

Visualising sound, annotating files and FRBRzing archive

Juha Korvenpää

My presentation deals with ideas from experiences of c-cassette digitisation. The theme of this conference is “re-evolution”. I too “re-evolve” ideas that could touch certain problems that we have encountered. New information about an audio object will emerge during digitisation process. How to pass that information to end-users? And how could the library benefit from user activity?

The National Library of Finland (NLF) digitises audio objects firstly on demand: the records needed immediately for research are digitised in the studio of the Music Library in Helsinki. Secondly, the National Library digitises and preserves material by concentrating on extensive bodies of material under threat, such as the legal deposit tapes. The Mikkeli-based Centre for Preservation and Digitisation has been digitising the National Library’s collection of cassette tapes over the past few years. Fragile materials will be digitised first, e.g. publications only in cassette format, which makes them unique and therefore a priority for preservation.

Digitising thousands of music c-cassettes makes it obvious that some songs or tracks will be digitised several times. There are songs so popular that they have been published on several different releases. And we haven’t even started digitising other formats, vinyl records or CD discs, which will contain other repetitions of these hit songs. In addition to our own digitisation, NLF obtains digital audio files also from other digitising partners (e.g. record companies) and in the future NLF probably receives born-digital files (donations, legal deposit). It is important to know what the source of digitisation is. The source is rarely the master recording, usually we digitise from commercial releases.

In the end we are in a situation where the Library’s audio archive contains many digital files of one track. These digital files will vary in their quality aspects. How to choose the best one for listening? One solution could be to use technical metadata in choosing the best copy among many versions. Possible digitisation defects should be documented in technical metadata and this should be made available for users. We could also gather feedback from digitisers and end-users to add information to our database about the condition of the sound files.

The National Library of Finland has already tested crowdsourcing for newspaper material. We should collect user’s comments concerning our audio files too, although audio material can’t be put freely on the web due to copyright restrictions. In-copyright material will be available for users in dedicated computers within the Music Library of NLF. Nevertheless, we should be able to annotate digital files that we produce. Annotating files could help users to find a better sound file for listening and to avoid unnecessary complaints from users - they might report about flaws that were already noticed during the digitisation process.

If files could be annotated, where would these annotations be saved? And more importantly, we should distinguish to which database record this annotation is attached. During digitisation we quite often find flaws that were originally on the tape and were not produced during digitisation. Digitisation may be flawed because of the poor condition of the original sound carrier. Digitisation may flaw also due to defects during the tape production stage (i.e. copying the commercial products from a master recording tape). One flaw can be present in all the releases of a particular song in a certain format (e.g. c-cassette) but absent in another format (vinyl record).

User experience could be improved also by visualising sound. I’d like to see a system where sound is represented in a visual form in addition to the album cover picture. Current music and sound

production systems represent sound as a two dimensional waveform, the sound level on the vertical and the time on the horizontal axis. Waveform indicates many aspects of sound files: the changes in time and volume, the balance between stereo channels, the overall content of a sound file. The visual representation makes sound files more concrete. One can “see” sound.

Songs and sounds created with personal computers can be uploaded to cloud storage services where this material can be listened to, downloaded and even commented by adding a free text note to a specific location in a waveform i.e. a certain timestamp in an audio file. These services clearly address some shortcomings of current library systems, at least when it comes to audiovisual material. Quite probably current practices are reflected in future requirements for library systems.

A step further would be taking these features; the waveform representation and sound file commenting, and implement them into a sound archive, which relies on a FRBR-like data model. FRBR is a conceptual entity-relationship model. I see it analogous to audio production hierarchy, where one master recording is the source from which other versions are produced. FRBR-model could be implemented (and already is, in Variations-project) for digital sound files. In FRBR terms every sound file is digitised from an item which itself is an instance of a certain manifestation. (Considering all aspects related to FRBR entities and their relations need more attention than given here.)

A sound archive may include many items from one manifestation. A particular item (digital file) will probably have some technical differences compared to other items. User commenting could help improving the usability of sound archives: by helping users to find correct sound files, by minimising cataloguing efforts in expression and manifestation level and by ensuring that technical comments are focused on correct sound files (items).