Audio Commons: bringing Creative Commons audio content to the creative industries

Frederic Font[1], Tim Brookes[2], George Fazekas[3], Martin Guerber[4], Amaury La Burthe[5], David Plans[6], Mark D. Plumbley[7], Meir Shaashua[8], Wenwu Wang[7], and Xavier Serra[1]

[1]*Music Technology Group, Universitat Pompeu Fabra, Spain*
[2]*Institute of Sound Recording, University of Surrey, United Kingdom*
[3]*Centre for Digital Music, Queen Mary University of London, United Kingdom*
[4]*Jamendo SA, Luxembourg*
[5]*AudioGaming, France*
[6]*Business School, University of Surrey, United Kingdom*
[7]*Centre for Vision, Speech and Signal Processing, University of Surrey, United Kingdom*
[8]*Waves Audio LTD, Israel*

Correspondence should be addressed to Frederic Font (`frederic.font@upf.edu`)

ABSTRACT

Significant amounts of user-generated audio content, such as sound effects, musical samples and music pieces, are uploaded to online repositories and made available under open licenses. Moreover, a constantly increasing amount of multimedia content, originally released with traditional licenses, is becoming public domain as its copyright expires. Nevertheless, the creative industries are not yet using much of all this content in their media productions. There is still a lack of familiarity and understanding of the legal context of all this open content, but there are also problems related with its accessibility. A big percentage of this content remains unreachable either because it is not published online or because it is not well organised and annotated. In this paper we present the Audio Commons Initiative, which is aimed at promoting the use of open audio content and at developing technologies with which to support the ecosystem composed by content repositories, production tools and users. These technologies should enable the reuse of this audio material, facilitating its integration in the production workflows used by the creative industries. This is a position paper in which we describe the core ideas behind this initiative and outline the ways in which we plan to address the challenges it poses.

1 INTRODUCTION

The democratisation of multimedia content creation and the social media revolution have changed the way in which multimedia content is created, shared and (re)used [1]. Online multimedia sharing has experienced a huge growth in the last decade, yielding significant amounts of user-generated multimedia resources, a big part of them available under open licenses [2, 3]. Furthermore, significant amounts of professionally created multimedia content released with closed licenses, are now becoming public domain as their copyright expires, but nevertheless remain unused. At the same

time, media industries (content users) need to reduce production costs in order to remain competitive. In the particular case of sound and music, an increasing amount of audio material is available and released under Creative Commons[1] licenses, both coming from amateur and professional content creators. For example, Jamendo[2] is a music sharing site which contains more than 470k music pieces released under Creative Commons licenses, and Freesound[3] [7] is a sound sharing site with more than 300k sound samples (including sound effects, instrument samples and field-recordings among others) also released under Creative Commons licenses. We refer to all this audio content plus any other content that could potentially be made available as public domain or under Creative Commons licenses as the *audio commons*. Using Creative Commons copyright licenses, content reuse possibilities are much broader than in traditional copyright models. The Creative Commons modular nature of its licenses provides a clear and powerful framework in which content creators can specify the rights for reusing their published content. For instance, a content creator might allow the reuse of its content provided that the content user attributes the source (CC-BY), or it might choose to allow only non-commercial uses (CC-BY-NC). A single piece of content can also be licensed to different content users under different terms, or be re-licensed to a particular content user. There is, therefore, an opportunity for media industries to incorporate audio commons content in their productions, but this is not yet an extended practice.

There are a number of reasons why such content is not yet extensively used in the professional sector. A major one is the lack of a shared culture and understanding within the creative industries of open content and its potential use. But there are also technical and practical issues that do not facilitate this usage. Despite the amount of audio commons content available in online repositories, such as those mentioned above, most potentially useful content remains scattered around the web (if available at all) and typically not properly labeled with specific licenses or reachable through search engines. Also, the nature of audio commons content, coming from a variety of sources and from authors with different levels of expertise, results in unstructured (or not uniformly structured) mass of content. As opposed

to, for example, professionally crafted sound libraries which have been manually curated and whose content has been carefully annotated, audio commons content is often unorganised, limiting its potential retrieval and reuse possibilities. Moreover, no tools are easily available to search and incorporate audio commons content in the production workflows of content users in the context of the creative industries.

We see these reasons as challenges that must be addressed in order to make the use of audio commons content an extended practice in the creative industries. In this paper we describe the vision of a technology supported ecosystem of content, users and tools that should facilitate the reuse of audio commons content in the professional sector and, by extension, in the independent or non-professional sectors too. We call this the *Audio Commons Ecosystem* (ACE), and its initial development is being funded as a research and innovation project grant. We propose a number of development actions to enable access and retrieval of audio commons content in innovative ways that fit the requirements of different use cases of media creation, benefiting content users by reducing production costs and benefiting content creators by exposing their works to professional environments and allowing them to licence their content. The proposed technologies fill in the existing gaps in order to make the ACE possible, from metadata specification to content annotation, content analysis, gathering of user feedback and licensing procedures. But perhaps more importantly, besides the particular research and innovation actions that will be carried out during the lifetime of the project and the initial development of the ACE, we frame this ecosystem in the broader context of what we call the Audio Commons Initiative[4]. The aim of the Audio Commons Initiative is, starting from the project consortium, engage as many stakeholders as possible, involving them in the definition, development and expansion of the ecosystem. We expect that the concept of Audio Commons and the ACE will have a big impact on the creative industries including, but not limited to, video games production, music production, sound design and audio-visual production.

The rest of this paper is organised as follows. In Sec. 2 we expose the main objectives of the Audio Commons Initiative and how it will be started as a funded project. In Sec. 3 we describe in more detail the core aspects of the Audio Commons Ecosystem. We end this paper in

---

[1]http://www.creativecommons.org
[2]http://www.jamendo.com
[3]http://www.freesound.org

[4]http://www.audiocommons.org

Sec. 4 with some general conclusions about the Audio Commons Initiative and the ACE.

2 THE AUDIO COMMONS INITIATIVE

The Audio Commons Initiative is the result of the realisation that audio commons content has a huge potential for being reused by the media industry and that is not yet being properly exploited. The main objectives that we envision are the following:

- Promote the publication of audio content as audio commons and foster its reuse by the creative industries, breaking the existing barriers between content users and content creators (either amateur or professional), and proposing a model for audio consumption in media production better aligned with the current trends in the digital economy.

- Develop technologies to support the ACE, making them open and widely available. These technologies will enable and promote the interaction between content creators, content providers, content users, tool developers and further potential actors, facilitating the publishing, licensing and consuming of audio commons content.

- Develop innovative technologies to automatically describe sound and music recordings in the context of search and retrieval applications. These technologies will also be made available under open licences and they will enhance the value of the audio content within media production applications.

- Bootstrap an Audio Commons Ecosystem, publishing a significant amount of audio commons content through it and building tools that can consume the content and be embedded in existing creative workflows of the industry.

- Define open standard procedures for new stakeholders to join and participate in the ACE, fostering its growth, maintenance and sustainability. These procedures will define how to produce and consume content within the ACE.

We are working towards these objectives with the support of the funded project *AudioCommons*, which runs for three years from February 2016. Besides the development of the particular tools and technologies of the ACE (which are described in more detail in Sec. 3), the project will address intellectual property and business-related challenges that emerge from the initiative. We will evaluate complex reuse scenarios using existing Creative Commons licenses, and make recommendations about their usage, putting a special emphasis on clarifying (re)licensing procedures for those pieces of content that require it (e.g., from CC-BY-NC to CC-BY). We will also propose business models that can arise from the concept of Audio Commons and make the ACE self-sustainable.

3 THE AUDIO COMMONS ECOSYSTEM

The main goal in developing the Audio Commons Ecosystem is to allow content creators to expose their audio commons content to the content users in the creative industries, and provide ways for its seamless integration in creative workflows. In Fig. 1 we show a conceptual diagram of the different components that will be interconnected in our vision of the ACE. As it can be seen in the diagram, audio content created by the content creators will be exposed in the ACE through a number of content providers. Content providers will host that content and publish it following a common metadata specification as defined by an audio ontology (the Audio Commons ontology), and make it available to content users through embeddable tools (e.g., audio plug-ins) developed by tool developers. Content providers will be able to communicate with the embeddable tools through a unified API specification (the Audio Commons API specification) which will not only manage information about audio content and retrieval but also will handle licensing procedures between content creators and content users. Because of the nature of Creative Commons licenses, content users will also be able to become content creators by feeding back derivative or new content to the ecosystem. The procedures for publishing and consuming content, as well as the tools for annotating content and the ontology and API specifications, will be open and made available to the public and to potential new actors willing to join the ecosystem.

Due to the variety, amount and availability of the content potentially published in the ACE, it can become an incredibly valuable resource for many multimedia applications. To name a few examples, the ACE can be useful as a source of content for procedural audio and soundscape generation in video games [4, 5], as a
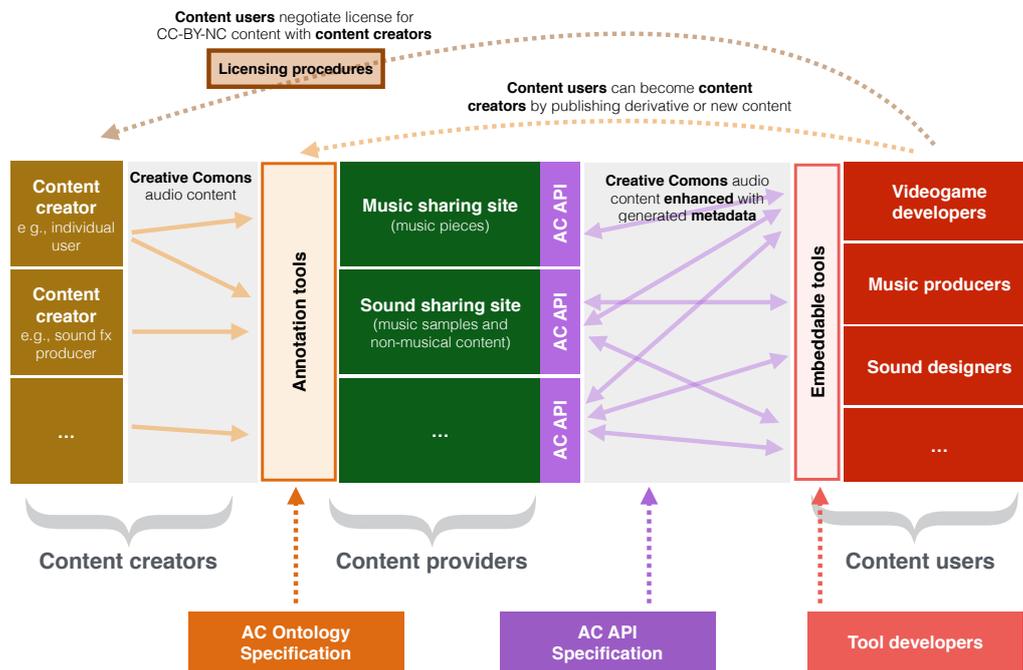
Fig. 1: Audio Commons Ecosystem conceptual diagram

source of music tracks for tv commercials, as a library of instrument samples and music loops for samplers and digital audio workstations, and as a resource of sound effects for video animations or other audiovisual productions. In the following subsections we describe in more detail some of the most important ideas behind the design of the Audio Commons Ecosystem.

3.1 Use of Creative Commons licenses

The amount of Creative Commons licensed works (of all kinds) available online has recently reached the significant landmark of 1 billion [6]. Part of it includes audio content such as sound effects, soundscapes, speech recordings, music samples and music pieces, which is hosted in online platforms such as the aforementioned Freesound and Jamendo. There also exist initiatives like Europeana[5] which aggregate content from national libraries and institutions and make it available to the general public (an important part of it being also released under Creative Commons or other kinds of open licenses), and other popular sites such as the Internet

Archive[6], Soundcloud[7] or ccMixter[8], that host content which already is or could be potentially released under Creative Commons licenses and hence be published in the ACE. Furthermore, there exist even more specific sources of Creative Commons audio material such as the Open Multitrack Testbed [8] and MedleyDB [9], academic databases for multitrack recordings. The ACE will be initially populated with at least 770k audio resources including music pieces from Jamendo and sound samples from Freesound (both acting as content providers and run by members of the Audio-Commons consortium). We expect them to serve as an example to other potential content providers for joining the ecosystem, which will ultimately motivate new content creators to publish their creations in the ACE.

3.2 Circular interaction model and co-creation

The same *open* sharing concepts represented by the Creative Commons licenses are applied to the whole design of the ACE. The ecosystem we propose allows content users to consume reusable content generated

---

[5]http://www.europeana.eu

[6]http://www.archive.org
[7]http://www.soundcloud.com
[8]http://www.ccmixter.org

by content creators and hosted by content providers. However, it also allows content users to easily become content creators and feed the ACE back with new material or derivatives of existing material. In this way, the Audio Commons Ecosystem also fosters co-creation of audio resources, through a model in which everyone can seamlessly contribute and benefit. The circular model will be supported by the tools and technologies of the ACE, allowing to both consume and contribute within the same environment.

3.3 Enhanced audio content

In the core of the Audio Commons vision, it lays the idea that audio content is much more that just audio files. The audio commons content will potentially come from very different sources and, particularly in the case of user-generated content, will typically lack a uniform structure to allow proper access and retrieval. The ACE targets different kinds of reusable audio content such as music pieces (i.e., audio recordings corresponding to complete songs), music samples (i.e., individual music elements such as single notes, percussive hits, chords, melodies and loops), sound effects (i.e., non-musical sound events such as foley, footsteps, opening and closing doors, alarm sounds, cars passing by, animals and all kinds of noises or artificially created glitches) and soundscapes (i.e., complex non-musical recordings such as environmental recordings, street ambiances or artificially constructed sonic environments).

There has been a lot of research in the sound and music information processing field to develop algorithms for automatically describing audio content. Existing descriptors are however limited and incomplete. Common audio features [10] allow us to compute characteristics of timbre, temporal evolution, tonality and loudness of audio pieces. However the majority of these features are very low-level without describing semantically relevant information [11]. In the case of music signals, specific algorithms try to address this issue and provide mid/high-level annotation for properties such as chord and key recognition, tempo/rhythm/meter characterisation, pitch estimation and multi-pitch estimation, melody extraction, instrument recognition, and music structure analysis [12, 13]. Research has also been carried out in the computational analysis of timbre perceptual qualities of audio recordings, oriented to sound samples rather than music compositions [14, 15, 16]. Furthermore, research on computing high-level semantic features from low-level audio descriptors exists. In

particular, in the context of classification problems, music genre [17], and artist identification [18] have gathered much research attention. Besides music classification, research has been also conducted for the classification of other audio content such as music samples [19] or non musical sounds [20].

To be able to successfully organise and represent audio commons content in the ACE, we will carry out a number of development actions to improve the state-of-the-art in sound and music description and semantic representation technologies. On the one side, we will focus our research on aspects that have been usually overlooked in existing literature (such as the development of descriptors targeted to short music samples). On the other side, we will stress the development of reliable high-level semantic descriptors with the use of bigger and crowd-sourced datasets. The combination of semantic representation of audio content with the use of an ontology (i.e., the Audio Commons Ontology) and the development of annotation tools (both manual and automatic) will enable searching and browsing audio commons content in unprecedented and innovative ways by, for example, browsing sound effects on the basis of their timbral characteristics. The tools that will be developed for content annotation will, in essence, enhance the content by providing extra metadata with a uniform representation compatible with the ACE.

3.4 Distributed framework

The Audio Commons Ecosystem is designed for incorporating many different agents that will be able to connect to the ecosystem for publishing, consuming, licensing or working with content. For this reason, we propose a distributed architecture in which an indeterminate number of content providers process and host resources created by the content creators, and offer API endpoints which follow a common API specification (i.e., the Audio Commons API specification). The tools embedded in production workflows will then access the different content providers by implementing a single API client. Within the AudioCommons project we will evaluate the ecosystem with initial content providers and "early adopters" of the technology (Jamendo and Freesound), and release embeddable tools that will allow the access to that content within production environments. The first embeddable tools for the ACE will include two audio plugins, an add-on for a well-known open source Digital Audio Workstation, and a

web interface for accessing music pieces in the ACE. Nevertheless, the aim of the Audio Commons Initiative is to expand the ecosystem and attract new actors. For this we will disseminate clear and standard procedures which will describe, to potential new actors of the ACE, how to create new embeddable tools for accessing (and publishing) content, how to become a content provider of the ACE, or how to interact with the ACE in any of the other ways enabled by our model (e.g., licensing content, providing metadata, etc.).

4 CONCLUSIONS

In this paper we have described the core ideas behind the Audio Commons Initiative and outlined the development of the Audio Commons Ecosystem. The ACE aims at directly connecting content creators and content users. We believe that closing the gap between content users' needs and the available audio commons resources generated by content creators will be beneficial to both parties, having an impact on the digital value chain that connects content creators and content users and which currently follows a traditional model that has not yet been adapted to the modern communication society. Our proposal wants to promote open data models in which reuse of distributed content takes a central spot, supporting the *remix culture* [21] and a sustainable digital economy. We expect that creative industries will engage with the initiative and that the ACE will attract new stakeholders, making the Audio Commons Initiative wider and ultimately contributing in making audio commons accessible to everyone and reusable everywhere. We understand that the model we propose poses challenges at many different levels, from technological development to intellectual property management, copyright policies and business models; and that by addressing them we can have a wide impact in the creative industries. The success of the initiative and its future sustainability largely depends on maximising take-up and adoption of the proposed concepts and ecosystem, and on other parties joining the initiative. We therefore encourage potential new stakeholders to contact us, support and become part of the initiative. We believe that the Audio Commons Initiative will serve as an example model for content reuse in the audio domain that could potentially be extrapolated to other fields and content types.

5 ACKNOWLEDGEMENTS

References

[1] Smith, T., "The social media revolution," *International Journal of Market Research*, 51(4), pp. 559–561, 2009, ISSN 1470-7853, doi:10.2501/S1470785309200773.

[2] Kaplan, A. M. and Haenlein, M., "Users of the world, unite! The challenges and opportunities of social media," *Business Horizons*, 53(1), pp. 59–68, 2010, ISSN 00076813, doi:10.1016/j.bushor.2009.09.003.

[3] Kietzmann, J. H., Hermkens, K., McCarthy, I. P., and Silvestre, B. S., "Social media? Get serious! Understanding the functional building blocks of social media," *Business Horizons*, 54(3), pp. 241–251, 2011, ISSN 00076813, doi:10.1016/j.bushor.2011.01.005.

[4] Hamilton, R., "Designing Next-Gen Academic Curricula for Game-Centric Procedural Audio and Music," in *56th AES International Conference: Audio for Games*, pp. 1–8, 2015.

[5] Janer, J., Kersten, S., Schirosa, M., and Roma, G., "An online platform for interactive soundscapes with user-contributed content," in *41st AES International Conference: Audio for Games*, pp. 1–6, 2011.

[6] Merkley, R., "State of the Commons 2015," 2015.

[7] Font, F., Roma, G., and Serra, X., "Freesound technical demo," in *Proceedings of the 21st ACM international conference on Multimedia - MM '13*, pp. 411–412, 2013, ISBN 9781450324045, doi:10.1145/2502081.2502245.

[8] Brecht, D. M., Mora-Mcginity, M., Fazekas, G., and D. Reiss, J., "The Open Multitrack Testbed," in *137th Convention of the Audio Engineering Society*, 2014.

[9] Bittner, R., Salamon, J., Tierney, M., Mauch, M., Cannam, C., and Bello, J., "MedleyDB: A Multi-track Dataset for Annotation-Intensive MIR Research," in *International Society for Music Information Retrieval Conference*, 2014.

[10] Peeters, G., "A large set of audio features for sound description (similarity and classification) in the CUIDADO project," *CUIDADO IST Project Report*, 54, pp. 1–25, 2004, ISSN nul.

[11] Wiggins, G., "Semantic gap??   Schemantic schmap!! Methodological considerations in the scientific study of music," in *IEEE International Symposium on Multimedia*, 2009.

[12] Lerch, A., *An introduction to audio content analysis: Applications in signal processing and music informatics*, John Wiley & Sons, 2012.

[13] Müller, M., *Fundamentals of Music Processing*, 2015, ISBN 978-3-319-21944-8, doi:10.1007/978-3-319-21945-5.

[14] Lakatos, S., "A common perceptual space for harmonic and percussive timbres." *Perception & psychophysics*, 62(7), pp. 1426–1439, 2000, ISSN 0031-5117, doi:10.3758/BF03212144.

[15] Terasawa, H., Slaney, M., and Berger, J., "Perceptual Distance in Timbre Space," in *Eleventh Meeting of the International Converence on Auditory Display*, volume 5, pp. 61–68, 2005.

[16] Caclin, A., McAdams, S., Smith, B. K., and Winsberg, S., "Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones," *Journal of the Acoustical Society of America*, 118, 2005.

[17] Laurier, C., Meyers, O., Serrà, J., Blech, M., Herrera, P., and Serra, X., "Indexing music by mood: design and integration of an automatic content-based annotator," *Multimedia Tools and Applications*, 48(1), 2009.

[18] Mandel, M. I. and Ellis, D. P. W., "Song-level features and support vector machines for music classification," in *International Conference on Music Information Retrieval*, 2005.

[19] Livshin, A., Peeters, G., and Rodet, X., "Studies and improvements in automatic classification of musical sound samples," in *Proceedings of the International Computer Music Conference (ICMC)*, pp. 171–178, 2003.

[20] Casey, M., "General Sound Classification and Similarity in MPEG-7," *Org. Sound*, 6(2), pp. 153–164, 2001, ISSN 1355-7718, doi:10.1017/S1355771801002126.

[21] Lessing, L., *Remix: Making art and commerce thrive in the hybrid economy*, Penguin Press, 2008, ISBN ISBN 978-1-59420-172-1.