The least common denominator does not satisfy user's needs: metadata schemes for digital audio archives

Elke Greifeneder
Royal School of Library and Information Science, University of Copenhagen, Denmark,
ljb301@iva.ku.dk

Bente Larsen
Royal School of Library and Information Science, University of Copenhagen, Denmark,
bentelar@hum.ku.dk

Birger Larsen
Department of Communication, Aalborg University, Copenhagen, Denmark,
birger@hum.aau.dk

Haakon Lund
Royal School of Library and Information Science, University of Copenhagen, Denmark,
tcm410@iva.ku.dk

Mette Skov
Department of Communication and Psychology, Aalborg University, Denmark,
skov@hum.aau.dk

Abstract:

This paper examines what metadata schemes are needed for radio material using the example of the LARM project. Based on several user studies in 2011, a metadata scheme has been implemented, which consists of three layers: archival metadata, LARM metadata and project-specific metadata. In late 2013 follow-up online interviews assessed the adoption of the metadata schemes and users' understanding of the underlying philosophy. The interviews showed two things: first, despite being involved in building the scheme up, participants used the current metadata scheme on a trial and error basis and did rarely understand the underlying philosophy. Secondly, participants showed a surprisingly high accountability working with metadata meaning that archive providers should worry less about misuse of metadata and instead embrace user participation.

Keywords: metadata, audio, user-generated content, information behaviour, e-infrastructure

1 Introduction

The digitalization and preservation of our cultural heritage has been and still is a major priority and the investment made in large-scale digitalization projects has been substantial. This considerable investment in digitalization of our cultural heritage and the foundation of research infrastructures within digital humanities has found support from the political system resulting in an increased awareness among politician of the importance of digitizing our cultural heritage to secure public access also in the future to these treasured resources.

Radio material has long been out of the focus of researchers, because these materials have essentially been locked away mainly due to copyright restrictions. An important part of our cultural heritage is broadcasted radio and the Danish LARM infrastructure project is an example of an eScience research initiative within this domain. However, this legacy of radio is very poorly indexed with metadata. Available metadata for most programs is the title of the program, the date of transmission, the channel, the length of the program and the name of the broadcaster. Occasionally it is possibly to find some additional information about the content of the often hour long programs. In the case of LARM, funding for in-depths indexing was not available and thus the challenge was to find ways to provide metadata that are easy to work with, easily extensible and that allow flexible data exchange (Lund et al., 2013).

The remainder of the paper is structured as follows: first, related literature and the LARM project are presented, followed by the research questions. Secondly, methodology and results of a first user study that elucidated metadata requirements and the resulting metadata schemes are discussed. The paper closes with the results of a follow-up assessment of these metadata schemes using synchronous qualitative online interviews.

2 Background

Retrieval of audio materials is a growing topic of interest as the increase of membership and publications of the annual conference of the International Society for Music Information Retrieval (ISMIR, 2014) demonstrate. There has been substantial research on how users seek music (for a good overview see Lee, 2010) and what metadata schemes might be needed in consequence. To the authors' knowledge, few of these results have been implemented, if any at all. The reason might be a disconnection between information retrieval and classification research. Exchange between research on users' seeking behaviour and classification research is most active around the implementation of the standards Resource Description and Access (RDA) and Functional Requirements for Bibliographic Records (FRBR) and several user studies on metadata requirements have been conducted (for example Ayres, 2005; Kanai, 2013; Riley, 2011). Further, Vainio et al. (2009) conducted a user study to assess metadata requirements in a mobile multimedia content and Gligorov (2012) examined crowdsourcing as a mechanism for metadata production in an audio-visual context.

User studies on metadata requirements in audio archives, in particular radio archives are sparse. Kim et al. (2003) examined search behaviour in large collections of recorded speech and Huurnink et al. (2010) conducted a transaction log analysis of search behaviour in an audio-visual archive.

The LARM research project explicitly chooses a bottom-up approach where metadata development took place in close cooperation with the researchers taking part in the project.

The LARM digital archive hosts one million hours of Danish broadcast radio, for which all Danish university researchers can register for free (see http://www.larm-archive.org/) and gain access to content trough the LARM.fm interface. LARM.fm is primarily a digital archive, but it also offers several services of an eScience environment such as search and annotation tools and also tools for collaboration i.e. upload and download of files, shared folders and project specific metadata schemes. As of today, almost 800 users are registered coming from various research disciplines including both researchers and students. The hope is that a wider audience could be allowed access in the future.

LARM was a joint project with participation from the Danish national broadcasting corporation, the State and University Library hosting the Danish Media Archive, University of Copenhagen, Aalborg University, University of Southern Denmark, Roskilde University, The Danish Media Museum and the national Danish e-infrastructure consortium (LARM.fm blog, 2014). The project was funded by the Danish Ministry of Science with 3.5 million Euros and ended in June 2014.

The purpose of the LARM project was to build a user driven research e-infrastructure providing access to the national Danish heritage of sound and radio. The technical backbone of the digital archive is the module based ICT platform CHAOS (Cultural Heritage Open Access System) and the user interface LARM.fm, which provides streaming access to archived sound as well as tools to search, annotate and collaborate with the sound files. Both are open source.

The main part of the radio programs stems from the national broadcasting company, Danmarks Radio. Programs from regional, community and commercial radio stations are also represented. The archive holds all Danish radio broadcast nationally in the period 1985 and onwards. The preserved radio broadcasts date back to the early 1930s, however, for radio broadcasts before 1985 only selected radio shows were recorded and kept for the future, and only a minor share of these are digitised.

3 Research questions

The digital archive faces two problems: first, the provided metadata is extremely sparse and because of lack of funding for internal in-depth indexing, collaborative enhancement of metadata is needed. However, due to copyright issues, large crowdsourcing activities are impossible, because access to the materials is restricted to registered university scholars and students in Denmark. Secondly, the user group is very heterogeneous and a simple metadata scheme would not be able to fulfil all users' needs for finding and working with audio material.

The metadata wish list between the various project researchers differs strongly. Two examples: Jacob Kreutzfeld, a researcher on urban soundscapes, investigates the relation between technology and the portrait of "the sound of the city" (Kreutzfeld, 2012). For his research, he needs a metadata scheme that includes a sound category (e.g. traffic, shop, car, bell, and goat), a category mix (sound level (fore/background), fade in/out), a category authenticity (real/scenography/effect) and a category anchorage (space, language, rhythm, and theme). In addition he needs a free field for annotations. The second example is Torben Sangild, a researcher on radio jingles (Stockfelt et al., 2012) who investigates changes of focus over time. He looks at theme songs (intro, separators and outro) and needs a metadata

scheme that includes these three categories as top-level categories and then further sub-categories on spots. He, too, requires an open annotation field.

As a result, the required metadata scheme for the digital archive LARM.fm would need to include the smallest common denominator of all user groups, but also take newly generated metadata into account and should facilitate researchers working with custom made coding manuals and metadata schemes. Researchers should be able to create their own metadata schemes, change existing metadata fields (i.e. alter wrong information or add information) and add free text in the form of annotations. Above all, ease of use and suitability to support research was the driving force behind the implementation of the metadata scheme.

Several user studies were conducted, which aimed at answering the following research questions:

1) What metadata scheme is needed for radio material?
2) Do users adopt and understand the chosen metadata schemes?
3) Does collaborative metadata enhancement work?

The present paper focuses on answering the second research question.

4 Questionnaire: Identification of user needs

In 2011, a questionnaire was distributed to 78 Danish and international audio culture researchers. The questionnaire consisted of 16 closed and open-ended questions and aimed at understanding the research interests, search behaviour and experiences, type of broadcast genres desired, relevant metadata types, level needed (part of or full broadcast; series of programmes) and use of the broadcast. In addition, an in-depth analysis of humanities research tasks and needs has been conducted, where participating researchers could describe the analytical concepts used and their coding wishes and practices. A detailed description can be found in Skov & Lykke (2012).

The participants indicated a number of research interests: radio, music, media, and cultural history, use of radio broadcast as a museum exhibition, web radio, radio jingles, language development, radio phonic idiom and cultural history (Skov & Lykke 2012). More particular, a number of information needs were identified that are related to audio materials. These can be classified as content and sound elements, as broadcast elements and radio production and structure.

The top seven search entries, marked as relevant or highly relevant and as such the most important metadata fields were broadcast channel, title, named person, genre, subject theme, abstract and topic. Participants considered audio media specific metadata as less relevant (this included recording technology, editing procedures, place of recording etc.).

The metadata provided by the Danish Radio broadcasting company describes radio programs as a whole entity. The participants were asked to indicate the desired indexing level and results indicated that 79% search for a specific radio programme. The results also showed that participants wished for the possibility that for their own purpose they could index on a more detailed analytical level of the radio shows.

5 Implemented metadata scheme

Based on this initial needs analysis, a metadata scheme has been implemented. The metadata scheme was built on top of the Dublin Core Metadata Element Set (Dublin Core Metadata Initiative, 2012). By using Dublin Core as foundation, data exchange with other cultural heritage repositories, such as Europeana, was simplified.

Three levels of metadata were developed: (1) core archival metadata, (2) LARM metadata, and (3) project-specific metadata. In addition, annotations were implemented. Administrative metadata were identified on all three levels. The three levels separated generic and research project specific metadata. The material providers deliver the core archival metadata. None of the users has the right to alter or delete this core archival metadata – even if it is wrong. The LARM metadata is user-generated, but is still of a more generic nature. Here users can indicate wrong information in the core archival data and propose corrections or add generic information, for example if a date or the name of the speaker is missing. In the LARM metadata, there are several fields available: Program Title, Description, Genre, Subject, Tags, Notes, Related Objects (such as links), and Cast. In Cast, users can add several persons each with subfields of: Name, Role Name, and Subject. Project-specific metadata schemes can be created by administrators on the request of users. They contain specific metadata fields that researchers expressed as necessary for their analysis. Annotations are considered to be part of the LARM metadata level. They can contain, for example, a transcript of a radio show or a researcher's thoughts on these. Users can create, alter or delete annotations, LARM metadata and project metadata. Again, archival metadata cannot be altered in any way.

Figure 1 shows the implementation of the three levels in the user interface. On the left side, the core archival metadata (Arkivmetadata) and the LARM metadata are listed. In addition, four project related metadata schemes have been created (LARM.radio.producent, LARM.radio.serie, Jingler and Lydkilde) as well as annotations. While Jingler (English: jingles) and Lydkilde (English: sound sources) are technically variations of the annotation field (here indicated by the same icon, that is the speech bubble), LARM.radio.producent and LARM.radio.serie are structured metadata fields.
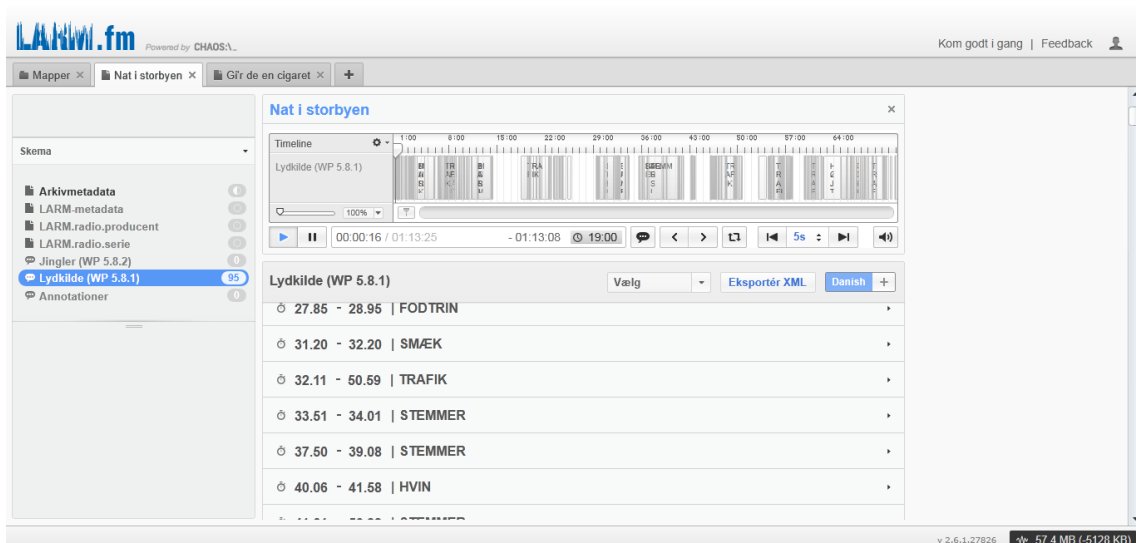

Figure 1. Screen shot of LARM.fm showing the levels of metadata schemes.

The example above shows the annotations to a radio show, which describe the choice of sound at a particular moment of the show. In that sense, the annotations are not free text, but resemble structured codes of a codebook know from qualitative research (Saldaña & Miles, 2013). The examples of Jingler and Lydkilde show that sometimes researchers chose to use the more adaptable annotations for their descriptions instead of falling back on creating project metadata schemes.

6 Online Interviews: Assessment of metadata schemes

In 2013, the project group decided to assess the metadata schemes and opted for synchronous online interviews with screen sharing. This approach allowed an exploration of user's attitudes and opinions towards the metadata schemes. Adobe Connect (http://www.adobe.com/products/adobeconnect.html) was used for the online interviews and youcanbook.me (https://gb.youcanbook.me/) for arranging interview times with participants. Invitations were sent out by choosing a random sample from all registered users of the interface LARM.fm and in addition by separately inviting a sub-set of the super users, who were funded by the project LARM. All participants received a headset and a camera as incentive. The interviews consisted of three parts: the first part included a set of questions about users' background, their expertise in working with audio files and their interest and experience in working with the interface LARM.fm. In the second part, participants were shown questions about their data usage behaviour on a screen and were asked to comment upon it. In the third part participants used screen sharing to show the interface LARM.fm on their device. All interviews were conducted in English. Figure 2 shows the research design setup.
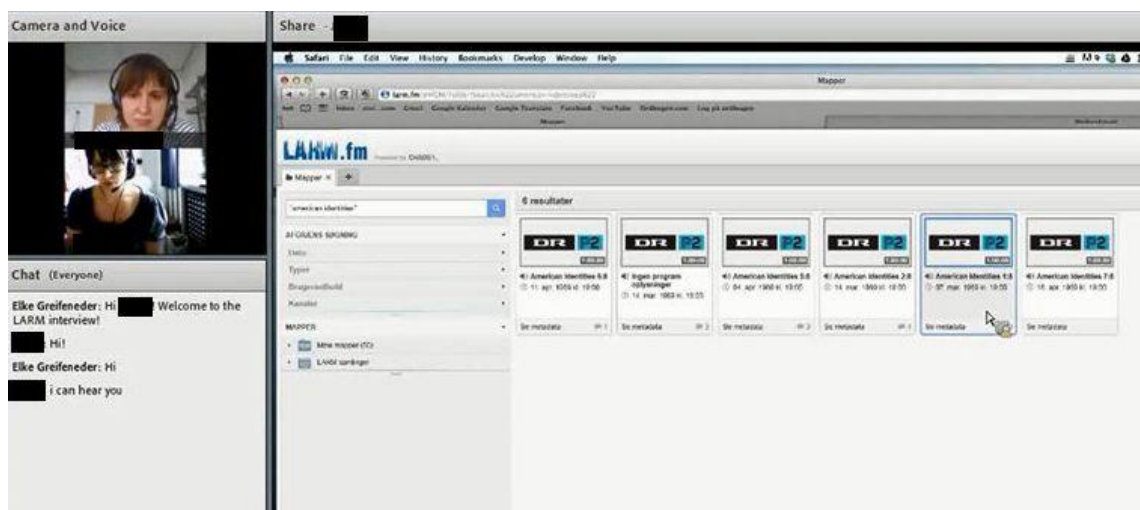

Figure 2. Research design setup for online interviews during screen-sharing.

Over a span of three months, 23 qualitative interviews were conducted which lasted 53 minutes on average. Fourteen participants had used the interface frequently, and 9 had used the interface only two or three times. A large majority of the participants were male (19 to 4 females), which matched the distribution of registered users in LARM.fm. A large group of participants were researchers including students on the bachelors and master level, PhD students, postdocs and professors. A smaller group of participants consisted of developers, administrative faculty and librarians. Discipline wise a wide range was represented spanning media studies, history, law, mathematics, Japanese studies, culture and

media studies, comparative literature and audio design. All names used below are randomly attributed, expressing only the gender.

The interviews showed that despite being involved in developing the metadata schemes, many participants used the current metadata schemes on a trial and error basis and did rarely understand the underlying philosophy. A few participants understood how it should work, but had trouble explaining it in their own words:

> *"The categories, well there's archive metadata which is just title and when it was broadcasted. Then there's LARM metadata. I'm not actually sure what that is. But I guess that would be if someone had tagged it. Then there's something called jingles, that would be project data but I don't know what it's doing out there. And also Lydkilde, I don't know why it's there. And then there is the annotation part which is obviously if anyone has annotated it."* (Peter)

> *[The metadata is] some kind of description. And annotations of course [...] the archive metadata is about the actual radio program and the LARM metadata seems like that is for the users to fill in."* (Jan)

Even the project members who attended the meetings in which the schemes were discussed had trouble recollecting the philosophy accurately.

> *"These three different layers or levels, it's not really clear what they do and why they are there. I know they exist because I sat in on the meetings but I've clearing forgotten."* (Walter)

This might be primarily the problem of a poor meeting, but it shows that the schemes are not intuitive enough to be remembered easily. For the occasional users, the differentiation was clearly more difficult to make as can be seen by the statement Chris made who described the archival metadata as *"the information that is in the archive that is searchable."*

Both sides, the occasional and the super user, are not entirely sure that they understand the underlying metadata philosophy, which means that on a long term perspective a different form of scheme needs to be found. Maybe it would already make a difference to change the terminology and make it clearer what each metadata scheme represents and can do. This would also help with the issue that participants expressed concern that the difference between annotations and metadata were not clear.

> *"I don't really know where, when something is an annotation and when it's metadata, I mean it's sort of. I guess that annotations are more, no I don't actually know that."* (Mike)

In general, participants showed during the screen sharing part that they used annotations to describe parts of a radio show and metadata to describe the radio show in its entirety. Despite the lack of understanding of the underlying philosophy, participants were indeed contributing metadata and annotations. There was a smaller group of the super users who said they were not contributing, but that was due to individual habits and not because of rejection of the schemes. One participant said that he made *"annotations on sheets because it doesn't really make sense here to do annotations"* (Leon) and another one considered the annotations as *"preliminary notes for myself in my own language"* (Mike). Some simply preferred to make

annotations on paper. One participant commented that the act of creating project metadata was simply too copious and that he therefore uses the more flexible annotations scheme for it.

The LARM project needed participant's input to enrich the archive – an activity that was clearly supported by some of the participants:

> *"You have these really really big digital archives like LARM. If you want to improve on them, you need to do some sort of crowdsourcing. You need to have everyone using it contributing, it is my opinion. "(*Anne)

As of September 2013 only 1.086 of 578.978 programs (that is 0.2 %) has been enriched with user-generated metadata. This number reflects that it was mostly the LARM project researchers who contributed to the archive. Since project funding ran out in June 2014, crowdsourcing of metadata enrichment may not be a sustainable solution. This can be supported by comments participants made on their willingness to contribute. For example, Mike said: *"I've just promised that I would do that, it is part of my job to feed the system with data"* or Chris who said that he "*didn't have the time to do that. But [he] will do that of course sometime*". There were also those like Gabriel who were not sure how his annotations might be useful to others: *"I haven't uploaded [the annotations] yet. I'm willing to do it but I can't see what people would use it for"*. Overall, a bad conscience of feeling that one has to contribute seems to overweight a willingness to contribute freely. The occasional users who were not part of the funded project did not contribute at all.

The metadata contributions and annotations users had made were full transcripts of radio shows, summary comments or time segmentations of radio shows, but also sometimes just one word. From the interviews, it seemed the biggest contributions users made were not additions to the archive, but corrections to it. Since users were not allowed to change the core archival metadata, they added, for example, a correct title to the LARM metadata, so that if another user searches for something, the file will be displayed in the search results, even if the title is still shown as the original archival title:

> *"LARM metadata is where I can write my comments and I have done that a lot of times. But the problem is I can't change the title of the file. I can only add my comments in the LARM metadata. And there I can write that the title is wrong and so on and the speaker is this but not this. But then the other users have to look under the LARM metadata and maybe they don't do it."* (Beate)

In many interviews, participants showed a surprisingly high accountability working with metadata. They reflected upon their actions, about the consequences of changing metadata and about the underlying question of user rights management in eScience environments. Several participants expressed the wish for a written policy on metadata changes.

> *"In this case I've actually enriched the archive by putting something that's not there in there. But it creates the problem of asking when is what I add better than what's already there? Do you see what I mean? It's kind of hard to describe but if what I add just overrules what was already there, what if I then add, what if what I had added was wrong?"* (Anne)

The low participation level in contributing to the archive is a deficiency on which future e-infrastructure projects will have to work if they want to succeed. But the high-level

accountability on users' side means that archive providers should worry less about misuse of metadata and instead embrace user participation.

7 Conclusion

This paper examined what metadata schemes are needed for radio material and if users will adopt and understand the chosen metadata schemes. Further the question of collaborative metadata enhancement was raised.

Based on several user studies, including a questionnaire sent to project members and potential users, the first research question was answered resulting in a metadata scheme, which consists of three layers. The first layer is the core archival metadata, which is delivered by the content providers and cannot be altered, except by administrators. The second layer is called the LARM metadata layer and contains metadata provided by LARM users. This metadata scheme can be altered by all users and describes entire files, for example a full 90 minutes radio show. The third layer is project-specific metadata, which can be created for users and is project specific such as for example metadata on sounds. In addition, annotations are offered as part of the metadata schemes, which have neither a restriction to user rights nor to content provided. Usually, participants expressed that they used annotations to describe parts of a file and metadata to describe the entire file.

In late 2013 follow-up online interviews assessed the adoption of the metadata schemes and users' understanding of the underlying philosophy. The results showed that some participants reported that they used the metadata schemes and the annotations, while others either preferred different ways of working with their data or felt that they had to contribute because of their involvement in the project. This also answers the question of the successfulness of crowdsourcing of metadata enhancement. Users contributed to the archive, in particular the users who were funded by the project, but a strong enthusiasm was lacking.

The metadata and annotations, which were provided by users ranged from one word contributions to full transcripts. Yet, most contributions were corrections to the archival metadata. In these contributions users showed a high degree of accountability, which is promising in regards to data misuse – an act, which did not happen at all in the present case.

In summary, developing the metadata scheme in close relation with future users helped getting a better metadata scheme, even if it has some flaws. In order to be successful with crowdsourcing, future e-infrastructure projects need to find more engaging ways to solicit user contributions. A consideration for future projects might be to focus the process of building the metadata scheme even stronger on how users describe data and not on how they search for data.

It became clear in all answers that there is no ideal metadata scheme for all users. If a platform wants to be more than just an archive with a search function, the least common denominator will not satisfy user's needs for e-infrastructures. On the other hand, trying to fulfil the various needs for such a diverse user group is impossible and will not result in a more user-friendly environment. In the end, users want to have access to material and will rely on tools they are accustomed to – most of them available outside of the e-infrastructure environment.

Acknowledgments

References

Ayres, M.-L. (2005), "Case Studies in implementing Functional Requirements for Bibliographic Records [FRBR]: AustLit and MusicAustralia", The Australian Library Journal, Vol. 54 No. 1, pp. 43–54.

Dublin Core Metadata Initiative (2012), "Dublin Core Metadata Element Set, Version 1.1", available at: http://dublincore.org/documents/dces/.

Gligorov, R. (2012), "User-generated Metadata in Audio-visual Collections", in Proceedings of the 21st International Conference Companion on World Wide Web, ACM, New York, NY, USA, pp. 139–144.

Lund, H.; Bogers, T.; Larsen, B & Marianne Lykke (2013), "CHAOS: User-driven Development of a Metadata Scheme for Radio Broadcast Archives", in Proceedings of the iConference 2013, pp. 990–994.

Huurnink, B., Hollink, L., van den Heuvel, W. and Rijke, M. de (2010), "Search behavior of media professionals at an audiovisual archive: A transaction log analysis", Journal of the American Society for Information Science and technology, Vol. 61 No. 6, pp. 1180–1197.

ISMIR – The International Society for Music Information Retrieval (2014), available at: http://www.ismir.net/.

Kanai, K. (2013), "Differentiation of Metadata Elements Based on Users' Music Information Needs: Focusing on FRBR, FRAD and Variations", LIBRARY AND INFORMATION SCIENCE No. 70, pp. 89–118.

Kim, J., Oard, D.W. and Soergel, D. (2003), "Searching Large Collections of Recorded Speech: A Preliminary Study", Proceedings of the American Society for Information Science and Technology, Vol. 40 No. 1, pp. 330–339.

Kreutzfeldt, J. (2012), "Street Cries and the urban refrain: A methodological investigation of street cries", Sound Effects, Vol. 1 No. 2, pp. 62–80.

LARM.fm blog (2014), available at: http://larm.blogs.ku.dk/about-larm/.

Lee, J.H. (2010), "Analysis of user needs and information features in natural language queries seeking music information", Journal of the American Society for Information Science and technology, Vol. 61 No. 5, pp. 1025–1045.

Stockfelt, O., Lønstrup, A. & Sangild, T. (2012), "Radiolab - three different approaches", SoundEffects - An Interdisciplinary Journal of Sound and Sound Experience, Vol. 2 No. 1.

Riley, J. (2011), "Leveraging the FRBR model for music discovery and data sharing: Autobiographical note", OCLC Systems & Services, Vol. 27 No. 3, pp. 175–189.

Saldaña, J. and Miles, M.B. (2013-2014), The coding manual for qualitative researchers + qualitative data analysis: A methods sourcebook, Sage Publications, London, Thousand Oaks, Calif.

Skov, M. and Lykke, M. (2012), "Unlocking Radio Broadcasts: User Needs in Sound Retrieval", in Proceedings of the 4th Information Interaction in Context Symposium, ACM, New York, USA, pp. 298–301.

Vainio, T., Väänänen-Vainio-Mattila, K., Kaakinen, A., Kärkkäinen, T. and Lehikoinen, J. (2009), "User Needs for Metadata Management in Mobile Multimedia Content Services", in Proceedings of the 6th International Conference on Mobile Technology, Application & Systems, ACM, New York, USA, pp. 51:1-51:8.