Sounding It Out: Sharing and Disseminating Audio-Visual Data

Louise Corti and Mus Ahmet
UK Data Archive
University of Essex, UK
Qualitative data

• various data types created:
  - interview/group recordings and transcripts
  - field notes and diaries
  - audio visual sources - still and moving

• interviews often captured as an audio source though recordings

• visual data increasingly being captured in the field
Sharing audio-visual data

Can be a challenge for social research data:

- ethical and consent issues
- quality of recording
- various proprietary formats
- storage capacity
What does UKDA currently do?

- we acquire audio-visual data as part of qualitative collections - but not routinely
- we deal with both analogue and born digital data
- we evaluate data and consider ethical issues
- we do NOT anonymise any audio visual data
- we process, document and catalogue the data
- we disseminate or provide links to the data
Key challenges

- minimising data storage requirements whilst simultaneously maximising audio quality

- using sustainable open formats for long term preservation

- using optimal and flexible formats for delivery
Analogue audio data

formats we deal with:

- reel-to-reel tape
- audio cassette
- micro cassette (e.g. dictaphone)

- we list the tapes with all the metadata we have (later...)

- we MAY digitise all or selected extracts

- latter can be labour intensive!
Digitising audio

• have digitised a large reel to reel collection (450 interviews each approx. 3 -10 hrs long)

• have digitised collections on audio-cassette

• typically we digitise at frequency 48kHz /24 Bits
Reel-to-reel digitising project

Our best loved study: The Edwardians

- conducted by the British Library with tight spec from us - agreed metadata
- 1203 reels converted by external supplier
- no filtering or post-processing applied
- dead air checks and long gaps edited out
- processing /metadatabase created
- batches returned every 3 weeks, audit and random listening. MD5 checksum confirmed
- 2517 .wav files produced
- HUGE files!!! 2 terabytes of files

- dissemination solution? Consent agreed, so just technical
Cassette Digitising Hardware

Plusdeck 2c is an innovative computer-controlled audiocassette recorder that lets you convert cassette into Wave or MP3 files (allows duplex).

Specifications:
- Cassette Speed: 4.75 cm/sec
- Frequency Response: 20 ~ 18,000 Hz
- Separation: 40dB
- Signal to Noise Ratio: 55 dB

Front View

- Tape Door
- Tape Availability, Playback Indicator LED
- A / B Side, Pause, Stop Buttons
- Output Volume
- Output, Input Jack
- Eject Tape, Fast Forward, Rewind Buttons
Analogue to Digital Conversion Software

The Plusdeck 2c is supplied with easy to use software that captures the analogue output from the cassette and converts it to Wave or MP3 files. Post-processing using Audacity
What’s The Difference Between ‘Lossy’ and ‘Lossless’ Compression?

There are two broad classes of compression algorithms:

Lossy

• Lossy compression algorithms discard data in order to compress it more than would normally be possible. Examples include JPEG, Vorbis, and MP3 compression.

Lossless

• Lossless compression algorithms produce compressed data that can be decoded to output that is identical to the original. Zip is a common general-purpose lossless compression format.
Audio File Formats

• it is important to distinguish between a file format and a codec

• a codec encodes and decodes the raw audio data, the data itself is stored in a file with a specific format

• most audio file formats support only one codec, however, some file formats support multiple codecs, for example AVI

• there are three major groups of audio file formats
  - uncompressed audio formats: WAV, AIFF
  - formats with lossless compression: FLAC, Apple
  - formats with lossy compression: MP3, Ogg Vorbis, and AAC
# Digital audio data we could deal with

<table>
<thead>
<tr>
<th>Format</th>
<th>Name</th>
<th>Open source?</th>
<th>Owner/IPR</th>
<th>Lossless?</th>
</tr>
</thead>
<tbody>
<tr>
<td>.wav</td>
<td>Waveform audio format</td>
<td>✔</td>
<td>Microsoft &amp; IBM</td>
<td>✔</td>
</tr>
<tr>
<td>.mp3</td>
<td>MPEG-1 Audio Layer 3</td>
<td>✗</td>
<td>Fraunhofer and Thomson</td>
<td>✗</td>
</tr>
<tr>
<td>.flac</td>
<td>Free Lossless Audio Codec</td>
<td>✔</td>
<td>Xiph.Org Foundation</td>
<td>✔</td>
</tr>
<tr>
<td>.aiff</td>
<td>Audio Interchange File Format</td>
<td>✗</td>
<td>Apple (Macs)</td>
<td>✔</td>
</tr>
<tr>
<td>.oma; ..omg; .atp</td>
<td>Minidisc (MD)</td>
<td>✗</td>
<td>Sony</td>
<td>✗</td>
</tr>
<tr>
<td>.m4a</td>
<td>Advanced Audio Coding (AAC)</td>
<td>✗</td>
<td>Apple (iTunes)</td>
<td>✗</td>
</tr>
<tr>
<td>.ogg</td>
<td>Ogg Vorbis</td>
<td>✔</td>
<td>Xiph.Org Foundation</td>
<td>✗</td>
</tr>
</tbody>
</table>
Preservation Formats for Audio

• preservation formats should, ideally be lossless, non-proprietary and open source, and cost-free to process

• the only two widely adopted formats that fulfil these criteria are Wave and FLAC
What’s so good about FLAC?

• FLAC is a file format for audio data compression
• like MP3 it is a compression format, but it is lossless
• compression works in a similar way to WinZip
• does not remove information from the audio stream as lossy compression formats
• reduction of bandwidth or storage requirements but without sacrificing the integrity of the audio source
• audio sources encoded to FLAC are typically reduced in size by 50%
• can be decompressed into an identical copy of the audio data
• it’s currently the fastest and most widely supported lossless audio codec
• oh, and it’s totally free and Open Source!
Dissemination Formats for Audio

- files disseminated via the internet by streaming or download need to be:
  - a reasonable size
  - in an accessible format
  - free, non-proprietary and open source - wherever possible

- unfortunately, the universal adoption of the MP3 standard, and lack of native support for other formats means we generally disseminate in MP3 format

- however, we also encode all audio files to the Ogg Vorbis format and will disseminate in that format if and when it becomes viable
What is an ‘ogg’?

• An ‘ogg’ is often incorrectly understood to refer to an audio file

• but it is actually wrapper for multimedia content codecs

• *a wrapper is a package which changes the interface to an existing package without substantially increasing its functionality*

• Ogg can be used to wrap lossless audio, compressed video, and various other multimedia codecs

• xiph.org (http://www.xiph.org)
What is "vorbis"?

- Vorbis is the name of a specific audio compression scheme that's designed to be contained in Ogg.
- For the average user, Vorbis is simply an alternate patent free version of mp3.
- It offers better quality at comparable file size.
- http://www.vorbis.org
## Summary Audio Formats

<table>
<thead>
<tr>
<th>Format</th>
<th>Compressed?</th>
<th>Preservation?</th>
<th>Dissemination?</th>
</tr>
</thead>
<tbody>
<tr>
<td>.wav</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>.flac</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>.mp3</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>vorbis</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>
How do we prepare data?

- Ingest
- *wav*
  - Preserve an original copy of data
  - Confirm the data is identical using Flac verify option
- *flac*
  - Preservation copy
  - Convert whole file or extracts
- *mp3*
  - Dissemination copy: download and streaming
Current UKDA audio download

- audio data in mp3 format is available from our web download system like other data

- download as zip with metadata in spreadsheet (DataList - interview level, NOT file level)

- is kept behind access control system

- is a very limited solution and OK for smaller collections but not good for collections of large files
Streaming and file-based download

- for our large Edwardians collection we have 450 interviews
  - we created multiple mp3s for each interview
  - too large for download zips

- will be enabling:
  - selective file-based download
  - streaming of extracts (a taster)
  - whole file streaming

- already have an online search and visualisation for this textual collection so easy to link in ESDS Qualidata Online
### Interview summary search (initial results)

Your search found the following record(s). Please tick those you are most interested in and click to retrieve the full interview summary record(s). You may bookmark this search.

You searched for ALL of the following terms:

**Free-text:** illness  
**Sex:** Female

<table>
<thead>
<tr>
<th>Name</th>
<th>Geographical Region</th>
<th>Occupational Class</th>
<th>Interview Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miss Algy P</td>
<td>Scotland</td>
<td>Semi-skilled Manual</td>
<td>163</td>
</tr>
<tr>
<td>Mrs D</td>
<td>London</td>
<td>Clerical and Foreman</td>
<td>263</td>
</tr>
<tr>
<td>Mrs H</td>
<td>Eastern</td>
<td>Unclassified</td>
<td>317</td>
</tr>
<tr>
<td>Mrs Alice Emma G</td>
<td>London</td>
<td>Unclassified</td>
<td>215</td>
</tr>
<tr>
<td>Mrs Alice Maud S</td>
<td>North West</td>
<td>Unclassified</td>
<td>68</td>
</tr>
<tr>
<td>Mrs Alison LS</td>
<td>Scotland</td>
<td>Unclassified</td>
<td>436</td>
</tr>
<tr>
<td>Mrs Amelia K</td>
<td>Eastern</td>
<td>Unclassified</td>
<td>125</td>
</tr>
<tr>
<td>Mrs Amelia W</td>
<td>West Midlands</td>
<td>Unclassified</td>
<td>319</td>
</tr>
<tr>
<td>Miss Amelia Margaret Ross S</td>
<td>North West</td>
<td>Skilled Manual</td>
<td>115</td>
</tr>
<tr>
<td>Mrs Anna M</td>
<td>Wales</td>
<td>Skilled Manual</td>
<td>421</td>
</tr>
</tbody>
</table>
Pioneers project

- also have a collection of in depth life stories with classic sociologists who have created data rich projects

- dedicated website
  - bios
  - key publications
  - datasets created
  - life interview
  - thematic extracts of interviews

- used as test bed for streaming
http://www.data-archive.ac.uk/qualidata/pioneers/thompson/becoming%20a%20researcher.asp

http://www.data-archive.ac.uk/qualidata/pioneers/thompson/interview.asp
Why Did We Choose Adobe Flash Player?

- Adobe Flash is ‘standard issue’ on nearly every computer
- no external programs like QuickTime are required
- visitors won't have to do anything special to play the media

Adobe® Flash® Player is the world’s most pervasive software platform and reaches over 98.8% of Internet-enabled desktops in mature markets.

Mature Markets include US, Canada, UK, France, Germany, Japan.
Why Wimpy Rave?

- uses Adobe Flash
- easy to install
- highly customizable
- readily & Seamlessly Integrated into websites
- universal - runs on any basic Web server
- flexible - runs under ASP, PHP Cold Fusion
- automatically lists and plays the contents of a folder
- can run off XML, RSS or Text playlists
- javascript Integration - control nearly every aspect of Wimpy Rave via Javascript
- encrypt File Names- Use URL encrypter or Playlister for Wimpy Rave to hide locations of file names
- Ecommerce Integration - put file into a shopping cart to allow download

only $89!
Words of advice

- use a dedicated streaming server as very resource intensive
A tiny bit about metadata (really)

- always aim for rich metadata

- collect audio-visual file metadata in a database/XML - do not rely on information embedded in the audio file as lost on conversion

- UKDA currently does not capture any file level metadata file level DDI fields are not utilised in house...but they could be. Are looking at this as very necessary

- Chosen to use Library of Congress (LoC) metadata based on current best practice

- Use in-house but logical file naming convention for audio visual data
  - Study numberinterview number 2000int001.rtf
  - Study numberinterview number_audio file 2000int001_1.rtf
  - Study numberinterview number_clip number
  - 2000int001_clip1.rtf
QUDeX

- XML schema for storing annotated multimedia data - version 3 released

- built to represent core functions of CAQDAS packages
  - coding, classifying, memoing, search and retrieval
  - text, audio-visual data, URIs etc
  - can relate/link files, segments, codes, notes, URIs and analysis

- significant input from ODaF
  - I

- hopefully under DDI committee for review/tools development

- volunteers welcomed and testers needed

- http://www.data-archive.ac.uk/dext/
QuDEx

• vendors have been consulted throughout the QuDex development process

• QuDEx aims to deliver core functionality while other metadata standards are used as appropriate

• recommended that all materials and metadata relating to a collection are packaged as ‘complex objects’ using a standard such as METS

• **METS: Metadata Encoding and Transmission Standard**

• used for encoding descriptive, administrative, and structural metadata regarding digital objects, expressed using the XML schema language (Library of Congress maintained)
Audio-visual XML

</segment> <!-- Audio segments -->
   label="Audio segment: games" displayLabel="First audio segment" creator="Angad"
   language="en">
  - <audio id="audio_1" src="document_2">
    <param id="p_1" clipType="time" clipBegin="00:02:22" clipEnd="00:04:44" />
  </audio>
</segment>

<!-- Video segments -->
   label="Video segment: time pass" displayLabel="First Video segment" creator="Angad"
   language="en">
  - <video id="video_1" src="document_2">
    <param id="p_2" clipType="smil" clipBegin="2.75min" clipEnd="5.34min" />
  </video>
</segment>

<!-- xml based segments -->
   label="Xml segment: childrens" displayLabel="First Xml segment" creator="Angad"
   language="en">
  - <xml id="xml_1">
    <node id="r1" src="document_4" xPtrExp="xpointer(/root/U[1]/range-to(/root/U[10]))" />
  </xml>
</segment>

<!-- Image segments -->
   label="Image segment: son-in-law" displayLabel="First Image segment" creator="Angad"
   language="en">
  - <image id="image_1" src="document_5">
    <area id="a_1" coords="12,34,54,67" shape="poly" />
  </image>
</segment>
Other solutions

- wish to synchronise text and audio
- other solutions possible - mostly proprietary
- Fedora/Dspace repository systems - multi media files can be ingested, previewed and downloaded as part of the whole collection
- wish to use QuDEx to reference text to audio - visual materials
- Open source tools that may be explored to create dynamic XML-based audio visual experience eg using Adobe FLEX AIR tools (ODaF)
FEDORA based viewing systems

Project Contextual Interview