Audiovisual and Multimedia Section
Training workshop
AV Collections for non-specialist librarians
Singapore, 2013

Access:
on site, online
metadata
access formats

James M Turner
Université de Montréal
Faculté des arts et des sciences
École de bibliothéconomie et des sciences de l'information

Outline-

- Introduction
- On-site access
- Online access
- Metadata
- Access formats

Introduction-

- Preservation and access : associated concepts
- Different kinds of access:
  - to physical object
  - to a particular part of an object (one song, a scene)
  - to a file on a computer or on the web
  - subject access, intellectual access
- We will look at access on site, online, metadata for access, and formats for access

On-site access

- For audiovisual materials, you need space and equipment
- Space:
  - costs money
  - is the space available needed more for other purposes?
  - staff need to manage, supervise, help with machines
  - a/v materials have sound, so users need cabins or earphones (alternately: they bring their own)
- Unlike books, all a/v materials require equipment for consulting

Equipment

- What equipment you have depends on:
  - what a/v materials you have
  - what you can afford to buy & maintain
- Using the equipment:
  - staff mount the a/v materials for viewing/listening
  - users mount their own materials, staff train them
  - cleaning, reporting malfunctions, repairs
  - service contracts or trained staff

Viewing / listening

- Preservation copies, viewing copies:
  - what kind of collection you have (deposit, public)
  - series of generations: master, sub-master, viewing copy
  - managing these copies, responsibility for making them
  - master copies should never or rarely be projected
  - no master = periodically replace viewing copy

Digital copies

- Digital materials have many advantages:
  - copies are easy to make
  - quality is maintained from one copy to the next
  - same equipment (a computer) for movies, video, sound
  - less training, users may already know how to use
  - you can use a web interface
  - computers relatively cheap to buy, maintain
- Several options for organising

Options for organising digital copies

- Server
  - copies are stored together, file format can be standard
  - security easier to maintain
  - users never touch the materials, can't damage them
  - costs lower than managing analogue materials
- Hard disk
  - relatively small, inexpensive
  - storage costs relatively low (<$100/Terabyte)
  - reliable, less handling, less damage than DVDs
- DVD (next slide)

Options for organising digital copies

- DVD
  - relatively cheap
  - buy commercial films, music
  - if you hold the rights, record your own films, sound
  - careful labelling, protection necessary
- For all organising options:
  - work out a timetable for replacing equipment
  - organise servicing the equipment
  - make sure staff are trained adequately

Online access

- If you already have a server for on-site access, you can use it for online access too
- A separate server is desirable (better security)
- One staff member as system administrator, or a service contract with a company
- Integrate with the library web site
- A decision: public, or members only?

Similar to in-house digital a/v

- If you already have digital copies for in-house use, most of the issues are the same
  - buying and maintaining equipment, materials
  - procedures for making copies
  - staff and user training
  - file formats determined by access policy

Good for statistics

- Online access will improve statistics:
  - number of users
  - number of times each document is consulted
  - collect data on user satisfaction
- Provide arguments for budget, service:
  - increased visibility
  - available any time, day or night
  - lower equipment, service costs = more acquisitions?
  - free up staff for other tasks

Learn and set up

- This is no longer new territory
- Lots of models available
- Lots of information about what to do and what NOT to do
- Experience of other libraries, colleagues helps you avoid mistakes
- Choose a good project leader to ensure success

Metadata

- Importance of metadata
- Kinds of metadata
- Costs
- Compromises
- Metadata for access

Importance of metadata

- Metadata (information about the a/v materials you collect) is essential for:
  - knowing what you have (inventory, catalogue)
  - finding anything (storing, retrieving)
  - communicating with users, colleagues, systems
  - making acquisitions, using, preserving
  - comparing different versions, recordings

Kinds of metadata

- Many, many kinds exist
- What you need depends on what you collect, who uses the collection
- What you can afford depends on staff, budget

Metadata services

- For films & tv on DVD, catalogue services like OCLC have the records
- Also for commercially recorded music
- Staff may have to adjust these records
- Original cataloguing for other materials
- Simple indexing often ok (genre, audience, general subject categories)
- Can users help, for example by adding tags?

Costs associated with metadata

- Lots of initial work to decide what to use
- You start from zero, expensive at first
- Cost for updating, adding new metadata as needed
- Don't add more than you need
- Automate as much as possible
- Read, talk with colleagues elsewhere about developments

Compromises

- Formerly, rigourous standards
- Now, much more flexibility because of the web
- Cataloguers are rigourous by nature, but detailed cataloguing is expensive now
- How much or how little do you need?
- Are simple standards like the Dublin Core ok for your collection?
- Is simple indexing good enough?

How things have changed

- Formerly, fewer photos, more detailed metadata
- Now, digital photos easier to take, too many to catalogue individually
- Lots of commercial movie and tv production, but also YouTube, Vimeo, Daily Motion, etc.
- Musicians no longer need the record companies to publish their creations
- Film and video makers no longer need the studios
- Do we still need the librarians?

More changes

- Artists can have their own web site
- Remix of commercial recorded music
- Video mashups
- Rights management, use, other issues difficult
- Google is used to search for everything
- Tagging instead of indexing

Metadata for access

- Metadata for physical / digital access
- Metadata for intellectual access

Physical/digital access

- This has to do with finding physical pieces or locating files in a database or on a computer
- Physical access: a call number, an address on a shelf
- Digital access:
  - searchable fields in a database (title, actors, running time…)
  - predictable organisation, names for file folders, files
  - an archival classification

Intellectual access

- Like for text, messy and not so predictable
- This is because interpretation can be different from one person to the next
- When materials are networked, we don't know the users so well

User needs

- Intellectual access for different collections:
  - art collections: artist, school, genre, theme…
  - documentary collections: who, what, how, when, where
  - ordinary photos: tags to name the objects, actions
- Indexing should meet user needs:
  - art collections: controlled vocabulary, authority lists, standardised tools (e.g. *Art & architecture thesaurus*, *Iconclass*)
  - documentary collections: in-house subject list, proper names important
  - ordinary photos: tags, *Thesaurus for graphic materials*

Intellectual access to sound

- For published music, use existing vocabularies:
  - publisher's, musicologist's catalogue for classical music
  - categories: jazz, pop, rock…
  - composer, musicians, titles
- For readings of plays, poetry, speeches, etc. :
  - cataloguing, subject headings, indexing as for books
- Stock sounds, sound effects for movies, etc. :
  - describe with words: "dog barking", "thunder"

Access formats

- Different file formats
- Preservation formats
- Access formats

Different file formats

- With all digital materials, we need different file formats for preservation than for access
- For preservation:
  - formats that will last a long time
  - open source, W3C recommendation, public, widely used
  - uncompressed data
- For access:
  - depends on user needs
  - more flexible, a variety of formats ok
  - compression ok

Preservation formats for audio

- As we saw in the audio module, .wav, BWF, MBWF / RF64 are recommended because they are the closest thing to a standard
- Physical carrier: hard disk, tape, LTO tape (better)
- For the short term, CD or DVD ok

Preservation formats for video

- For archiving production collections, such as a tv archive:
  - MXF (metadata wrapper) + JPEG2000
  - physical carrier: LTO tape
- For mixed library collections:
  - Already compressed, probably acquired on DVD
  - Preserve this copy, replace when damaged
  - National library: preservation copy, distribution copy

A word on Motion JPEG

- Everyone is familiar with JPEG
- JPEG2000 is an improvement on JPEG:
  - compression can be reversed, no loss of data
  - came after JPEG was already widely used, so people don't know about it
- Motion JPEG /Motion JPEG2000 is a way to compress moving images:
  - each frame of a film or video is compressed like a photo
  - much better quality than MPEG2, which destroys much of the data

Digital materials not permanent

- No stable physical carriers for digital files
- So digital materials can never "reside" anywhere for very long
- They are somewhat "homeless", have no fixed address
- Digital preservation activities revolve around keeping the files "alive", readable

A word on digital preservation

- With digital materials, don't try to preserve the physical carrier (tape, disc)
- Instead, copy to new carrier from time to time ("refreshing")
- Move to new format, new version as necessary ("migration")
- Keep the material "alive" so newer computers, operating systems, software can still play it

A word on digital master copies

- The original or "master" copy is the one to preserve
- But with a/v materials, what this means exactly is not clear
- One way to understand it: the master copy is the best one you have
- Since digital copying makes an exact copy, copies are identical
- However, different "generations" for access

Access formats

- Ideally, the original or "master" copy is not compressed (but: JPEG2000 is ok)
- Access formats can be compressed because they are not used for preservation
- Different degrees, types of compression
- Different formats, depending on user needs

One model (e.g. for an a/v archive)

- Master copy (not compressed)
  - Submaster copy (not compressed)
    - Access copy (compressed a little)
      - Access copy (compressed more)
        - Access copy (compressed more, for web)

Example: a photograph

- Master copy (untouched)
  - Submaster copy (used to make access copies)
    - Access copy (to use in a book, little compression)
      - Access copy (to use in a presentation more compression ok)
        - Access copy (to use on the web, highly compressed)

Compressing access copies

- Creates smaller file sizes
- More flexibility with physical carriers
- Users have different needs
- Lower quality helps protect ownership, copyright
- On the web, small file sizes means:
  - faster transmission time across the net
  - shorter time to load the file to your computer
  - smoother viewing, listening, no pauses

Access formats for photos

- JPEG
- JPEG2000
- GIF
- PNG

Access formats for moving images

- Recommended
  - .mp4 (MPEG 4)
  - .mov (QuickTime)
  - .wmv (Windows Media Video)
- Acceptable
  - .swf (Macromedia Flash)
  - .flv (Flash Video)
  - .rm (Real Media)

Access formats for sound

- Recommended
  - .mp3 (MPEG 4)
  - .AIFF (Apple Computer)
  - .wav (Windows)
- Acceptable
  - .ogg (Ogg Vorbis)
  - .flv (Macromedia Flash Video)
  - .ra (Real Audio)

Conclusion

- Everything changes: keep reading
- Talk, visit with colleagues
- The web a good source of information