While organised sound can be captured by means of physical objects, its perception and interpretation in general cannot (though certain aspects can be encoded). Digital libraries of music are often multimodal, they include various types of objects and relationships between them. In that they are much more complex than digital libraries based on text or on any other single data type. The infrastructural demands for this kind of digital repositories present a large number of unsolved problems. Also, music is in most cases copyrighted material, and the legal aspect of digital music libraries plays a very important (and mostly restrictive) role. Audiovisual recordings of music which are part of digital libraries present a whole new series of challenges, related to capturing sound/video as well as long-term preservation of the recording and its migration to new formats and carriers in order for it to remain playable.

This section opens with an overview of the seminal guide on digital audio objects produced by the International Association of Sound- and Audiovisual Archives. Digitisation of analogue recordings as one way of acquiring digital content is outlined. Then music information research is introduced – an academic field searching for new methods to manage music collections and to extract information from them. The section closes with a summary of available digital music datasets.

IASA-TC04 Guidelines on the Production and Preservation of Digital Audio Objects

The International Association of Sound- and Audiovisual Archives (IASA) released its landmark publication called "Guidelines on the Production and Preservation of Digital Audio Objects" (IASA-TC04). It discusses all issues related to producing born-digital and digitising analogue recordings as well as their long-term preservation, access and collection interoperability.

---

13 Google Books: http://books.google.com/books/

14 See http://en.wikipedia.org/wiki/Google_Books and http://books.google.com/googlebooks/history.html for the history of GoogleBooks

15 Europeana: http://europeana.eu/

16 See http://pro.europeana.eu/about?utm_source=portalmenu&utm_medium=portal&utm_campaign=Portal%2Bmenu

17 http://www.dljournal.org/

18 http://www.dlib.org/

The key digital principles outlined in this document include:

1. Standards – using formats, resolutions, carriers and technology that adhere to international standards is crucial if mid-term or long-term preservation of recorded content is intended. Non-standard formats and technology may not in future be included in the preservation pathways that will enable long-term access and future format migration.
2. Sampling rate: IASA recommends a minimum sampling rate of 48 kHz for any audio recordings. This will ensure a lossless migration on future information carriers.
3. Bit depth: at least 24 bit depth is recommended for audio recordings, since it is the smallest bit depth encoding a dynamic range that approaches physical limits of listening.

It is essential to use a lossless format such as WAV, FLAC, AVI, AIFF for recordings. Any lossy formats like MP3, AAC, ATRAC (mini-disc format), MPEG are not appropriate.

IASA-TC04 understands metadata in a narrow sense of "structured data that provides intelligence in support of more efficient operations on resources" (p. 12). Schematic or encoded statements (also referred to as metadata instances) may be very simple, e.g. a simple URI within angle brackets; typically the become highly evolved and modular over en extended period of time.

There are two approaches of storing metadata recommended by IASA-TC04. The first approach is applicable for small collections. It involves using BWF files (Broadcast Wave format, EBU Tech 3285[19]) which are extensions of the widely spread WAV format and are supported by most recent audio technology. BWF incorporates metadata into the headers which are part of the file. In most basic exchange and archiving scenarios this is advantageous, because the danger to lose metadata or to mislabel a recording is eliminated. The Library of Congress has been working on formalising and expanding the BWF file format.

For larger institutional collections aimed at preservation, on the other hand, the disadvantages of this approach outweigh the advantages: the fixed nature of the embedded information may become a liability. Though editing metadata in a file is possible, it does not scale up as an appropriate approach for larger collections. The concept is then to always store the representation of one resource as two bundled files: one including the "contents" and the other including the metadata associated to that content.

In practice combinations of both approaches are common.

The file format recommended for externally stored metadata instances in general is XML. For time-based media, in particular for audiovisual objects, METS (Metadata Encoding and Transmission Standard) can be useful. METS includes components like descriptive metadata, the structural map (particularly useful for audiovisual streams), administrative metadata, technical metadata, digital provenance and rights metadata. Descriptive metadata is often encoded in accordance with Dublin Core metadata scheme, for compliance with the Open Archive Initiative (OAI-PMH). Dublin Core is a simple template of just 15 elements, which has been developed by the Dublin Core Metadata Initiative for cross-domain resources discovery on the internet. In some cases, for example when metadata is harvested from traditional library catalogue records based on MARC format, a more sophisticated metadata scheme such as MODS (Metadata Object Description Scheme) developed by Library of Congress is more appropriate.

The next stage of metadata design is an information model or an ontology. An ontology defines the terms used to describe and represent an area of knowledge. Ontologies are used by people, databases and applications that need to share domain information. Ontologies include computer-usable definitions of basic concepts in the domain and the relationships among them.[20] There are

---

[19] BWF user guide: http://www.ebu.ch/fr/technical/publications/userguides/bwf_user_guide.php , EBU 3285 specifications: http://tech.ebu.ch/docs/tech/tech3285s5.pdf

[20] Definition by W3C: http://www.w3.org/TR/webont-req/

existing ontologies in various domains, such as CIDOC CRM for the cultural heritage sector (museums, libraries and archives); FRBR (Functional Requirements for Bibliographic Records) will be appropriate mainly for recorded performances of musical or literary work; COA (Contextual Ontology Architecture will be fit for purpose if rights management is paramount as will be the MPEG-21 standard. RDF, a versatile and light-weight specification should be a component especially where web resources are being created. Existing thesauruses or controlled vocabulary lists can be used in place of complete ontologies. New computer-usable ontologies can be created using OWL (Web Ontology Language) using Protege – an open source tool from Stanford University.

Digitization

Digitization of analogue materials contributes an important part to the content creation of audiovisual digital libraries. Though a large proportion of new content is born digital[21], it is the digitized versions of older recordings that are often the most popular search queries. Today digitizing audiovisual holdings is considered the first priority for their long term preservation: multiplication and recovering strategies have been developed by IASA Technical Committee to guarantee the preservation of the content (Schüller, IASA-TC-03).

Notably, it has not always been evident that digitisation can be a form of preservation. In the 1970s-1980s expert associations such as Audio Engineering Society (AES) and Association for Recorded Sound Collections (ARSC) recommended re-recording of endangered materials on magnetic tapes. They were concerned about the use of digital recording technology and storage media because of rapid obsolescence of hardware, digital format and storage media. They also pointed out the questionable stability and durability of the digital media (Smith 1999).

Today digitization is in most cases the preferred form of long-term preservation. It has become obvious that preserving the carriers and maintaining the dedicated equipment is hopeless. At the end of the twentieth century the traditional "preserve the original" paradigm shifted via "preserve the content, not the carrier" to the "distribution is preservation" (Cohen, 2001) idea of digitising audiovisual content and making it available via digital libraries.

The philosophy of audio preservation and editing process has also been transformed at the end of the Twentieth century. The recent stance on decision making during the editing process has been influenced by the publications by Dietrich Schüller (Schüller, 1991) and is summarised in the UNESCO preservation guide (Boston, 1998): save history, not re-write it. Schüller appeals for defining procedures which guarantee the re-recording of the signal's best quality by limiting audio processing to the minimum. For more information on the audio preservation debate see (Orcalli, 2001).

Music Information Research – creating new infrastructure for digital libraries of music

A digital library is a more than a bunch of data, it is also defined by the infrastructure provided for the data. Someone has to be able to access, manage and share the data, and, on the next level, to extract new knowledge from it. When the data in question is music in all its forms and representations the task of providing such infrastructure becomes non-trivial an in some cases notoriously difficult. This is the reason why musical digital libraries are usually homogeneous and only equipped with some basic infrastructural elements today.

The area of academic research that is capable of changing this situation is called Music Information Research (MIR). It is concerned with extraction of information from music recordings or from

[21] Already in 2003 it was estimated that about 92% of new information was born digital, see (Lyman, 2003) at http://www.sims.berkeley.edu/how-much-info-2003

symbolic representations of musical pieces by means of (computer-aided) computation. MIR is a young discipline: the first official MIR conference took place in 2000. Today the International Society for Music Information Research[22] has hundreds of members around the world and the research community is thriving. In the days of iPods carrying thousands of songs and digital libraries like iTunes containing millions of tracks, the tasks of managing such collections became apparent and urgent: tracks retrieval by similarity of content, automatic genre classification and music recommendation have become the main research areas of MIR.

MIR faces a number of challenges: music can be represented in many forms, from an audio recording over a MIDI format to a score notation. Music is an interplay of physical (sound, image), psychological (listening) and cultural (remembering, interpreting) aspects and no aspect can be understood fully on its own. Thus MIR is a highly interdisciplinary field involving computer scientists, musicologists, information scientists and psychologists among others (Downie, 2003).

Common MIR problems include:

- music recommendation – suggest music tracks to a listener based on his preferences such as previously chosen tracks
- source separation and instrument recognition in audio recordings
- automatic music categorisation such as genre recognition
- music summarisation, automatic extraction of musical structure and of most relevant fragments
- query-by-humming (e.g. Shazam!)
- Optical score recognition, online score following
- Automatic transcription, automatic score creation

MIR community has developed an evaluation framework for algorithms which is comparable with the TREC framework in the field of text retrieval. The framework is called MIREX (Music Information Research Evaluation and eXchange)[23], the competitions have been running for seven years identifying the best performing new tools.

As opposed to text retrieval, MIR deals primarily with copyrighted material (music recordings, score editions, etc.). This creates significant problems in accessing the data necessary for research. Currently a new approach is being implemented and tested[24], where original music recordings are stored on a secure server and researchers can only access them through services like feature extraction (i.e. only retrieving information necessary for the particular research task and not complete recordings). This should provide protection against copyright violation accusations for researchers and open access to more copyrighted data (West, 2010).

MIR is primarily focused on Western popular and Western art music; it is only recently that non-Western musical repertoires have attracted attention of MIR scholars. The relatively sparse interest in non-Western music is due to the lack of data (recordings and annotations) available to MIR scholars and also to insufficient knowledge and understanding of other musical cultures (Downie, 2003; Tzanetakis, 2007). In recent years the interest to non-Western collections has grown considerably among MIR researchers and this is reflected in the increasing number of publications on subjects like melodic characterisation in Flamenco (Mora, 2010), automatic maqam or raga identification (Bozkurt, 2008; Sridhar, 2009), exploration of tuning and of Western influence in African music (Six, 2011). An ethnoMIR interest group and mailing list have been established and

---

[22] ISMIR homepage: http://www.ismir.net/

[23] MIREX homepage: http://www.music-ir.org/?q=node/13 , MIREX Wiki: http://www.music-ir.org/mirex/wiki/2011:Main_Page

[24] Networked Environment for Music Analysis: http://www.music-ir.org/?q=node/12

some significant projects in MIR for non-Western music are being funded[25].

MIR researchers have addressed many inherently difficult tasks in relation to infrastructure in music digital libraries. Tim Crawford et al suggested methods to recognise works of Western classical music in historical recordings (Crawford, 2010). By means of audio content matching incorrect metadata can be detected and corrected and duplicates can be identified and eliminated. This approach works even for poor quality historical recordings where other fingerprinting mechanisms fail.

A notoriously difficult task of automatic segmentation of audiovisual streams (such as ethnomusicological field recordings) into semantically separate fragments was addressed by Matija Marolt (Marolt, 2009). He experimented with field recordings from the Slovenian Academy of Sciences which he was able to segment into solo singing, group singing, speech, instrumental music and chime bells ringing.

Emilia Gomez et al. looked into automatic classification of musical recordings according to the culture of their origin, in particular distinguishing between Western and non-Western music (Gomez, 2009).

Available digital music collections

Because music recordings are copyrighted material with highly sensitive issues arising about any non-commercial use, finding datasets of audio recordings which are available for research is a challenge. Music Information Research (MIR) is an academic field that depends on such datasets and researchers in this field have put together the best resources. User data such as genre/mood tags and listening statistics are also of great importance for MIR.

Along with audio collections there are numerous symbolic music representation formats. Music is captured by means of Western score notation as well as by other forms of score notation. MIDI format provides much more precise information about sound events and musical instruments than the score notation. There are a number of music encoding languages that make music readable by a computer, the best known of them is Humdrum by David Huron. For all of these symbolic representations of music datasets are also available.

The Million Song Dataset[26] is the newest sensation in the MIR world. It offers audio descriptors, user and similarity tags, lyrics and more for over a million Western popular songs. The audio is not provided for copyright reasons, but it can be fetched from other services (the code for this is provided). This dataset for the first time offers an opportunity for a targeted and collaborative effort on scalability of MIR tools and strategies (Bertin-Mahieux, 2011). The LabROSA team that released the dataset partners with the highly successful internet radio last.fm[27] which has made its user statistics available for the 1 million tracks[28].

MIREX (The Music Information Retrieval Evaluation and eXchange) is a benchmarking competition of MIR tools and algorithms similar to the TREC contest in text retrieval. The aim of the competition is to name the best performing methods/tools in the field. The participants send in their code and it is run on the same dataset under the same conditions, thus providing comparability. The tasks include: music similarity for audio recordings as well as for symbolic representations of music (e.g. cover song detection), automatic genre and mood classification, melody extraction, key detection and chord estimation, beat tracking, real-time audio-to-score alignment, and many more (Downie, 2008).

---

[25] e.g. CompMusic: Computational Models for the Discovery of the World's Music at Music Technology Group Barcelona: http://compmusic.upf.edu/

[26] Million Song Dataset: http://labrosa.ee.columbia.edu/millionsong/

[27] http://www.last.fm/

[28] http://labrosa.ee.columbia.edu/millionsong/lastfm

MIREX started in 2011 and has been running annually. The datasets on which the competitors' code is run are often publicly available together with tags and labels for the given MIR task. These are available through the Wiki pages of the MIREX tasks, which can be accessed through the main Wiki page[29]. A summary of some of these collections can be found on the IMIRSEL website[30]

Georg Holzmann presents a comprehensive list of available datasets for MIR research (Holzmann, 2009)[31]. He lists all major audio repositories of popular music which are available under Creative Commons license including Magnatunes and USPOP2002. Above that his list also includes all kinds of symbolic music representation datasets, such as music scores, chords, themes, tags and metadata.

CHARM (Centre of the History and Analysis of Recorded Music)[32] was the first of its kind large musicological endeavour to study Western classical music through audio recordings as opposed to texts (scores) studies (Leech-Wilkinson, 2009). It involved creating and analysing large datasets of recordings, also by means of MIR methods. Some of the datasets are publicly available on CHARM's sound files website[33], among them a large collection of recordings of Schubert's songs from King's College in London. Altogether almost 5000 recordings are available.

Donald Byrd compiled a list of music datasets of primarily symbolic representations such as common score notation and MIDI (Byrd, 2006-2009).[34]

Of particular interest to AIRSPLACE choral projects will be the Choral Public Domain Library, a sheet music archive that focuses on choral and vocal music in public domain.[35] The site CPDL.org was launched in December 1998 by Rafael Ornes. Currently it is run by a group of volunteer administrators as a Wiki page.

Another important resource of sheet music is the International Music Score Library Project.[36] It is more musicology and research oriented and offers publications and recordings samples along with the scores.

Streaming services allow for listening of music recordings online though download is not permitted. The most popular streaming services are Spotify[37] and last.fm[38] for popular music and Naxos[39] for classical music.

Wikipedia offers a list of online music databases and on-demand streaming services.[40]

III Ethnomusicological archives and their electronic holdings

Ethnomusicological archives are the largest publicly available repositories of music recordings from non-Western cultures, holding hundreds of thousands of hours of original and commercially released ethnomusicological recordings. Most states maintain national archives concerned with preservation of local and national musics. Specifically designated ethnomusicological archives in the West are comprised of recordings from all around the world. Together these archives form the documented musical memory of mankind.

---

29 MIREX 2011 Wiki page: http://www.music-ir.org/mirex/wiki/2011:Main_Page#Tasks_with_a_2_September_2011_deadline:
30 IMIRSEL datasets: http://www.music-ir.org/?q=node/6
31 Georg Holzmann's list of MIR datasets: http://grh.mur.at/sites/default/files/mir_datasets_0.html
32 CHARM: http://www.charm.rhul.ac.uk/index.html
33 CHARM sound files: http://www.charm.rhul.ac.uk/sound/sound.html
34 Donald Byrd's Candidate Music IR Test Collections List: http://www.informatics.indiana.edu/donbyrd/MusicTestCollections.HTML
35 ChoralWiki: http://www3.cpdl.org/wiki/index.php/Main_Page
36 International Music Score Library Project: http://imslp.org/
37 Spotify: http://www.spotify.com/
38 Last-fm internet radio: http://www.last.fm/
39 Naxos music library: http://www.naxosmusiclibrary.com/
40 List of online music databases on Wikipedia: http://en.wikipedia.org/wiki/List_of_online_music_databases

Though most states commit to preserving their local and national cultures, very often archives holding ethnomusicological content suffer from insufficient budget, lack of professional staff and organisational dependencies. These archives are open to the public in most cases, but the access is hindered by lack of staff and the copyright laws. Online access to the holdings or the catalogues is still extremely rare.

In the recent years many ethnomusicological archives have experienced a rapid growth of their holdings due to the fact that audiovisual recording equipment has become affordable for many more researchers and non-professionals. Further growth of the repositories often combined with decreasing budgets is to expect. Unfortunately any infrastructure for automated processing of audiovisual materials, their cataloguing and making them accessible to the users is missing. In the absolute majority of scenarios every recording needs to be processed and catalogued manually by the staff, and the access is limited to the visitors physically present at the archive, who depend almost entirely on the advice by the archivist in search for their queries. Numerous digitisation initiatives have generated considerable digital repositories of earlier recordings, but since metadata and other additional textual information have not always been within the scope of these digitisation projects or haven't been paid necessary attention, the new digital repositories have not become more easily to navigate. The next order of complexity is added by the fact that there are no cataloguing standards for ethnomusicological holdings, the content of each archive is highly unique and metadata formats are the result of the archive's specific history combined with the archival practices of the enclosing institution (e.g. a national library or an ethnological museum), which are usually not designed for the use with ethnomusicological content. These inconsistencies and the lack of practically applicable and generally accepted ontologies make interoperability between ethnomusicological archives a huge challenge.

Ethnomusicological archives

Ethnomusicological archives build a systematic, often very well documented repository of music recordings, covering geographically all regions of the world and historically reaching in some cases as far back as the end of the 19th century. The core of their holdings are original audio and video recordings made by ethnomusicologists during their field trips. Many of the archives also collect published or commercial music recordings relevant for the subject. Trying to capture as much of the social context of music traditions as possible, ethnomusicological archives preserve also still images, field notes, books and publications on music cultures, music instruments, cultural and domestic objects brought from field trips (Proutskova, 2008).

and environments on all kinds of media

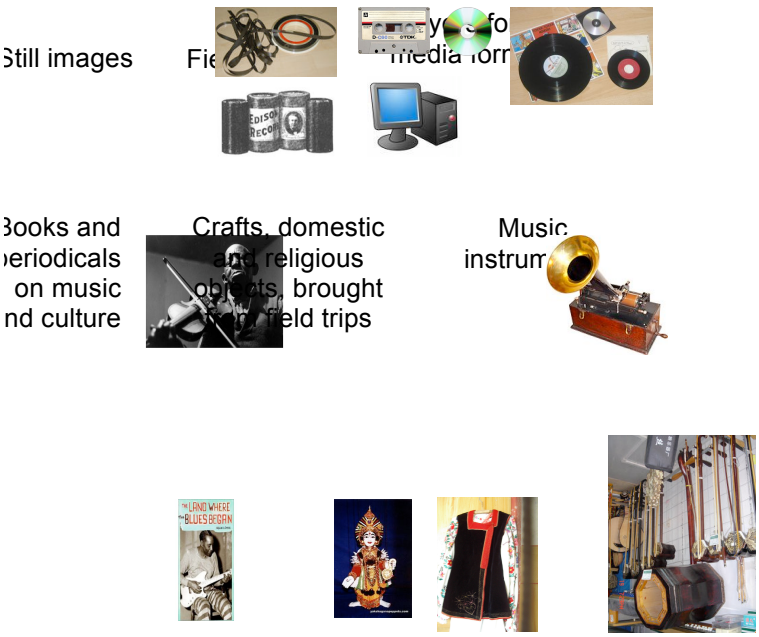


Figure 1: What ethnomusicological archives preserve[41]

The label "ethnomusicological archive" is a category describing the “cultural” content of an archive and not the type of data, in contrast to the category ‘sound- and audiovisual archives’, that is used, for example, in the title of the International Association of Sound- and Audiovisual Archives (IASA). Nevertheless the main part of their holdings is audiovisual data.

At the same time, because of their cultural content and the diversity of archived object types, they have very much in common with public institutions of the cultural heritage sector, such as museums, libraries and general archives.

There is a long standing tradition of documenting musical cultures by means of audio (nowadays also audio-visual) recordings. Western institutions hold historical recordings of key importance not only for European cultures but also musical traditions outside Europe. The first ethnomusicological audio recording was made in 1900 by Erich M. von Hornbostel, when Siamese (Thai) musicians were giving a performance in Berlin. The Berlin Phonogram Archive was founded in the same year in order to collect, preserve and study music recordings of world cultures. We can still listen to this first recording today, as well as to many other wax cylinder recordings preserved at the Berlin Phonogram Archive.

There are huge, often culturally diverse, repositories of ethnomusicological recordings residing at various Western institutions. The National Sound Archive (Traditional Music Section) in London hosts one of the largest ethnomusicological collections in the world, containing recordings from all around the globe, which is surpassed only by ethnomusicological holdings of the Library of Congress in Washington. The Berlin Phonogram Archive and the Phonogram Archive in Saint Petersburg (Russia) are also very large, historically grown, culturally spread collections. Many educational institutions hold collections of recordings made by their staff and students, which can be very significant, here Indiana University and UCLA have to be mentioned (Magas, 2012).

Many countries fund national sound/music archives, which are the main documenting bodies of local and national music cultures of the state (Polish National Sound Archive, Bulgarian Academy of Sciences Music Archive, etc.).

[41] Picture credits: Wikipedia

Fig. 1 shows the amount of ethnomusicological recordings in some of the largest archives of the world.

	Items	Hours (incl. commercial)	Original collections
Berlin Phonogram Archive	Over 150.000	18.000	Over 1.000
Library of Congress, Archive of Folk Culture		Over 100.000	Over 4.000
National Sound Archive of the British Library, World and Traditional Music section	300.000		370
Archives for Traditional Music, Indiana University	128.550	Over 250.000	2.000

Table 1: Amount of recordings in ethnomusicological archives (from Proutskova, 2007)

Anthony Seeger (Seeger, 1986, 1996) describes the significance of the archives for the field of ethnomusicology.

Retrieval strategies and user interfaces

The growth of digital ethnomusicological resources requires new scalable retrieval strategies as well as their intuitive graphical representations (Proutskova 2009). Ethnomusicological collections constitute a challenge with respect to both metadata and content management. The audio content of these collections is still a less studied terrain in Music Information Retrieval (Tzanetakis, 2007; Downie, 2003). Today most of the online ethnomusicological archives' catalogues offer bibliographic text searches that may sometimes include categories specific to music (e.g. the British Library Sound Archive Catalogue simple search[42]). Finding salient results is hindered by the incompatibility of controlled vocabularies between archives and metadata schemes. Locating recordings in an ethnomusicological archive remains mostly a domain of an experienced archivist, knowing their archive by heart. To enable users to search archives more effectively on their own, extensive, reliable annotation must be provided together with effective search tools and browsing interfaces (Proutskova 2008).

In ethnomusicological archives search and retrieval differ significantly from music archives containing Western repertories: there is usually no composer because many recordings are of traditional music; the artist is in most cases not known to the user; there is no such thing as album title and the song title may be just the lyrics incipit; genre and functional descriptions may exist, but they are most definitely not compatible between cultures. The first descriptor according to which an ethnomusicological collection is usually searched is the cultural origin of the music (Mengel, 2007), which often (but not always) coincides with the geographic location where the recording was made. It also may be the language, the name of the collector, a social function or context (like a specific ritual), etc. Often the searches are performed combining several criteria (Mengel, 2007).

Alan Lomax was the first to suggest a user interface for an ethnomusicological collection based on geographic location, using recordings from the Cantometrics dataset (Lomax, 1976). The Cantometrics project, started in the 1960s and pursued by Lomax until the 1990s, aimed at discovering correlations between singing styles and societal traits such as hierarchy structures and gender relations. For that purpose Lomax and his colleagues gathered and analysed around 7,000 recordings from more than 500 cultures (Lomax, 1976). In his Global Jukebox[43] clickable icons on

[42] British Library Sound Archive: http://cadensa.bl.uk/cgi-bin/webcat

[43] Lomax's dream of a Global Jukebox was realised by Association for Cultural Equity and was available on their website http://www.ace.org . Currently the name Global Jukebox has been given to the music label associated with ACE that releases Lomax's field recordings. The geo-archive from the original Global Jukebox concept is available

the world map trigger the audio content of the recording made at the respective location. A data similarity search is based on the 36 Cantometrics parameters, and relies on the manual annotation of each track by expert listeners. The results of this search can be shown on the geographical map or used to map out charts showing correlations between patterns of song and patterns of society.

Portals such as the National Geographic Nat Geo Music portal[44] or Mondomix[45] provide a relatively comprehensive overview of music from all over the globe, including audio and video searches, though they are limited to commercial releases. Their interfaces offer the user lists of genres, artists or regions to browse through, and thus rely on the user's knowledge of music-related metadata, rather than on the intuitive exploration of audio via audio-visual data representation.

Visualisations of commercial recordings rely mostly on clustering by genre or mapping of artist relationships (Gleich, 2005; Goussevskaia, 2008), which make them inappropriate for ethnomusicological recordings. With classification by genre, attempts to classify music often fail because of reasons like ambiguities, subjective judgment and marketing interests (Hilliges, 2006).

Musicovery (originally developed as Music Plasma or Live Plasma by Frederic Vavrille)[46] provides a colourful visualization of tracks according to genre and links between similar artists. It provides the option of selecting and listening to the genre of 'world' music, though the interface relies on standard music interface data (artist and name of track) leaving no room for ethnomusicological descriptions. It provides no information about the geographical or cultural origin of the tracks, or why particular tracks have been visually linked, and presents no opportunity for data management and annotation.

Collaborative database projects, such as The Freesound Project[47] which focuses on environmental sounds (Martınez, 2009), and Xeno-Canto[48] for bird song (Wellinga, 2009), rely on voluntary contributions of recordings, and use the Google Maps geotagging system as the basis for exploration of recordings from around the world. Clicking on a pin provides a link to a list of recordings related to the chosen location, which in turn provide links to audio samples. Recordings can be downloaded under the Creative Commons license and freely studied and explored. Tags and descriptions can be collaboratively added. Even though they are not designed specifically for ethnomusicological recordings, these collaborative audio-collecting systems offer greater scope for visualising and exploring ethnomusicological collections.

Geotagging is also used by the British Library Sound Archive online interactive maps which allow researchers to sample, download and study a variety of recordings with related texts and metadata.[49]

The above examples however do not focus on the feature extraction and audio similarity tools made available by the latest developments in MIR research.

Magas and Proutskova present a prototype for a dynamic user interface which offers an audio content-based search for ethnomusicological recordings (Magas, 2012). It makes use of the latest findings on audio similarity search in the field of MIR and adds a novel dynamic visual feature for location-tracking the similarity results, thus offering a carefully designed, intuitive gateway for exploring and searching a music collection of any cultural and geographical scope (Figure 2).

---

at http://www.culturalequity.org/lomaxgeo/ . The Cantometrics part of the Global Jukebox isn't currently available.

[44] National Geographic world music portal: http://worldmusic.nationalgeographic.com/

[45] Mondomix: http://www.mondomix.com/

[46] Musicovery: http://www.musicovery.com/

[47] http://www.freesound.org/

[48] http://www.xeno-canto.org/

[49] British Library Sound Archive: http://sounds.bl.uk/Maps.aspx

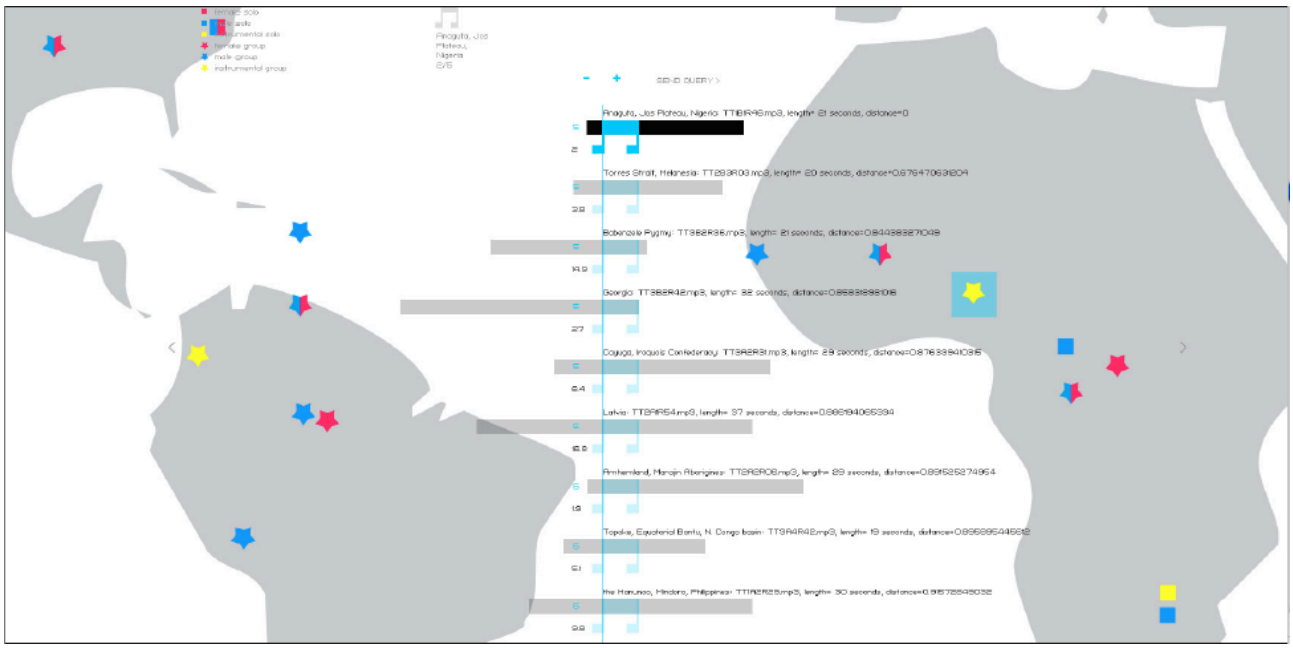


Figure 2: Fragment A location-tracking interface for ethnomusicological recordings. The square location tracker is visible on the right in Nigeria, as the site of the Anaguta, which have been chosen as the initial query for the audio similarity results displayed in the centre. The highlighted query results are displayed on bars corresponding to the relative lengths of tracks, as segments relative to the query length, and are aligned to the query point. Clicking on a query result will play the corresponding audio and launch the location tracker on a trajectory to the corresponding location on the map. Artwork copyright 2009 Michela Magas.

The work of annotating ethnomusicological recordings is very time-consuming and requires deep knowledge of the recorded material and its cultural context. Practically all ethnomusicological archives possess recordings where annotation is missing. Some archives report that less than 10% of their holdings are annotated appropriately (Proutskova, 2008).

This is particularly true of their electronic holdings: very often digitization projects have been focussing on digitising audiovisual content in the first place, paying insufficient attention to annotations. This has sometimes led to omitting metadata from the digitisation process or to missing links between digital audiovisual artefacts and its editorial metadata.

One way to deal with the shortcomings of annotation could be the principle of social tagging: an online editorial system allowing registered users to add, edit and negotiate information on catalogue entries, which then becomes available to all users. Naturally people interested in particular recordings might be better informed about their cultural context, than the archivists are. The process can be supervised by an archivist or, if the number of participants is high, it might become self-regulated, as the experience of Wikipedia[50] shows. In any case, it must be clearly sited, who and when added a particular information and what his sources are (Proutskova, 2008).

Since editorial metadata like cultural origin and social context is essential for ethnomusicological recordings, retrieval strategies combining both metadata and audio content search promise the best results. Such retrieval strategies would introduce entirely new search and research scenarios. They will enable data-driven research on music, relying on archives' holdings as corpora (Proutskova, 2008; Magas, 2012).

Access to data and its political aspects

In present-day ethnomusicological archives, the main obstacles causing restrictions in access to

[50] www.wikipedia.org

recordings after they have been successfully located in the catalogue is the lack of manpower and of an effective infrastructure facilitating fast processing of users' requests. Still there are other obstacles of political nature, specific to ethnomusicological archives (Proutskova, 2008).

1. A possessive attitude of some collectors and archivists – it is sometimes hard for a field researcher, who went through all possible difficulties to obtain a recording, to give it generously to others without expecting even a "Thank you!".

2. A state or administrative policy – if for example an archive holds recordings of local music traditions that are considered as an important part of the national identity, the state sometimes may put restrictions on the use of its content. If the archive is part of a large institution, general restrictive access policies of the institution are likely to affect the work of the archive.

3. Yet the main reason to restrict the access to recordings is a complicated proprietary rights situation (Seeger, 1996, 2004). The rights status of many of the recordings in ethnomusicological archives is not clear. On one hand ethnomusicological recordings are part of cultural heritage, often produced and maintained with public money, which implies that they should be accessible for at least the common needs of society like research. Ethical considerations make it necessary to provide access to the communities that were recorded. On the other hand we are talking about music recordings, for which the proprietary rights of the producers are strongly guarded in the West. The songs are often traditional, thus there is no composer, the performers living far away from the Western law space. The right to play and reproduce the recording usually remains by the collector. Unfortunately, field researchers are often reluctant to make their recordings accessible in general, without them being asked for permission for each use. In many cases the ownership is unknown (orphan works) or the owner cannot be located, which also restrains archives from making recordings accessible. These complications lead to the sad situation, that the recordings can neither be used for commercial goals nor in an educational context or for research purposes.

Existing resources

Though many archives catalogue their holdings electronically today, the majority of the catalogues are still not online, the most prominent example being Berlin Phonogram Archive[51]. Yet there have been some promising developments in the last years.

The WebFolk Project of the Bulgarian Academy of Sciences seems to have been the first online ethnomusicological catalogue, launched before 1997[52].

Some bigger archives' catalogues are now available within the online catalogues of the enclosing libraries: The UK National Sound Archive, which is part of the British Library,[53] and the Archives for Traditional Music at Indiana University[54]. Recently the Archive of Folk Culture of the Library of Congress has published its online catalogue on the library's website[55].

Only few archives that hold distribution rights for their collections are bold enough to offer online access to their recordings. One model is set by Smithsonian Global Sound portal. It offers all the songs from the Folkways Recordings collection for download for 99 cent – a kind of iTunes model for traditional musics[56], with the major part of the benefits going to artists recorded. The Global Music Archive at the Vanderbilt University's Blair School of Music offers streamed audio for listening[57].

---

51 It is planned to make the electronic catalogue of the Berliner Phonogram Archive available online within internet portals of the EU projects DISMARC and EthnoArc

52 http://arts.bas.bg/EN/Default.htm

53 http://cadensa.bl.uk/cgi-bin/webcat

54 http://www.iucat.iu.edu/authenticate.cgi?status=remote&select1=WEBSERVER

55 http://lcweb2.loc.gov/diglib/ihas/html/afccards/afccards-home.html

56 Smithsonian Global Sound: http://www.smithsonianglobalsound.org/

57 Global Music Archive: http://www.globalmusicarchive.org/

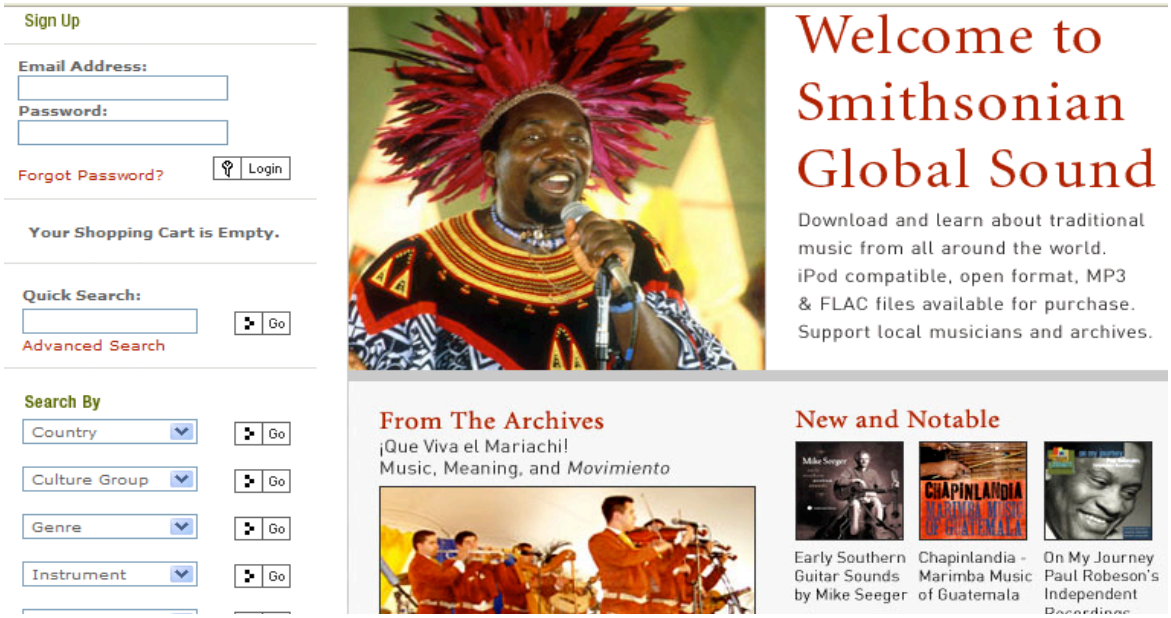


Figure 3: Fragment of Smithsonian Global Sound homepage

EVIADA stands for "Ethnographic Video for Instruction and Analysis Digital Archive".[58] Funded by Mellon Foundation between 2000 and 2009, the EVIADA project team developed a new generation of archival framework which combines detailed annotation of collections, their long-term preservation, peer-reviewed online publication and provides access to the collections specifically for use in the classroom. About 40 collectors from various countries were invited to contribute their collections of video field recordings to the EVIADA archive. The condition was that the researcher provides complete annotation of their collection. A small number of collections have been electronically released on the EVIADA website, other collections are currently under peer-review.

Association for Cultural Equity[59] offers access to Alan Lomax's geo-archive[60] with field recordings from all over the world. A wealth of materials about Alan Lomax, his research and his numerous field trips is available, including recordings of his lectures and recording sessions (not just recorded songs).

Interoperability and metadata standardization

Since the founding of the first sound archive in Vienna in 1899 in many countries of this world ethnomusicological archives have been established, typically using different documentation strategies, focusing on different aspects of the recordings. There are no widespread standards. Metadata structures often grew historically, reflecting documenting policies and types of media a specific archive once used for cataloguing. An ethnomusicological archive is often a part of a bigger institution and is obliged to use the standards of the institution, which are not always well adapted to non-Western musics.

The diversity of archives' histories, types of archived objects and metadata structures makes networking among archives and data integration a great challenge.

Existing metadata standards and best practices in related fields, like sound archives, cultural heritage sector, digital archives, etc. may suggest practical solutions. Some of the existing standards provide well organised, abstracted high level structure (ISAD(G), SIDOC Concept Reference

---

58 EVIADA: http://www.eviada.org/ . To be granted access to published collections, please contact the EVIA team.
59 Association for Cultural Equity: http://www.culturalequity.org/
60 Alan Lomax's geo archive: http://www.culturalequity.org/lomaxgeo/

Model, Dublin Core, EAD). Others suggest encoding schemes and vocabularies (ICAT for sound documents, MARC for books and published materials). Unfortunately none of them can sufficiently describe all data and object types in ethnomusicological archives, providing both the main focus on original recordings and controlled vocabularies for musical and ethnological content.

There is no consensus among ethnomusicological archives about cultural groups, nor geographic locations, nor ethnographic or anthropologic taxonomy for cultural context and objects. For musical instruments Hornbostel-Sachs classification is sometimes used, but since it is almost 100 years old, it has often been modified for the needs of a particular archive. For languages, Ethnologue[61][Gordon 2005] is a frequent choice, but many institutions have their own tables and labels for languages.

The need for exchange and integration of metadata from different ethnomusicological collections makes a certain amount of metadata standardisation inevitable. Straightforward solutions like flat data exchange or forcing participating archives to use the same metadata structure may be useful at some stage. There is a tool developed by an EU project ethnoArc, which helps to browse different archives without reducing their metadata structures to a single common metadata scheme[62].

But these solutions come very soon to their limits. A sustainable, flexible and scalable solution is sought for a semantically rich integration of larger networks.

A common ontology for recordings of non-Western musics would be a way to overcome these obstacles. It would allow deep catalogue metasearch and browsing or data exchange and linking between the archives using it, even if they apply different metadata schemes. The Ethnographic Thesaurus project might have the potential to suggest a basis for a common ontology in the near future[63]. If standards for digital archiving like Dublin Core and OAI are supported, an integration of the catalogues and associated digital sources into the Semantic Web will be possible, making its resources available for ethnomusicological recordings (Proutskova, 2008).

VI. Best practices

Decisions to be made prior to establishing a digital music library:

1. Naming and naming spaces for resources and metadata, e.g. persistent URIs;
2. Whether metadata is stored with the content or externally, or a combination of both;
3. Metadata scheme to be applied to best meet the needs of the production team, the repository itself, and the users, e.g. Dublin Core, METS;
4. The graunlarity of metadata, a trade-off between developing and managing costs and creating sufficient descriptions that will serve future demands;
5. meeting interoperability demands, e.g. among institutions, or OAI-PMH compliance for general web resources;
6. creator and user rights management;
7. metadata harvesting;
8. repository software
9. A lossless file format for preservation of original resources, e.g. TIFF for images or FLAC for audio files;
10. technical specifications such as resolution for images and sampling rate and bit depth for audio;

---

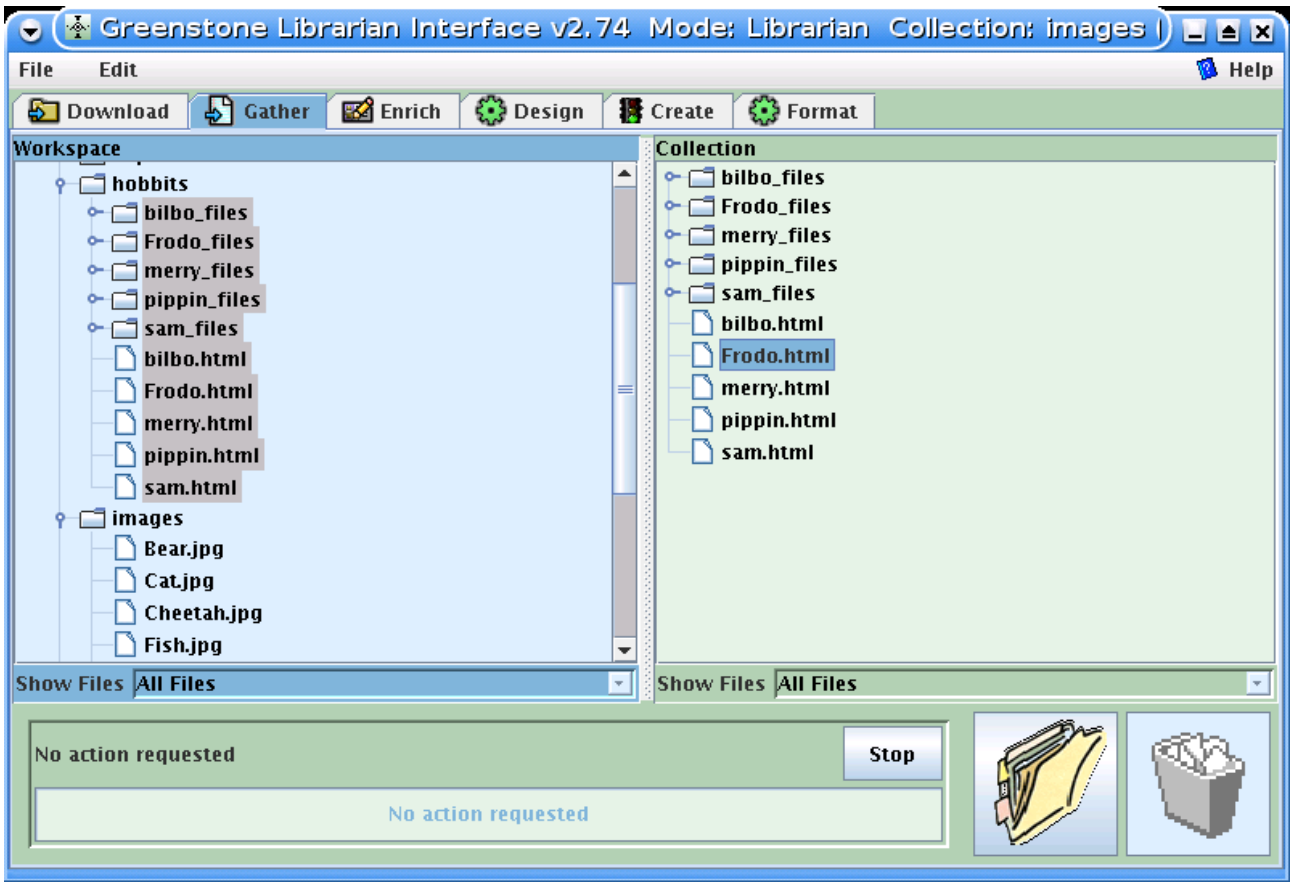
61 http://www.ethnologue.com/
62 http://www.ethnoarc.org/
63 http://www.afsnet.org/thesaurus/

11. Delivery formats (e.g. MP3 for sound or MPEG for video) and the required network capacity;
12. Term of preservation and future migration strategies;
13. Physical data carriers (avoid CDs and DVDs for preservation);
14. Backup;
15. Authenticity checks (fingerprinting or checksums).

For more guidance and information consult IASA-TC04.

Repository software

Greenstone was one of the earliest public software suites addressing the needs of institutional digital libraries of all sizes (Witten, 2001)[64]. It was developed by the New Zealand Digital Library Project at the University of Waikato. Development and distribution was in cooperation with UNESCO and the Human Info NGO in Belgium since 2000. It is an open source, multilingual software enabling institutions to build digital libraries and to distribute them via internet or on CD-ROM. There is a graphical user interface called GLI (Greenstone Librarian Interface) for building collections and assigning metadata; for the endusers (readers) of the collections there is a web interface. Greenstone is probably the best tested and supported environment with a large user and developer community which has stood the test of time. It can handle all types of text documents in a great variety of formats; many audio formats, MPEG (moving image) and MIDI can be imported via a plugin. There is also a plugin for the most popular image formats.



Greenstone supports metadata import and export in a number of important formats such as Dublin

[64] Greenstone home page: http://www.greenstone.org/

Core (qualified and unqualified), XML. MARC, OAI, METS. New metadata sets can be defined using Greenstone's Metadata Set Editor.

"How to build a digital library" by Witten, Bainbridge and Nichols (Witten, 2009) is a book-length introduction into building and maintaining digital libraries with Greenstone. It contains a chapter on ingesting and curating multimedia content. The book was written with both the technician and the librarian in mind.

DSpace is a US based alternative to Greenstone (Smith, 2003)[65]. It also has a very large community of users and developers and is being used by a large number of educational, government and private institutions. It can be installed out of the box and customised to fit the needs of an individual collection. DSpace was first released in November 2002, following a joint effort by developers from MIT and HP Labs in Cambridge, Massachusetts.

The main strength of DSpace is at the same time its weakness: its integrated feature set allows institutional users to quickly establish a repository and start adding new items to the collection; because of that DSpace has evolved into a monolithic software application and a complex code base, that introduces potential scaling and capacity constrains for some large institutional users. This presents no problem for most small to medium scale digital libraries and should not be an issue for any digital audio collection (IASA-TC04).

Witten at el (2005) analyse the differences between the two systems. They mention the following points[66]:

- Preservation. DSpace is explicitly oriented towards long-term preservation, while Greenstone is not (Tansley, 2003).
- Support infrastructure. DSpace is designed for institutional use, where there are centralized computing facilities and a competent infrastructure for software support. Greenstone is designed to be easy for anyone with basic computer-literacy skills to install, in a laptop, desktop, or institutional environment.
- Author-oriented vs. librarian-oriented. DSpace incorporates an interface whereby users (typically authors, though some institutions choose to have librarians do this on behalf of the faculty) can submit documents to the system, and define metadata for them. Greenstone does not. Greenstone supplies an end-user interface with which collections can be designed, customized, and built (by librarians). DSpace provides a generic design that can be tailored – but not by typical end users.
- Built-in metadata standard. DSpace imposes a single metadata standard on all collections. Greenstone provides a widely used standard (Dublin Core) but also allows collection-builders to use their own metadata scheme.
- Distribution on removable media. Those who create Greenstone collections can write them to a self-installing CD-ROM that operates on all Windows systems (even obsolete ones right down to Windows 3.1/3.11, still in use in developing countries).
- International users. Greenstone provides interfaces for readers in 35 languages, including many minority ones, and has a scheme that helps language maintainers keep the interfaces up to date when new interface features are added (Bainbridge, 2003).

Further Witten et al (2005) describe an implementation of a migration software StoneD between the two systems, which ultimately allows to use both of them simultaneously.

A third alternative to Greenstone and DSpace is FEDORA (Flexible Extensible Digital Object and

---

[65] DSpace home page: http://www.dspace.org/

[66] A blog website bringing a more recent comparison of both systems: http://rajabiswas.wordpress.com/2011/07/02/lib-automation/

Repository Architecture)[67]. It is designed as a base software architecture upon which a wide range of repository services can be built. It originated from a research project at Cornell University in 1997, and the first article was published by its creators in 1998 (Payette, 1998). FEDORA has been slower to gain adopters because it lacks a dedicated user interface and access services out of the box. There are now a number of commercial and open source providers of web-based front-ends for FEDORA.

The main strength of FEDORA are its flexible and scalable architecture. FEDORA can scale to cope with large collections, yet it is sufficiently flexible to store multiply types of digital objects and their complex relationships. There are few limitations to the features that can be added to a FEDORA repository, and it still remains interoperable with other applications and systems. It can be configured to support virtually any metadata profile.

The main disadvantage of FEDORA is the high level of software engineering expertise required to contribute to its core development, and it is not readily installed and implemented "out-of-the-box" (Bradley, 2007).

Tools have been developed to migrate content from DSpace to FEDORA and vice versa, which theoretically negates any future compatibility issues and supports sharing.[68] There have even been considerations of a closer integration: for all the content and metadata to be stored in Fedora with DSpace continuing to provide the user interface for workflow, discovery and administration (retain the out-of-the-box experience of DSpace while exposing the versioning, object relationship, and flexible architecture features provided by Fedora).[69]

IASA recommendations for small-scale digital music libraries

For a small-scale audio libraries IASA-TC04 recommends the use of BWF (Broadcast Wave Format) audio format which is a container for a WAV file that can include metadata for the audio.

The minimal metadata set should include:

- Unique Identifier, which should be structured, meaningful and human-readable as well as unique
- a description of the source, a small amount of free text to identify the content of the audio file
- technical data: format, sampling rate, bit rate, file size; this information can be harvested automatically in most cases; though it can be acquired later, making it an explicit part of the record allows management and preservation planning of the collection
- coding history, or, for born-digital object, production history, describing the process and technology of creating the digital file
- process error messages collected during creation or transfer

If audiovisual objects are supposed to remain accessible over a medium or longer term, planning and preparation needs to be done for the time point when the current computing storage and access environment becomes obsolete. The choice of BWF or any other standard format ensures the longest time possible before any format migration is necessary.

IASA also lays out basic principles of archival storage:

- There should be multiple copies. The system should support a number of duplicate copies of the same item.

---

[67] FEDORA homepage: http://fedora-commons.org/
[68] Repository Interoperability Framework: http://www.apsr.edu.au/currentprojects/index.htm#riff
[69] See http://dltj.org/article/or11-report-3/

- Copies should be remote from the main or original system and from each other. The greater the physical distance between the copies the safer in the event of disaster.
- There should be copies on different types of media. If all the copies are on a single type of carrier, such as hard disc, the risk of a single failure mechanism destroying all the copies is great. IT professionals commonly use data tape as the second (and subsequent) copy.

No single component of a digital system can be considered reliable, instead reliability is achieved through multiple redundant copies at every stage. A system which contains all of the data on a hard disc array such as RAID, all of which is also stored on a number of tapes (e.g. LTO) is recommended. A minimum of two sets of data tapes must be produced, to be stored physically in different places. This solution will minimise the disaster risk and will also enable the repository manager to implement a low-cost error checking – see IASA-TC04 Chapter 7 for more details.

Annotations of time-related media

Annotation of the content is one of the most common and most labour-intensive tasks associated with creating or enriching audiovisual digital libraries. Annotations are usually stored extern to the recordings and in this instance they constitute metadata. On the other hand, annotations can be considered part of the content, since they carry essential information about the recordings without which the content is not complete (Agosti, 2007). For example, if a singing lesson was recorded, and the teacher made notes during or after the recording, these notes are not part of the recordings; at the same time, they are as much part of the lesson as anything else that was recorded.

Audiovisual recordings are data streams. In order to be archived and retrieved these streams have to be segmented into events which can be annotated and associated with corresponding metadata. The process of stream segmentation is usually done manually and is in fact so time consuming, that many archivists, in large institutional archives as well as in small private collections, never actually manage to segment their streams. All of them face the problem of rapidly growing quantities of unindexed content, exploitation of which depends solely on the quality of editorial metadata if any has been provided. Hundreds of thousands of hours of audio and video recordings remain invisible, because their content is “hidden” in a stream (Proutskova, 2008).

There are various reasons why the task of annotating recordings is largely ignored by academic scholars: the work of media annotations is very time-consuming and it usually does not result directly in an academic product such as an article; this work is not rewarded in any way; academic scholars often do not want to deal with archives considering their media recordings to be their own property and of value for their research in the first place. At the same time, in fifty, hundred, two hundred years time our theories will be long forgotten, whereas our artefacts (audiovisual recordings with respective annotations) will gain importance while the time passes. Well-preserved archives and their holdings will outlive us while our recordings and annotations preserved by them will be our most valuable legacy to future generations (Seeger 1986).

There are endless ways to annotate time-related data such as audiovisual recordings. Here we consider four of them which are particularly relevant for AIRSPACE, together with software tools for their realisation. Here, Annotator's Workbench from the EVIADA project is described in more detail to provide a general sense of the annotation process. Then, alternative applications for other annotation scenarios are outline.

EVIADA's Annotator's Workbench (AWB) is intended for segmentation and annotation of audiovisual streams such as field recordings or, in case of AIRSPACE, documentations of music lessons. It best suits for complete audiovisual projects which are academic or ethnographic in nature, with a possible perspective of peer-reviewing and electronic publication.

Annotator's Workbench is open source and publicly available on the EVIADA website. It runs under Windows and MacOS X. It allows for manual segmentation of an audiovisual stream as well as for various modes of annotations for the segments including controlled vocabulary. A unique

feature of AWB is a possibility to provide within segment descriptions links to other segments (e.g. to other performances of the same song). Glossary and bibliography can be created for each project.

Since AWB is intended for annotating and archiving anthropological field recordings, there is no possibility to delete parts of the recordings, which is in compliance with archival practice. Instead, public access to segments can be managed: any segment can be closed for access for an indicated period of time.

Segmentation in EVIADA is a hierarchical process: first the stream is divided into main events (such as a concert or a lesson), which in turn are segmented into smaller bits (e.g. songs, interviews), etc. Also specific parts of the stream can be highlighted (e.g. a particularly beautiful part of a performance, a certain bit that needs work or teacher's explanations on a given subject). Segmentation is visualised in a timeline view, quick navigation as well as frame-precise settings are facilitated.

There is no one "correct" segmentation for a collection, each annotator would provide his/her own version, a segmentation can evolve when new information becomes available or more work has been done on the annotation.

After the stream has been segmented, each segment needs to be annotated. It would usually receive a title and a textual description, which can range from a short statement to a lengthy field diary entry. Usually the date/time and the place at which the recording was made is indicated and information is given about participants. For the purposes of professional cataloguing and interoperability editorial descriptions using ontologies or other kinds of controlled vocabularies can be added. Technical metadata can be added about the recording quality/conditions. Additional information and resources can be linked to the segments. They may include other formats like still images, links to relevant publications, music/lyrics transcriptions, etc. Textual transcriptions can be made in place and they do not need to comply with segments' boundaries.

Analysis of the recordings may follow the annotation process. It may involve transcriptions as well as investigating relationships and creating links between segments.

AWB project is an XML file containing all the information about created annotations. XML is an open format which is easily customised and presented online. EVIADA has a tool for displaying AWB projects which they use for their website, there are plans to make this tool public and open source. Online presentation of EVIADA content using this tool can be seen on their website.

For more modular projects or where more complex segmentation is required, the software of choice would ELAN[70]. It is aimed originally at language annotators and is a mature and widely used application. It allows segmentation and annotation of audio and video recordings. Any kind of annotation hierarchy or number of annotation layers is allowed. Also, tools for automatic audio analysis for ELAN have started to emerge, which facilitate manual segmentation and annotation.

iMovie[71] is Apple's software which is included into all Mac OSX distributions (thus it is free for all Mac users). It remains by far the best free video editing tool. It is ideal for creating new films and stories from original recordings as opposed to archiving everything within EVIADA. Recording snippets are easily cut and glued together in any order or fashion. Subtitles can be inserted and voice-over or other sounds can be added to the project. iMovie also features zooming in into single frames as well as changing playback speed. Many other effects and video manipulations are possible.

Variations project[72] at Indiana University has made available their Audio Timeliner application[73]. It is aimed at annotating a structure of a musical piece using bubble diagrams. It can be used with

---

[70] ELAN: http://www.lat-mpi.eu/tools/elan/
[71] iMovie: http://www.apple.com/uk/apps/imovie/?cid=mc-features-uk-g-fea-imo-imovie
[72] Variations3: http://www.dlib.indiana.edu/projects/variations3/index.html
[73] Variations Audio Timeliner: http://variations.sourceforge.net/vat/

audio files in many formats. Created structures can then be used to navigate audio.

V. Conclusions

Digital music libraries are complex, often multimodal repositories, including various object types. Among them are audiovisual recordings, which require special procedures for producing and preservation. Legal aspects are very important for all music collections; licensing and user rights management have to be implemented. Due to copyright restrictions, access to digital music resources for research is rare, which has created a real bottleneck in research progress and development.

The situation becomes even more complex for non-standard musical repertoires such as ethnomusicological collections, where conventional retrieval strategies (e.g. search for title and composer) have no relevance. Lack of standards makes each such collection an island and interoperability is largely hindered. Though a number of new digital resources in this field have emerged in recent years, they are still an exception rather then the rule.

This document presented a literature and electronic resources survey of digital music libraries and of ethnomusicological online repositories. It also provided an overview of practical recommendation for creating a new digital library of music, with a focus on non-standard repertoires and audiovisual collections. It was conducted for the AIRSPLACE project.

VI. References

[Agosti et al., 2007] Agosti, M., Bonfiglio-Dosio, G., and Ferro, N. (2007). A historical and contemporary study on annotations to derive key features for systems design. *International Journal on Digital Libraries*, 8:1–19. 10.1007/s00799-007-0010-0.

[Bainbridge et al., 2003] Bainbridge, D., Edgar, K., McPherson, J., and Witten, I. (2003). Managing change in a digital library system with many interface languages. In *Proc European Conference on Digital Libraries*, Trondheim, Norway.

[Balnaves, 2005] Balnaves, E. (2005). Systematic approaches to long term digital collection management. *Literary and Linguistic Computing*, 20:399–413.

[Becker et al., 2009] Becker, C., Kulovits, H., Guttenbrunner, M., Strodl, S., Rauber, A., and Hofman, H. (2009). Systematic planning for digital preservation: evaluating potential strategies and building preservation plans. *International Journal on Digital Libraries*, 10:133–157. 10.1007/s00799-009-0057-1.

[Berman, 2006] Berman, H. (2006). *Law And Revolution, II: The Impact of the Protestant Reformations on the Western Legal Tradition*. Harvard University Press.

[Bertin-Mahieux et al., 2011] Bertin-Mahieux, T., Ellis, D. P., Whitman, B., and Lamere, P. (2011). The million song dataset. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*.

[Bozkurt, 2008] Bozkurt, B. (2008). An automatic pitch analysis method for turkish maqam music. *Journal of New Music Research*, 37:1–13.

[Bradley and Lei, 2007] Bradley, K. and Lei, Junran; Blackall, C. (2007). Memory of the world: towards an open source repository and preservation system; recommendations on the implementation of an open source digital archival and preservation system and on related software development. Technical report, UNESCO.

[Byrd,] Byrd, D. Candidate music ir test collections: A list.

[Candela et al., 2010] Candela, L., Castelli, D., Fox, E., and Ioannidis, Y. (2010). On digital library foundations. *International Journal on Digital Libraries*, 11:37–39. 10.1007/s00799-011-0063-y.

[Choudhury et al., 2002] Choudhury, S., Hobbs, B., Lori, M., and Flores, N. (2002). A framework for evaluating digital library services. *D-Lib Magazin*, 8(7/8).

[Cornelis et al., 2010] Cornelis, O., Lesaffre, M., Moelants, D., and Leman, M. (2010). Access to ethnic music: Advances and perspectives in content-based music information retrieval. *Signal Processing*, 90(4):1008 – 1031. Special Section: Ethnic Music Audio Documents: From the Preservation to the Fruition.

[Crawford et al., 2010] Crawford, T., Mauch, M., and Rhodes, C. (2010). Recognizing classical works in historical recordings. In *11th International Society for Music Information Retrieval Conference (ISMIR 2010)*.

[D. Gleich and Zhukov, 2005] D. Gleich, M. Rasmussen, K. L. and Zhukov, L. (2005). The world of music: Sdp layout of high dimensional data. *Information Visualization*.

[Dattolo, 2010] Dattolo, A. (2010). Structuring and navigating ethnic music archives. *Signal Processing*, 90(4):1077 – 1082. Special Section: Ethnic Music Audio Documents: From the Preservation to the Fruition.

[Day and Wortman, 2000] Day, B. H. and Wortman, W. A. (2000). *Literature in English: A Guide for Librarians in the Digital Age*. Association of College and Research Libraries, Chicago.

[Diet and Kurth, 2007] Diet, J. and Kurth, F. (2007). The probado music repository at the bavarian state library. *Proceedings of the International Symposium on Music Information Retrieval*.

[Downie, 2003] Downie, J. S. (2003). Music information retrieval. *Annual Review of Information Science and Technology*, 37:295–340.

[Downie, 2008] Downie, J. S. (2008). The music information retrieval evaluation exchange (2005-2007): A window into music information retrieval research. *Acoustical Science and Technology*, 4(29):247–255.

[Fuhr et al., 2007] Fuhr, N., Tsakonas, G., Aalberg, T., Agosti, M., Hansen, P., Kapidakis, S., Klas, C.-P., Kovács, L., Landoni, M., Micsik, A., Papatheodorou, C., Peters, C., and Sølvberg, I. (2007). Evaluation of digital libraries. *International Journal on Digital Libraries*, 8:21–38. 10.1007/s00799-007-0011-z.

[Gomez et al., 2009] Gomez, E., Haro, M., and Herrera, P. (2009). Music and geography: Content description of musical audio from different parts of the world. *Proc of the 10th International Society for Music Information Retrieval Conference (ISMIR2009)*.

[Gonçalves et al., 2008] Gonçalves, M., Fox, E., and Watson, L. (2008). Towards a digital library theory: a formal digital library ontology. *International Journal on Digital Libraries*, 8:91–114. 10.1007/s00799-008-0033-1.

[Goussevskaia et al., 2008] Goussevskaia, O., Kuhn, M., Lorenzi, M., and Wattenhofer, R. (2008). From web to map: Exploring the world of music. In *Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT'08. IEEE/WIC/ACM International Conference on*, volume 1, pages 242–248. IEEE.

[Hart,] Hart, M. Project gutenberg mission statement. October 23, 2004, last checked 26/03/2012. http://www.gutenberg.org/wiki/Gutenberg:Project_Gutenberg_Mission_Statement_by_Michael_Hart

[Herwig, 2007] Herwig, M. (2007). Google's total library. *Spiegel Online International*

[Hilliges et al., 2006] Hilliges, O., Holzer, P., Klüber, R., and Butz, A. (2006). Audioradar: A metaphorical visualization for the navigation of large music collections. In Butz, A., Fisher, B., Krüger, A., and Olivier, P., editors, *Smart Graphics*, volume 4073 of *Lecture Notes in Computer Science*, pages 82–92. Springer Berlin / Heidelberg.

[Holzmann, 2009] Holzmann, G. (2009). Music information retrieval datasets. Technical report, Neural Information Processing Group, TU Berlin, Germany.

[IASA-TC03, 2005] (2005). *The safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy. Standards, Recommended Practices and Strategies (IASA-TC 03)*. IASA: International Association of Sound and Audiovisual Archives, South Africa.

[IASA-TC04, 2009] (2009). *Guidelines on the Production and Preservation of Digital Audio Objects: Standards, Recommended Practices and Strategies (IASA-TC 04)*. IASA (International Association for Sound- and Audiovisual Archives) Technical Committee, 2 edition.

[Koch et al., 2004] Koch, L.-C., Wiedmann, A., and Ziegler, S. (2004). The berlin phonogramm-archiv: A treasury of sound recordings. *Acoustical science and technology*, 25(4):227–231.

[Lai et al., 2008] Lai, C., Fujinaga, I., Descheneau, D., Frishkopf, M., Riley, J., Hafner, J., and McMillan, B. (2008). Metadata infrastructure for sound recordings. *Proceedings of the International Symposium on Music Information Retrieval*.

[Landone et al., 2007] Landone, C., Harrop, J., and Reiss, J. (2007). Enabling access to sound archives through integration, enrichment and retrieval: the easaier project. *Proceedings of the International Symposium on Music Information Retrieval*.

[Leech-Wilkinson, 2009] Leech-Wilkinson, D. (2009). *The Changing Sound of Music: Approaches to Studying Recorded Musical Performance*. CHARM, London.

[Licklider and Bolt, 1965] Licklider, J. and Bolt, B. (1965). *Libraries of the future*. M.I.T. Press.

[Lomax, 1976] Lomax, A. (1976). *Cantometrics: An Approach To The Anthropology Of Music*. Number 94720. The University of California, Extension Media Center, Berkeley, California. accompanied by 7 cassettes.

[Lyman and Varian,] Lyman, P. and Varian, H. R. (2003) How much information. Last checked: 28/03/2012. http://www.sims.berkeley.edu/how-much-info-2003

[Magas and Proutskova, pted] Magas, M. and Proutskova, P. (2012 accepted). A location-tracking interface for ethnomusicological collections. *International Journal on Digital Libraries, Special Issue on "Music Digital Libraries"*.

[Marolt, 2009] Marolt, M. (2009). Probabilistic segmentation and labeling of ethnomusicological field recordings. *Proc of the 10th International Society for Music Information Retrieval Conference (ISMIR2009)*.

[Martinez et al., 2009] Martinez, E., Celma, O., Sordo, M., de Jong, B., and Serra, X. (2009). Extending the folksonomies of freesound.org using content-based audio analysis. In *Proceedings of the SMC 2009 - 6th Sound and Music Computing Conference*.

[Mengel, 2007] Mengel, M. (2007). ethnoarc, linked european archives for ethnomusicological research. Technical

Report D4, Berlin Phonogram Archive.

[Mora et al., 2010] Mora, J., Gómez, F., Gómez, E., Escobar, F.-B., and Dìaz, J. M.-B. (2010). Characterization and melodic similarity of a cappella flamenco cantes. In *11th International Society for Music Information Retrieval Conference (ISMIR 2010)*.

[Orio et al., 2009] Orio, N., Snidaro, L., Canazza, S., and Foresti, G. (2009). Methodologies and tools for audio digital archives. *International Journal on Digital Libraries*, 10:201–220. 10.1007/s00799-010-0060-6.

[Payette and Lagoze, 1998] Payette, S. and Lagoze, C. (1998). Flexible and extensible digital object and repository architecture (fedora). In *Lecture Notes in Computer Science*, volume 1513 of *Research and Advanced Technology for Digital Libraries*. Springer.

[Pomerantz et al., 2009] Pomerantz, J., Abbas, J., and Mostafa, J. (2009). Teaching digital library concepts using digital library applications. *International Journal on Digital Libraries*, 10:1–13. 10.1007/s00799-008-0049-6.

[Proud, 1991] Proud, J. K. (1991). *The Oxford Text Archive*. Number 5985 in British Library research and development reports. British Library, Research and Development Dept., London.

[Proutskova and Magas, 2009] Proutskova, P. and Magas, M. (2009). Beyond the metadata: new intelligent audio content search for large music collections. presented at the Unlocking Audio 2 conference at the British Library.

[Proutskova, 2007] Proutskova, P. (2007). Musical memory of the world - data infrastructure in ethnomusicological archives. *Proceedings of the International Symposium on Music Information Retrieval*.

[Proutskova, 2008] Proutskova, P. (2008). Data infrastructure for ethnomusicological archives - current situation and future perspectives. *Journal of International Society of Sound- and Audiovisual Archives*, (31):45–54.

[Reigle, 2009] Reigle, R. (2009). Humanistic motivations in ethnomusicological recordings. In Dogantan-Dack, M., editor, *Recorded Music - Philosophical and Critical Reflections*. Middlesex University Press. with CD.

[Riley, 2008] Riley, J. (2008). Application of the functional requirements for bibliographic records (frbr) to music. *Proceedings of the International Symposium on Music Information Retrieval*.

[Seeger, 1986] Seeger, A. (1986). The role of sound archives in ethnomusicology today. *Ethnomusicology*, Spring/ Summer.

[Seeger, 1996] Seeger, A. (1996). Ethnomusicologists, archives, professional organisations, and the shifting ethics of intellectual property. *Yearbook for Traditional Music*, pages 87–107.

[Seeger and Chaudhuri, 2004] Seeger, A. and Chaudhuri, S., editors (2004). *Archives for the Future: Global Perspectives on Audiovisual Archives in the 21st Century*. Seagull Press, Calcutta.

[Six and Cornelis, 2011] Six, J. and Cornelis, O. (2011). Tarsos - a platform to explore pitch scales in non-western and western music. In *Proceedings of the 12th International Society for Music Information Retrieval Conference, ISMIR 2011*. International Society for Music Information Retrieval.

[Sridhar and Geetha, 2009] Sridhar, R. and Geetha, T. (2009). Raga identification of carnatic music for music information retrieva. *International Journal of Recent Trends in Engineering*, 1(1).

[Stumpf, 1908] Stumpf, C. (1908). Das Berliner Phonogrammarchiv. *Internationale Wochenschrift für Wissenschaft, Kunst und Technik*, 2:225–246.

[Tansley et al., 2003] Tansley, R., Bass, M., and Smith, M. (2003). Dspace as an open archival information system: Current status and future directions. In *Proc European Conference on Digital Libraries*, pages 446–460, Trondheim, Norway.

[Tzanetakis et al., 2007] Tzanetakis, G., Kapur, A., Schloss, W. A., and Wright, M. (2007). Computational ethnomusicology. *journal of interdisciplinary music studies*, 1(2):1–24.

[Váradi et al., 2008] Váradi, T., Krauwer, S., Wittenburg, P., Wynne, M., and Koskenniemi, K. (2008). Clarin: Common language resources and technology infrastructure. In Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odjik, J., Piperidis, S., and Tapias, D., editors, *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC08)*, Marrakech, Morocco. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2008/.

[Wellinga and Planqué, 2009] Wellinga, W.-P. and Planqué, B. (2009). Xeno-canto: web-based sound collecting and identification. presented at the Unlocking Audio 2 conference, British Library, London, UK.

[West et al., 2010] West, K., Kumar, A., Shirk, A., Downie, J. S., Ehmann, A., and Bay, M. (2010). The networked environment for music analysis (nema). *IEEE 2010 Fourth International Workshop on Scientific Workflows (SWF 2010)*.

[Witten et al., 2009] Witten, I., Bainbridge, D., and Nichols, D. (2009). *How to Build a Digital Library*. The Morgan Kaufmann Series in Multimedia Information and Systems. Morgan Kaufmann Publishers.

[Witten et al., 2005] Witten, I. H., Bainbridge, D., Tansley, R., Huang, C.-Y., and Don, K. J. (2005). Stoned - a bridge between greenstone and dspace. *D-Lib Magazine*, 11(9).