Backend Infrastructure Supporting Audio Augmented Reality and Storytelling

Kari Salo[1(✉)], Diana Giova[1], and Tommi Mikkonen[2]

[1] Helsinki Metropolia University of Applied Sciences, Helsinki, Finland
kari.salo@metropolia.fi, dianakaal@gmail.com
[2] Tampere University of Technology, Tampere, Finland
tommi.mikkonen@tut.fi

Abstract. Today, museums are looking for new ways to attract and engage audience. These include virtual exhibitions, augmented reality and 3D modelling based applications, and interactive digital storytelling. The target of all these activities is to provide better experiences for audiences that are very familiar with the digital world. In augmented reality (AR) and interactive digital storytelling (IDS) systems, visual presentation has been dominant. In contrast to this trend, we have chosen to concentrate on auditory presentation. A key element for this is a backend service supporting different client applications. This paper discusses our experiences from designing a portable open source based audio digital asset management system (ADAM), which supports interaction with smart phones and tablets containing audio augmented reality and audio story applications. We have successfully implemented ADAM system and evaluated it in the Museum of Technology in Helsinki, Finland.

Keywords: Digital asset management · Metadata · Open source DAM · Audio augmented reality · Soundscape

1 Introduction

Today, museums are looking for new ways to attract and engage audience. There are lots of activities in virtual exhibitions [1, 2], augmented reality and 3D modelling based applications [3, 4], and interactive digital storytelling [2, 5]. The common target of all these activities is to provide better experiences for audience that is very familiar with the digital world.

The Neighborhood Living Room sub-project, which is a part of Creative Europe funded People's Smart Sculpture Project, studies different methods to build a more dynamic and participatory audience relationship in a museum. The vision is that the Museum of Technology could be integrated as a part of the Arabianranta district community in Helsinki, Finland. The museum is also aiming at offering an emotional and participatory experience for the residents, especially youth.

When we look at the young residents and their natural way of using information technology (IT) it is obvious that somehow IT should be involved in exhibitions. Many of the residents have a smart phone, a tablet or both. In Finland 88 % of the age group

16–24 years old have used internet several times a day. 87 % are using mobile phones and 35 % tablets when accessing internet outside of home or office [6].

In augmented reality (AR) systems and interactive digital storytelling (IDS) systems visual presentation has been dominant. We decided to concentrate to auditory presentation as there is not that much activities ongoing. We also decided to use soundscapes as a way to augment reality in museum environment [7]. One key element when developing AR, soundscape and IDS systems is a backend service supporting different client applications. In our case, the backend service is an audio digital asset management system (ADAM).

In this paper we propose an open source based ADAM that is portable and also affordable for smaller museums and other culture sector actors and events. ADAM supports interaction with smart phones and tablets containing audio augmented reality and audio story applications, which are targeted to young visitors in culture institutions and events.

The rest of the paper is organized as follows. In Sect. 2, we discuss about related works and concepts. Then, in Sect. 3 we describe the overall system, and in Sect. 4 we derive corresponding system requirements. In Sect. 5, we present the implementation of the system, including the design of the core system and its rationale. In Sect. 6 we evaluate how well we succeeded. In Sect. 7, we will further discuss the results, and in Sect. 8, we draw some final conclusions.

2   Background

Audio augmented reality systems have been planned for different purposes, and the complexity of these systems has varied from special purpose equipment to smartphone. Audio-Haptic Navigation Environment (AHNE) is a physical space which contains virtual 3D-objects that can be searched and moved. To enable interactions, Kinect depth camera is used to track user's movement, and feedback is provided through audio and custom-implemented haptic gloves [8]. Augmented and Tangible Sonic Interaction (ATSI) does not require special space and user does not have to wear any device except headphones. The user may attach sounds to physical objects that already exist in a selected space, like living room. ATSI will follow user's head and hands movement using Kinect and maintain the correct spatial auditory perspective when picking and moving sounding objects [9]. Mobile Audio Augmented Reality System (MAARS) relies only on mobile device's orientation sensor when creating the impression of virtual sound sources being located in the physical space. The user will experience this virtual audio space using headphones [10]. Common to all of these systems is that they need to know user's location in order to produce sound that is modified according to user's movement. In our case the user is the active party. She knows her location related to the particular environment, where the sounds have a role as augmenting the reality and either she will with the help of mobile apps search relevant sounds or produce the acoustic environment using her creativity and imagination.

A soundscape can be a musical composition, a radio program or an acoustic environment [11]. In our case, it is the acoustic environment. Klang.Reise is an installation of video and audio recordings inside a closed spherical space [7]. The goal

is to demonstrate different sounds of a selected place and how these sounds change over the time. The Sound Design Accelerator (SoDA) project provides software for soundscape generation. Targeted to sound designers [12], SoDA contains a storage of annotated audio files, a semantic search engine, an automated soundscape composer combining interpretation of semantical analysis, a geographical and acoustic space modeler, user defined semantic filters, and a soundscape synthesis engine. Our approach is somewhere in between these two. Our target audience is museum visitors, like in Klang.Reise. However, in addition of fixed soundscapes we are asking visitors to build simple soundscapes. We do not expect visitors being familiar with reverberation, resonance, acoustic absorption, bit depth and other acoustic terms.

Interactive storytelling systems could be a mix of human produced stories and digitally produced interaction or fully digital environments with user participation. A fully digital prototype consisting of a game engine based narrative environment, an AI-based interactive digital storytelling system and communication protocol, and a language between these two components was developed. The user will interact through IDS, and narrative environment will reflect the interaction [13]. Sarajevo Survival Tools contains a digital story in the form of video. The video is divided into story segments. After each segment, the user is given a possibility to browse segment related objects and material in a virtual museum [2]. A prototype containing animation based user interaction, recorded actor readings, improvisations, and expert commentaries was developed for a museum exhibition on the medieval historian Jean Froissart. Metadata was attached to recorded audio clips to facilitate relevant audio file retrieval and play based on user interaction and location [5]. Our target is to rely on non-modifiable audio material on storytelling; thus the last two examples are closest to our approach. However, we are also prepared to save user's narratives, so that visitors should be able share their own stories.

Digital Asset Management System (DAM) should be able to manipulate as well as protect from unintentional alteration those digital assets stored in it. Digital asset could be defined as a file which is tagged with the information about it. This definition – an asset is a file plus metadata – is commonly used by large companies [14]. The second definition is that an asset is a file and its rights, which essentially means that content has value as long as its owner has the right to utilize it [15]. The two definitions are complementary. Thus a DAM should contain digital files along with their metadata as well as usage rights. The management aspect of the system is fairly straightforward to understand as it is the actions which are required to be executed onto these assets. The aforementioned assets could be adding, removing, and editing assets' data or metadata. This in turn would ensure that the digital integrity of the asset is maintained while providing re-purposed files for every media need.

To utilize, search, and find relevant media files, it essential to utilize metadata. There are several metadata standards for different purposes, like metadata exchange between systems [16], general metadata for broad range of domains [17], audio and video resources for a wide range of broadcasting applications [18, 19], series of interfaces for interchange information about multimedia content in the audio domain [20], and audio specific structural and administrative metadata [21]. Different standards have been evaluated from digital audio-visual library point of view using four selection criteria [15]. There are several studies regarding audio specific metadata and

annotation, like Telemeta [22] and GlobalMusic2one [23]. Telemeta is very similar to what we aim at. The differences lie in the target group and related implications. Telemeta is targeted to music researchers and thus it relies on TimeSide audio processing framework, audio analysis plugins and audio extraction libraries. GlobalMusic2one annotation tool has the same target group as Telemeta, i.e. musicologists. Thus the tool aims at describing audio file in very detailed level. In contrast, our view is that individual audio file is a whole.

3   System Overview

Figure 1 describes the overall system including also mobile applications that will utilize audio digital asset management system. The overall system is a distributed system consisting of an audio digital asset management system (ADAM), management application, and mobile applications. ADAM provides functionalities to manage assets and offers interfaces for both for management application and mobile applications over internet. The management application is more or less an administration console to manage assets and users. Mobile applications, which we will not address in this paper, are for example audio augmented reality, soundscape design, audio story recording and listening, or audio memory sharing applications.
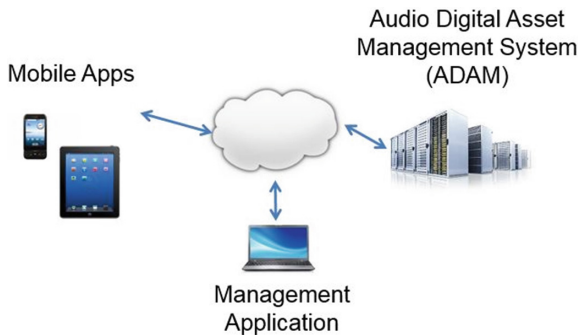
Fig. 1. System overview

4   Requirements

In May 2015, we decided that the first version the backend system should be up and running in Museum of Technology's server within 4–5 months. This meant that we could not start from scratch and that server requirements could not be too limiting. In addition, the system should be available also for other parties without license fees, and hence we would build our system using open source components. At functional level it was required that the system should provide the following functionalities:

- create, read, update and delete audio and related metadata content;
- search content based on metadata;

- manage access groups and rights;
- authenticate users;
- provide easy to use admin console for non-IT personnel;
- provide APIs: authenticate, search content, download content, and upload content.

Metadata requirements were left partially for further studies. Initially, we defined the following metadata requirements:

- Title of the audio file;
- Description of the audio file;
- Tags (keywords) assigned to audio file;
- Date when audio file was saved;
- Link to audio file;
- Format of the audio file: wav, mp3, and PCM;
- Length of the audio file in seconds;
- Category: nature, human, machine, and story;
- Sound type: soundscape, ambience, and effect;
- Location: longitude and latitude.

5   Implementation

Comparison of Digital Asset Management Systems. There are many different kinds DAMs. Some might be specifically designed for a specific file type, while others might be designed to handle a multitude of file types. In this case, the comparison is regarding systems which can accommodate audio files. Table 1 shows some of the open source DAMs along with their features, license and programming language(s). Based on the table, we determined that for us the best candidates would be Telemeta and Resource Space. As described in Sect. 2, Telemeta is mainly targeted to musicologists. In addition, Telemeta was still in development mode and it was not compatible with Windows systems. Consequently, we ended up using ResourceSpace as the basis for ADAM. This selection was also supported by the fact that we have a large pool of students with PHP skills.

Storage of Resources. Digital assets, called resources, are stored according to their identities in a table. The metadata values are stored into the row of that particular resource, and the metadata fields are stored in a separate table. This design simplifies modifying metadata fields and values. Moreover, there are further tables that are used to specify which collection resources are in, who has the rights to these resources, how many users there are, various reports, how many users have accessed the APIs, etc. In general, the tables are informative, which simplifies the programming. For example, relevant data can be fetched by joining or pivoting tables with MySQL. In ADAM resources are based on their identities. If a user sends an audio file into another user's collection, then only a record of that action is saved into the database – the files are not duplicated or physically copied on the server space.

Table 1. Detailed overview of DAMs

Software name	Features	License	Language
Telemeta	• open source • web audio archiving software • metadata • user management • English and French support • REST API • Dublin Core compatibility	CeCILL	Python and JavaScript
ResourceSpace	• fully featured DAM system • user management • API available • plugins available • metadata • many file types	BSD	PHP and SQL
Phrasenet	• DAM system • user management • images/video/documents support	GPL3	PHP
EnterMedia	• typical DAM system • uses XML, but database possible • plugins available • metadata	LGPL	Java

API Design. According to the requirements listed above, three APIs were needed: authentication API, upload API, and search API. Developing APIs was the most time consuming part of modifying ResourceSpace to fulfill the requirements. The authentication API is needed by the mobile users to receive a token, which in turn will be used with search and upload APIs. Authentication is security feature and ensures that only authorized users have access to token. The search API is a HTTP GET request containing token and predefined search parameters. The response given in JSON format contains audio files' metadata based on search parameters that are set along with the search request. Thus, the search API also enables downloading, as the link to audio file is a part of the metadata. The upload API lets users with a valid token upload their audio files along with metadata they choose to transmit to ADAM as a multi-part form using HTTP post. Metadata and token will be encoded in a part of the URL, and the audio file in the body of HTTP post.

Metadata. During the requirement specification phase, it was decided that we need to study metadata standards more deeply. Based on metadata evaluation criteria introduced in [23] we defined our criteria as following:

• Internal metadata model now, readiness for external model;
• Flat metadata model;
• Support identification, description, technical and rights types;
• Syntax of supported metadata.

Based on the criteria, we ended up adding more metadata fields to enable in the future the exchange of assets by supporting Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), which requires compatibility with unqualified Dublin Core [15, 16]. In addition, it was clear that we cannot call our audio files as assets unless we introduce at least a rights metadata field. Thus our final metadata is as described in Table 2. In total, 7 new metadata fields were introduced. Most of the metadata will be input manually during the storage of audio file, but some will however be extracted automatically from the audio file properties.

Table 2. The Dublin Core Metadata Element Set mapping to ADAM

Dublin Core Metadata Element Set	AudioResourceSpace	Notes
Title	Title	
Contributor	Contributor	new
Source	Source	new
Creator	Creator	new
Date	Creation Date	automatic
Language	Language	new
Subject	Tags	
Type	Category + Sound Type	
Relation	Relation	new
Description	Description	
Format	File extension	automatic
Coverage	Location (longitude + latitude)	
Publisher	Publisher	new
Identifier	Link	
Rights	Rights	new
	Length	
	File Size	automatic

Management Application. The management application that a user of ADAM sees once she logs into the system, is straightforward to understand and use. Figure 2 shows the modified dashboard suited to serve the requirements for the audio storage system with some notes regarding its functionality. Some changes had to be made to accommodate only audio files, necessary user rights, and custom metadata fields. These changes were made by accessing the Team Centre menu (the top navigation bar in Fig. 2), which provides options for modifying user rights, file types, and metadata fields.

Installation Requirements. ADAM can be installed on Linux/Unix, Windows, Mac OS X, and Synology DSM systems and it works with the most web servers including Apache and Internet Information Services (IIS) for Windows Server. ADAM requires PHP greater than or equal to version 5, and MySQL greater than or equal to version 5.0.15.

Fig. 2. ADAM dashboard

6 Evaluation

We produced five separate user guides: installation guide, admin guide, and one guide per API. Based on this documentation, we were able to evaluate how we initially succeeded in ADAM implementation.

The installation guide was given to a person who was not developing ADAM. He was able to first check together with IT staff in Museum of Technology that a compatible server environment was available. Then he followed the installation guide and successfully installed the system. After running basic functionality tests defined in the installation guide, it was clear that the installation succeeded.

Two persons from museum were given a short introduction regarding ADAM main concepts. They were given the admin guide to check if the management of audio files and users is easy enough. Their response was that the admin guide is sufficient to handle the tasks. The same guide was also handed to sound design students who were planning what kind of audio files can be used as building elements when designing soundscapes in museum environment. The feedback from the students was mainly positive. They were able to accomplish their task to store soundscape building elements with metadata. The only drawback was the time consuming manual input of metadata.

All the APIs were initially tested by a person who was not developing them. Authentication API was tested using Chrome Advanced REST Client. HTTP post containing the username and password were sent to ADAM, request succeeded and response contained token. Search API was tested also using Chrome Advanced REST Client. Defining different search combinations resulted into right response. Upload API was tested using simple Android app which was developed for testing purposes. Uploading an audio file together with relevant metadata worked as expected.

API testing continued with the group of students who were developing Android apps which would be accessing ADAM. The feedback received from this group was that they wished the search API would output an empty array in JSON if the results for search were null, which was changed accordingly. Otherwise all APIs were working as defined in the respective guides.

To summarize, we have succeeded in providing the backend system as expected (Table 3). The only drawback was the two weeks delay in installing the final system in the museum's server.

Table 3. Satisfying the requirements

General Requirements	Yes/No
Backend system working in museum within 4–5 months	2 weeks delay
Open source system	Yes
Functional Requirements	
create, read, update and delete audio and related metadata content	Yes
search content based on metadata	Yes
manage access groups and rights	Yes
authenticate users	Yes
provide easy to use admin console	Yes
provide the following APIs for mobile applications: authenticate, search content, download content, and upload content	Yes
Metadata Requirements	See Table 2

7   Discussion

Initially ADAM was installed and developed in Metropolia's (development organization's) server. As we wanted to prove the portability the system was installed in the Museum of Techology's server, and this version was used to get feedback from the different stakeholders. Two (none-IT) persons from the museum took the role administrator and tested usability and functionality of ADAM utilizing solely the respective user guide. We involved sound designers to evaluate system as a part of planning audio material for soundscape workshops. We also involved 7 Android application development teams to utilize ADAM as a part of their audio mixing apps. Based on the feedback from different parties ADAM was modified accordingly. So far the only feature that requires more attention is the metadata input. We need to study more which metadata is really required and reconsider if OAI-PMH support is relevant.

We will utilize ADAM at least in the following cases together with the Museum of Technology:

- audio augmented reality/soundscape workshops,
- audio stories connected to museum's artefacts,
- sharing memories in the form audio stories

We strongly believe that ADAM is a viable innovation platform also for smaller museums and other culture sector's actors who run on a tight budget and at the same time want to utilize audio as a part of their creative activities. One proof of this claim is that already now three of the People's Smart Sculpture partners have expressed their willingness to utilize the platform.

8 Conclusions

In this paper we have proposed an open source based audio digital asset management system that is portable and also affordable for smaller museums and other culture sector's actors and events. Audio digital asset management system supports interaction with smart phones and tablets containing audio augmented reality and audio story applications, which are targeted to young visitors in cultural institutions and events.

We have successfully implemented ADAM system and evaluated it in the Museum of Technology in Helsinki, Finland. Next, we are ready test different use cases where the backend system is heavily utilized for building innovative apps.

References

1. Ardissono, L., Kuflik, T., Petrelli, D.: Personalization in cultural heritage: the road travelled and the one ahead. User Model. User-Adap. Inter. **22**(1–2), 73–99 (2012)
2. Rizvic, S., Sadzak, A., Hulusic, V., Karahasanovic, A.: Interactive digital storytelling in the Sarajevo survival tools virtual environment. In: Proceedings of the 28th Spring Conference on Computer Graphics, SCCG 2012, pp. 109–116. ACM, New York (2012)
3. Demiris, A.M., Vlahakis, V., Ioannidis, N.: System and infrastructure considerations for the successful introduction of augmented reality guides in cultural heritage sites. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST 2006, pp. 141–144. ACM, New York (2006)
4. De Sa, M., Churchill, E.: Mobile augmented reality: exploring design and prototyping techniques. In: Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2012, pp. 221–230. ACM, New York (2012)
5. Blythe, M., McCarthy, J., Wright, P., Petrelli, D.: History and experience: storytelling and interaction design. In: Proceedings of the 25th BCS Conference on Human-Computer Interaction, BCS-HCI 2011, pp. 395–404. British Computer Society, Swinton (2011)
6. Tilastokeskus, Internetin käytön yleiset muutokset (in Finnish). http://www.stat.fi/til/sutivi/2014/sutivi_2014_2014-11-06_kat_001_fi.html
7. Drechsler, A., Raffaseder, H., Rubisch, B.: Klang.Reise: new scientific approaches through an artistic soundscape installation? In: Proceedings of the 7th Audio Mostly Conference: A Conference on Interaction with Sound, AM 2012, New York, NY, USA, pp. 44–46 (2012)
8. Väänänen-Vainio-Mattila, K., Suhonen, K., Laaksonen, J., Kildal, J., Tahiroğlu, K.: User experience and usage scenarios of audio-tactile interaction with virtual objects in a physical environment. In: Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces, DPPI 2013, pp. 67–76. ACM, New York (2013)
9. Pugliese, R., Politis, A., Takala, T.: ATSI: augmented and tangible sonic interaction. In: Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction, TEI 2015, pp. 97–104. ACM, New York (2015)

10. Heller, F., Krämer, A., Borchers, J.: Simplifying orientation measurement for mobile audio augmented reality applications. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2014, pp. 615–624. ACM, New York (2014)
11. Schaefer, R.M.: The Soundscape: Our Sonic Environment and the Tuning of the World. Inner Traditions International/Destiny Books, Rochester (1993)
12. Casu, M., Koutsomichalis, M., Valle, A.: Imaginary soundscapes: the SoDA project. In: Proceedings of the 9th Audio Mostly: A Conference on Interaction with Sound, AM 2014. ACM, New York (2014). Article no. 5
13. Peinado, F., Navarro, Á., Gervás, P.: A testbed environment for interactive storytellers. In: Proceedings of the 2nd International Conference on INtelligent TEchnologies for Interactive Entertainment, INTETAIN 2008. ICST, Brussels (2008). Article no. 3
14. Jacobsen, J., Schlenker, T., Edwards, L.: Implementing a Digital Asset Management System: For Animation, Computer Games, and Web Development. Focal Press, Burlington (2005)
15. De Sutter, R., Notebaert, S., Van de Walle, R.: Evaluation of metadata standards in the context of digital audio-visual libraries. In: Gonzalo, J., Thanos, C., Verdejo, M., Carrasco, R.C. (eds.) ECDL 2006. LNCS, vol. 4172, pp. 220–231. Springer, Heidelberg (2006)
16. Open Archives Initiative Protocol for Metadata Harvesting. https://www.openarchives.org/pmh/
17. Dublin Core Metadata Element Set, Version 1.1. http://dublincore.org/documents/dces/
18. AES60-2011: AES standard for audio metadata - Core audio metadata. http://www.aes.org/publications/standards/search.cfm?docID=85
19. Ebu Core Metadata Set, Specification v. 1.6. https://tech.ebu.ch/docs/tech/tech3293.pdf
20. ISO/IEC 15938-4:2002(en) Information technology — Multimedia content description interface — Part 4: Audio. https://www.iso.org/obp/ui/#iso:std:iso-iec:15938:-4:ed-1:v1:en
21. AES57-2011: AES standard for audio metadata - Audio object structures for preservation and restoration. http://www.aes.org/publications/standards/search.cfm?docID=84
22. Fillon, T., Simonnot, J., Mifune, M., Khoury, S., Pellerin, G., Le Coz, M.: Telemeta: an open-source web framework for ethnomusicological audio archives management and automatic analysis. In: Proceedings of the 1st International Workshop on Digital Libraries for Musicology, DLfM 2014, pp. 1–8. ACM, New York (2014)
23. Woitek, P., Bräuer, P., Grossmann, H.: A novel tool for capturing conceptualized audio annotations. In: Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound, AM 2010. ACM, New York (2010). Article no. 15