Enabling Access to Unpublished Audio Archives

Presentation to ALA Members, 17 January 2005

Why am I talking about this?

- Alexander Street Press; an academic publisher of electronic full-text databases and indexes in the humanities and social sciences

- We are experts in delivering digital music services:
  - world's first legal online music subscription service in 2001
  - World's first digital music library service in 2003
  - Now the largest and most exciting provider of digital music products to libraries
  - 150,000+ recordings licensed to date
  - Licenses from 35+ music labels and audio archives

Classical Music Library

Smithsonian Global Sound (June05)

African American Song (Autumn '05)

What's this showcase about?

- New technology has made it possible for libraries to make previously unpublished audio archives available online. However, this involves many important Legal and Technical considerations
- Alexander Street Press is in the process of creating 2 major new music collections:
  - Both collections feature previously unpublished audio recordings that have been digitised from a variety of original media
  - All recordings will be made available online
- This presentation gives you an insight into some of the issues involved

ALEXANDER STREET PRESS

What are unpublished audio archives?

What scale/type of audio is involved?

Example; Library of Congress

- 2.5m recordings, including >100k audio cassettes and 170k open-reel tapes
- Primary recordings of key artists e.g. Caruso
- Field recordings of sources which underlie much of modern music e.g. American and European folk traditions…
- Recordings of cultures lost due to modernization, anthropological and linguistic data e.g. traditional songs/chants
- Speeches & spoken words of historical figures e.g. Churchill, Roosevelt
- Early radio broadcasts (acetates)
- Live performances and events

ALEXANDER STREET PRESS

What is the opportunity (and risk!) now?

Technology has created an Opportunity

- The rapid pace of technological development in the last decade has created 2 valuable opportunities for audio archives:

FALLING COST OF DIGITIZATION

- Hardware for storing and delivering audio is much cheaper as technologies are standardized
- Software for processing and manipulating audio is much better/simpler e.g. Berkley initiative to take 3D scans of media surfaces
- Costly expert input may be reduced

INCREASING RETURN FROM COMMERCIALIZATION

- The development of a cheap infrastructure to deliver digitised audio (the internet!) and to listen to it (multimedia PCs) has created a growing market
- An increasing number of companies are able to create commercial opportunities for special collections of audio material

ALEXANDER STREET PRESS

What is the opportunity (and risk!) now?

But there are also some major risks …

- **Natural deterioration** of the underlying media; most recordings are designed for short term commercial use and not long term/permanent preservation e.g. magnetic tape has an expected lifespan of 10-30 years, recordable CDs are more prone to degradation than manufactured discs.
- **Wear and tear** of use (and misuse) e.g. a stylus will gradually deteriorate the quality of sound as the groove is worn away, very high heat can destroy the information on a MiniDisc.
- **Obsolescence of the technology** to capture and deliver the audio e.g. who still has equipment to play a Video Disc from 1978?
- **Obsolescence of standards** within the same media e.g. tape playback heads need to be in the same configuration as the track to prevent loss of sound quality
- **Increasing cost of supporting older formats** as trained engineers, and the spare parts they need, become rarer and more expensive

Legal & Technical Issues

1) Who can you provide access to?

Assuming you want to provide access to scholars and/or the public …

- US Copyright law is TORTUOUS
- Determining if an archive is in the public domain is complex and can depend on a mixture of US law, foreign law and State law
- If the archives are NOT in the public domain, then you have no right to digitise without permission from the copyright holder (save for preservation copies)

For example: The University of Chicago's High Fidelity Digitization On-line (Hi-Fido) project to provide public access to their Jazz recordings was stalled by copyright issues

For example: Hyperion, a classical music label based in the UK, was successfully sued by an academic who had produced amendments to a score for a work composed by a French composer who died in 1726. The academic was found to have created a new copyright because of the creativity required in reconstructing his version of the original work.

Legal & Technical Issues

1) Who can you provide access to?

Assuming you want to provide access to scholars and/or the public …

- US Copyright law is TORTUOUS
- Determining if an archive is in the public domain is complex and can depend on a mixture of US law, foreign law and State law
- If the archives are NOT in the public domain, then you have no right to digitise without permission from the copyright holder (save for preservation copies)

For example: The University of Chicago's High Fidelity Digitization On-line (Hi-Fido) project to provide public access to their Jazz recordings was stalled by copyright issues

For example: Hyperion, a classical music label based in the UK, was successfully sued by an academic who had produced amendments to a score for a work composed by a French composer who died in 1726. The academic was found to have created a new copyright because of the creativity required in reconstructing his version of the original work.

Legal & Technical Issues

2) The artists are uncontactable!

What if you can't contact the original artists (performers or composers)?

- They, or their estates/relatives (if dead), may be uncontactable?
- You may not even know their names or that they were involved?
  e.g. Field recordists seldom took notes of performer details
- If you use the music without permission, you open yourself to litigation if they turn up at a later date and contest your usage!
- Clearing licences in these situations can involve a lot of research, legal advice and working closely with the labels involved – experience definitely helps …

For example: One major label's policy is to assume permission is given if artists are uncontactable and then to remove their material if they turn up later and complain

Legal & Technical Issues

3) Cultural issues

- The artists may have cultural concerns over your use of their audio recordings
- These concerns may have been unknown/ignored at the original time of recording, but are now recognised as valid issues in today's society
- Consultation with artists and any relevant cultural organisations to better understand their concerns, and equally better explain your intentions, can overcome some objections
- Digitisation can benefit a cultural group by allowing their audio history to be preserved

For example: The LOC catalog includes some recordings of Peyote chants. These were recorded and released on LP early last century against the wishes of the native american elders, who saw these chants as sacred and only to be heard at special times and locations. They have yet to be released on CD because of this issue.

ALEXANDER STREET PRESS

Legal & Technical Issues

4) Concerns over making (or losing!) money

- Some copyright owners still believe the "dotcom dream" that the internet will make them rich and demand significant sums to provide online access
- They are often unaware that digital delivery is not free and has its own unique set of costs:
  - *conversion from original media*
  - *creation of metadata*
  - *cost of delivery*
  - *"cleaning"*
  - *storage of data*
  - *maintenance*
- Some copyright owners are simply not comfortable with providing digital access to their works/performances because of the risk of piracy
- Access to digital audio can be controlled in many ways:
  - At source e.g. password/IP authentication for authorised users only
  - After distribution e.g. Digital Rights Management tools can limit the use of the file by location, number of plays, and conversion to other formats
  - Ironically digital formats with DRM are far more secure than CDs!

ALEXANDER STREET PRESS

Legal & Technical Issues

5) Confusion over 'Mechanicals'

- Mechanical Royalties* are paid to composers for audio in copyright but …
- It's not always clear that something is in copyright … (see below)
- Composers can choose to collect their payments directly or through (a) their record label or (b) one of several mechanical rights agencies so there is no "single clearing house" to go to
- there is no agreement between the US collection societies as to how to approach this
- Some collection societies collect on behalf of other countries, most can't!

* Name derives from the original use as the right to mechanically reproduce a song on a piano roll

ALEXANDER STREET PRESS

Legal & Technical Issues

6) Confusion over standards & formats

SAMPLE RATE	Digitising an analog recording requires you to choose a sample rate; both the frequency of samples (kHz) and the size of each simple (bits) e.g. CDs are sampled at 44.1kHz and 16 bits, but LOC produce audio masters at 96kHz and 24 bits
CODEC	A CODEC is a system for compression and decompression of digital files and can be lossless or lossy (allows for smaller file sizes) e.g. MP3 (royalty bearing!), AAC (used by iTunes), Windows Media Player (one of the best), Vorbis (high quality open source codec)
BITRATE	Digital music can be streamed at a Constant Bitrate (best for limited bandwidth e.g. 22kbps or 64kbps), Variable Bitrate (best for downloads) or Multiple Bitrates (if you have the right audio player …)
AUDIO FORMAT	The most successful tend to be proprietary e.g. Windows Media Player (.wma, .wmv), Quicktime (.qt, .mov), iTunes (.m4p), Real (.ram, .ra)

ALEXANDER STREET PRESS

Legal & Technical Issues

7) Trade-offs in Sound Quality

The process of digitising a sound recording is an art, rather than a science:

CLEANING THE SOUND

- Audio software can automatically remove distortions/artifacts in the noise quality resulting from imperfections in the original medium and damage over the years
  - the more automated the process, the rougher the clean i.e. you also remove legitimate noise that can deaden/distort the original recording
  - the more manual the process, the more it costs!

COMPRESSING THE DATA

- LOSSY compression reduces the file size but also the sound quality by cutting out sounds beyond our range/processing capability sounds that are 'masked' by other sounds
- However, even LOSSLESS compression can create audible artifacts as a result of the electronic processes.

Legal & Technical Issues

8) Managing metadata, text & images

- An online archive of music is almost useless without accurate and comprehensive metadata to allow you to retrieve relevant recordings
  - Creation of metadata is labor intensive and costly because it often requires research e.g. to correctly classify the genre or recording date
- Often ignored in planning is the information that comes with the recordings, such as cover art, liner notes and other textual information
  - Each element should to be catalogued to ensure it can be linked to the original recording
- A relational database is necessary to deliver the functionality users will want e.g. separate tables for artists, compositions, performances etc
  - A controlled vocabulary is essential but can be costly to create
- Advanced indexing can enable searches that would otherwise be impossible
  - Alexander Street Press has developed Semantic Indexing™ to answer questions like “Find songs associated with the Freedom Rides of the Civil Rights era”

ALEXANDER STREET PRESS

Summary

1. **Get professional advice!** There is a wealth of expertise out there to be tapped for free and many organisations dedicated to assisting audio archives; e.g. National Digital Information Infrastructure and Preservation Program (NDIIPP) and the National Recording Preservation Board
2. **Conserve your existing media** e.g. professional standards exist to advise on the right way to store archive audio
3. **Preserve the recordings themselves**; digitisation is recommended as the best long term method for preservation and a well designed Digital Repository should assume media obsolescence and plan for reformatting, as well as periodically validating data integrity
4. **Ensure you have sufficient funding**! This can come both from external grants as well as commercial exploitation of the digital copies

Summary

5. CONTACT US!
   - If you have an audio archive containing material of value to either the academic community or to the general public please contact us and we can see if there is an opportunity for Alexander Street Press to help you get it online …

   - Go to www.alexanderstreet.com for more details