This is an Accepted Manuscript of a book chapter published by Routledge/CRC Press in "Foundations in Sound Design for Linear Media: A Multidisciplinary Approach" on 2019:

Xambó, A., Font, F., Fazekas, G. and Barthet, M. "Leveraging Online Audio Commons Content For Media Production". In Michael Filimowicz (ed.) Foundations in Sound Design for Linear Media: An Interdisciplinary Approach, Routledge. pp. 248-282. ISBN 978-1-138-09396-6.

Please refer to the published version here: https://www.routledge.com/9781138093966 or http://www.crcpress.com/9781138093966

Leveraging Online Audio Commons Content For Media Production

Anna Xambó, Frederic Font, György Fazekas, and Mathieu Barthet

Abstract

With the advent of online audio resources and web technologies, digital tools for sound designers and music producers are changing. The Internet provides access to hundreds of thousands of digital audio files, from human- and nature-related environmental sounds, instrument samples and sound effects, to produced songs ready to use in media production. In relation to the vast amount of creative content available online, an emerging community has forged a culture of sharing. Creative Commons (CC) appears as a legal framework to support such initiative enabling the reuse and remix of creative artefacts. In this chapter, we discuss key concepts and challenges related to the use of CC online audio content (Audio Commons content) for linear media production. We present five use cases connected to the Audio Commons Initiative, illustrating how the gap between audio content creators, digital content providers, sound designers and music producers can be bridged using a web infrastructure and user-friendly tools. The use cases cover various creative production workflows from composition to performance. This chapter discusses novel tools enabling users to "surf" the web in search of sounds matching a creative brief, to import and process CC-licensed audio in the DAW, or to play live performances with laptop ensembles making use of responsive web audio technologies.

1 Introduction

Since the popularization of the Internet in the late 1990s, within the World Wide Web (WWW) ecosystem there has been an exponential growth in storage capacities, *semantic technologies* (i.e. data structured to be understood by both human and machine agents) and social activities (Gruber 2008; Shadbolt et al. 2006). There exists a range of online services that offer both free or paid access to a varied range of multimedia content (e.g. SoundCloud[1] for music, Freesound[2] for sounds, YouTube[3] for videos, Flickr[4] for photos, and so on). New ways of managing this content have emerged (e.g. sharing, reusing, remixing and repurposing), which has led to a new community of *prosumers* who both produce and consume online digital content (Ritzer and Jurgenson 2010). In the field of audio recording, prosumers are those who work in *project studios,* which are professional studios built at home using affordable digital technologies, where prosumers both consume the equipment when purchasing it and produce content from using this equipment (Cole 2011). Another change brought by the Internet has been Creative Commons (CC). CC is a mechanism founded in 2001 to establish a legal and technical infrastructure for sharing content. CC offers a range of licenses and has helped to foster the WWW as we know it nowadays (Merkley 2015). The development of CC licenses has offered a finer-grained level of licensing possibilities, compared to the classical copyright model, which was too strict for the new practices around the generation and reuse of digital content (Lessig 2004).

In this chapter, we present general concepts and technologies aiming at consuming or producing crowdsourced online audio content in the context of linear media production. We will cover CC sound and music content, which we refer to as *Audio Commons* content (Font et al. 2016). This chapter targets sound designers, music composers, researchers, developers, or anyone passionate about sounds and the Internet, with a will to learn how Audio Commons content can be leveraged for media production. In particular, this chapter focuses on the challenges and opportunities of using Audio Commons content and what it can bring to the traditional digital audio

[1] https://soundcloud.com
[2] https://freesound.org
[3] https://www.youtube.com
[4] https://www.flickr.com

workstation (DAW). It is worth noting that Audio Commons content can also be repurposed for interactive media. For those interested in sound interaction, this chapter can be complemented by the second volume from this series entitled "Foundations in Sound Design for Interactive Media".

The remainder of the chapter is organized as follows. First, the concepts of uploading, retrieving, licensing and attributing Audio Commons content are presented, before discussing new workflows enabled by online audio content and services. The second part of the chapter presents five use cases showing how Audio Commons can serve the generation of sound textures, music production, soundscape composition, live coding, and collaborative music making. Lastly, future directions and challenges for sound design and music production are discussed. By the end of this chapter, the reader should be able to (1) identify the key concepts related to Audio Commons and how media production can be changed by using online audio content, (2) get inspired by existing tools and practices repurposing Audio Commons content, and (3) create, share and reuse Audio Commons content. Although equally interesting, it is out of the scope of this chapter to present the web technologies behind online audio databases (e.g. search engines, client-server architecture, semantic web, and so on); more information on these can be found in online tutorials and textbooks.

2 Uploading, Retrieving, and Consuming Online Audio Content

In this section, we present the main concepts related to sound design and music production using online audio content. We will discuss terms linked to the description of audio content (*metadata* and *folksonomies*), their retrieval (*semantic audio*, *text-based queries* vs. *content-based queries*), *Creative Commons licenses*, as well as *online digital audio workstations* (DAWs) and *web application programming interfaces* (APIs).

2.1 Uploading audio content

Storing sound and music online and sharing it instantly with people all over the world might have sounded like science fiction a few years ago. Nevertheless, after the *social media revolution* (Smith 2009), it has become part of our daily routine. Examples of this practice are demonstrated by the success of online sound and music sharing websites like SoundCloud, Bandcamp, Soundsnap, Looperman, Jamendo, Freesound, and many others. These sites host hundreds of thousands of audio files which need to be indexed to become accessible and reusable.

Audio files stored online can be of different audio quality and may be encoded using different file formats. We can distinguish between lossless formats (the audio is preserved in its full quality, typically formats with extensions like *.wav, .aiff and .flac*), and lossy formats (the least important information in the audio is removed so the file size can be reduced, typically formats with extensions like *.ogg* or *.mp3*). To avoid online sound and music collections becoming long lists of audio files which are only identifiable by their filenames, extra *metadata* must be provided. In other words, they need to be described so that sound and music sharing platforms can allow users to find them. Even though some metadata can be automatically derived by computers analyzing audio files (e.g. file format properties, duration, number of channels), richer descriptions still need to be manually provided by humans (e.g. music genre for a song, the microphone used for recording a sound effect, or the location where a sound was recorded). This is also true for other kinds of multimedia items without an intrinsic textual representation, such as video and images (Bischoff et al. 2008).

As one can imagine, there is no single and definitive way to describe all kinds of sounds and music content. For example, describing the recordings of a musical instrument, a sound effect, or a speech, will likely require different sets of information related to production, context and perception.

Sound and music sharing platforms will generally allow users to provide metadata for their sounds in the form of, at least, some keywords (or *tags*) and a textual description. Some platforms let users provide specific metadata fields such as music genre, tempo expressed in beats per minute (BPM) or the artist’s name. Nevertheless, it remains a real challenge to come up with useful keywords and textual descriptions that summarize well the contents of an audio file, especially in the case of non-musical audio content where the most important perceptual aspects are less well established compared to the case of music, which is more easily associated with notation and music theory systems.

A good practice when describing a sound is to look at the following information levels:[5]

- Semantic [S]: information about the sound sources corresponding to the different events appearing in the sound, the actions that generate the sounds, and the meaning of the sounds for the listener.
- Perceptual [P]: description of the sounds' perceptual qualities (e.g. timbre, timing), not necessarily tied to the source(s) of the sound.
- Technical [T]: information about the gear and techniques that were used to create or record the sound.
- Contextual [C]: other aspects like the location where a sound was recorded (if relevant al all) and the purpose of the sound.

All levels are complementary and bear relevant information for indexing and retrieval purposes (Marcell et al. 2001). Figure 1 shows an example of a textual description for a sound where the information levels mentioned above are indicated next to each sentence:

[5] The sound description guidelines presented here are based on those provided by Freesound:: https://freesound.org/help/faq/#hey-i-got-this-bad-description-moderated-file-can-you-help-me-create-a-better-description

An interesting field recording of birds, led by Blue Jays, doing their alarm calls in a peaceful forest ambience [S,C]. High-pitched sounds and chirps reverberating in the woods [P,C]. Recorded with a Sound Devices recorder and stereo microphone (Rode NT4) [T]. I was in the middle of the woods when I noticed that the birds seemed to be alarmed like there was a predator [C, S]. I finally spotted a big Barred Owl sitting high in a tree about 50 yards from my recorder. Pretty amazing how the bird community sounds the alarm [S,C].

Figure 1: Example of rich sound description including Semantic (S), Perceptual (P), Technical (T) and Contextual (C) information (adapted from https://freesound.org/s/151599/).

A description such as the one from Figure 1 can be summarized with a number of tags falling into the semantic, perceptual, technical and contextual categories mentioned above. An example of list of tags related to the sound previously described is shown in Figure 2. One of the benefits of describing a sound with tags is that it can become easier to find similar sounds e.g. through interactive tag clouds. It can also facilitate visualizing tag patterns among a user's own sounds, as later discussed in Section 2.2.

field-recording birds bird-calls blue-jays ambience peaceful forest woods sound-devices rode-nt4

Figure 2: Example of tags describing a sound.

2.2 Retrieving audio content

Existing services for sound and music distribution, including those mentioned in the previous section, use a variety of search and retrieval methods. In current online audio search engines, the most common model for retrieval is the use of a single search box that provides free-text entry. In this approach, keywords, search terms or more general text entered by end users are matched with descriptive file names and metadata, such as title, artist or contributor where appropriate (see Section 2.1). Interactive *tag clouds* with weighted keywords related to occurrences are also employed to help the user finding sounds by more visual means (see Figure 3). It is also common to provide a mechanism for filtering results, i.e., provide a way to show only items in the search results that match designated categories. For instance, this can restrict result sets to the kind of sound library or package the user is interested in, or focus on other metadata elements such as genre, theme or license (see Section 2.3). The result of queries is most commonly presented as a flat list of media items (e.g. sounds, music files) that are typically ranked by a set of relatively simple criteria. Figure 4 illustrates an example of a sequential list of sounds retrieved with Jamendo's Music and Sound Search Tool (MuSST).[6] These items often include popularity, for instance, the number of downloads, some quality rating provided by the users, the duration of the sound files, or the date at which they were uploaded or added to the catalogue of the provider.

Figure 3: Example of an interactive (weighted) tag cloud from Freesound's frontpage website. The size of the tags are proportional to their popularity in Freesound. Selecting a tag yields a list of all the sounds described with this tag.

[6] http://audiocommons.jamendo.com

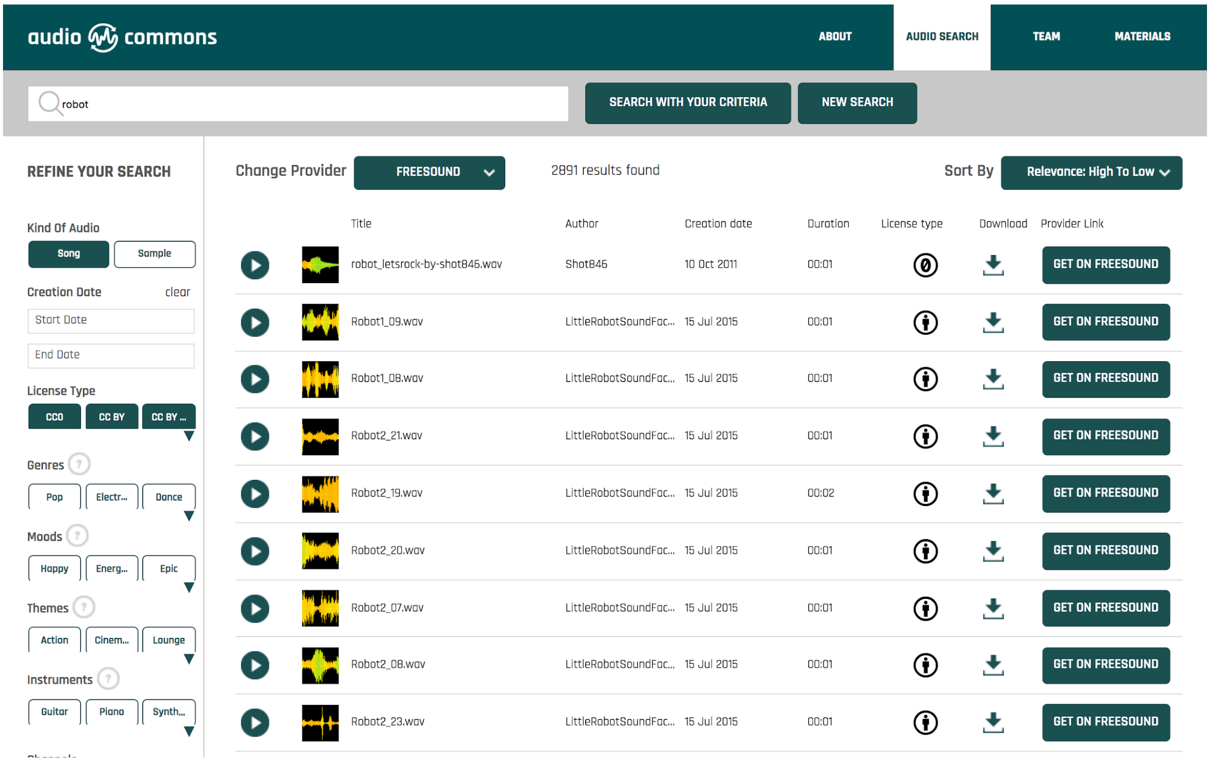

Figure 4: A screenshot example of the query results from Jamendo's Music and Sound Search Tool.

While the search mechanisms described above appear sufficiently rich, navigating large sound or music libraries can become daunting, especially when we consider the large amount of content in existing platforms. In addition, audio is time-based and linear, as opposed to many other forms of media. Consequently, the quality and/or fit for purpose of a retrieved item cannot be judged at a glance, similarly to looking at an image, gleaning over a short piece of text or reviewing a couple of key frames in video. Many services provide an embedded media player that allow users to audition items displayed in the search results prior to downloading or listening to the entire recording. Showing information related to the audio content is also relatively common. This may include a reduced representation of the audio waveform or its frequency content, such as a time-frequency representation or spectrogram. These displays help to judge the complexity of the recording for those who know how to interpret them, e.g., show how rich or sparse the content is in time and frequency, and allow for guessing whether it is a short sound effect, speech or music, for example. However these visual displays tell us surprisingly little about many important criteria of interest, such as the audio quality, the sound sources present in a recording, the meaning of a spoken word content, or the mood which is expressed in a recording.

These problems are commonly addressed using free-form tags aiming to describe audio content thoroughly, a technique discussed in Section 2.1. Audio may be tagged by experts or collected from contributors or end users of sound sharing services. Tags collected this way are often referred to as *folksonomies* (Lamere 2008). Expert annotation is a slow and expensive process that does not scale with rapidly growing collections. Crowdsourcing allows for solving this issue, but it has its own problems, most prominently, folksonomies are plagued by lack of consistency, biases and other sources of “noise” in the tagging process (Choi 2018).

Automatic analyses of the audio content and its meaning, a procedure often identified as *semantic audio*, allows for the extraction of metadata such as tempo, instruments or perceptual qualities (e.g. rough or soft) directly from the audio recordings. This is becoming increasingly common in sound sharing services, despite audio feature extraction, and the representation of the resulting metadata, is still an active field of research (Fazekas et al. 2011). Although either many high-level or complex semantic audio descriptors cannot reliably be extracted yet, several services employ automatic analysis tools. These tools provide search or filtering functionality based on tempo or key estimation and even higher-level analyses aiming to determine musical genre or mood.

A possible way to overcome the limitations of complex semantic feature extraction explained above is to navigate by similarity to a selected seed item from a collection. There are machine learning and visualization techniques that enable placing similar items close to each other in low-dimensional spaces. When these items are visualized, they can facilitate retrieval by navigating the collection by similarity amongst its items. For example, Freesound Explorer[7] is a visual interface for making queries to Freesound and exploring the results in a two-dimensional space where sounds are organized according to timbral similarity (see Figure 5). The map is computed using a dimensionality reduction technique over automatically extracted

[7] https://labs.freesound.org/fse/

spectral audio features (Font and Bandiera 2017). In this way, closer sounds in the timbre space tend to have similar timbre, and clusters of search results naturally emerge in different parts of the visual space. This allows users to navigate Freesound content by combining a standard text-based search mechanism and a visual method for exploring the results.

Figure 5: A screenshot example of Freesound Explorer.

Studies on the needs of modern music consumers have revealed a strong interest in being able to search and browse music by mood (Lee and Downie 2004). Researchers have investigated semantic mood models aimed at music recommender systems (e.g. Barthet et al. 2013). A common psychological model devised to characterise human emotions is Russell's arousal/valence (A/V) model (1980) which represents levels of activation or excitation (*arousal*) and positiveness (*valence*). The A/V model has been applied to music retrieval in a number of works. We provide here examples related to the Moodplay web-based social music player described in (Barthet et al. 2016). Figure 6 shows an example of a search interface letting users search for songs according to emotional indications expressed using the two-dimensional A/V model (in the model used, "uplifting" music is represented as light positive music). Figure 7 displays a map of the songs available in the

Moodplay music player's library, resulting from predictions of the emotions expressed by the songs in the A/V space using Semantic Audio.

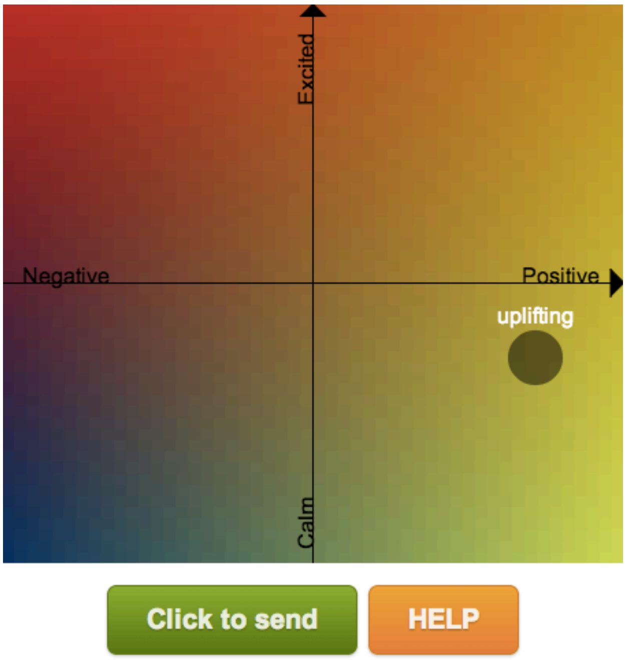

Figure 6: Example of Moodplay's user interface to select music based on the arousal/valence emotion model (Barthet et al. 2016).

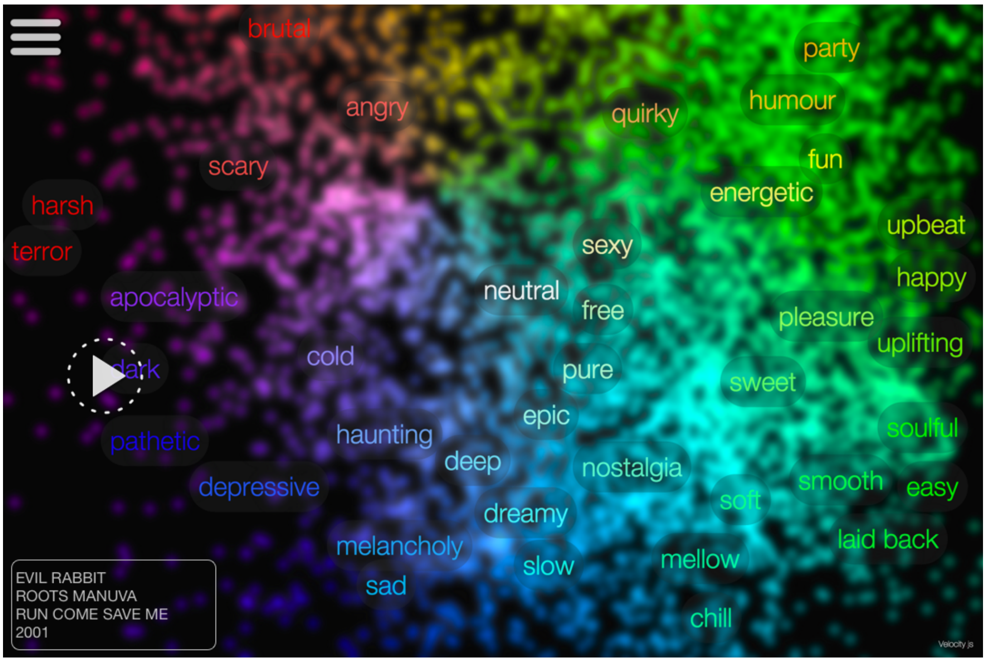

Figure 7: Example of Moodplay's user interface displaying music tracks in the two-dimensional arousal/valence space (Barthet et al. 2016).

2.3 Licensing and attribution

The 'right to copy' or *copyright* is a legal right that exists in many countries in relation to an original piece of work and its creator(s). The legal framework covers aspects such as reproduction, derivative works, distribution and public exhibit of an original work. The copyright has a limited amount of time (typically the *copyright length* is 50-100 years after the passing of the creator, a number of years that varies depending on the country). For example, since 1998, the copyright in the United States lasts a creator's life span plus 70 years, while this extends to 95 years for copyrights owned by corporations. This means that royalties need to be paid for the use of intellectual works during this period (Lessig 2000). Once copyright expires, the original work becomes part of the public domain and the public right of common use takes effect enabling its free and unrestricted use. Copyright has a number of benefits, namely giving the creator(s) credit to their original work, legal protection against plagiarism, and potential revenues from distribution and derivative works. However, copyright law can limit the growth of creative arts and culture because of the long wait required until having free access to these works. The use of derivative works (e.g. unpaid, unnegotiated) can be penalized unless it is released to the public domain. The risk that Lessig refers to as *copyright perpetuity* (Lessig 2000), and which prevents an original piece to become part of the public domain and the free circulation of creative ideas, was a relevant topic of debate during the foundations of the WWW. This debate, motivated by the advent of digital content and network capabilities, fostered the creation of Creative Commons (CC).

CC licensing relates to the concept of *free culture*. Here, *free* is understood not as in "free beer", but as in e.g. "free speech," "free markets," and "free will,"' (Lessig 2004, p. xiv). Inspired by Stallman's ideas on free software and free society, Lessig distinguishes a *free culture* (the desired CC model) from a *permissions culture* (the existing copyright model). In a free culture, creators and innovators are supported and protected by granting intellectual property rights. In a permissions culture, creators and innovators can only create with the permission of creators from the past. CC licensing is thus connected to the notion of providing as much freedom as possible to the creators within a legal framework. It is worth mentioning that Lessig emphasizes that free culture is not a culture without property, or where creators are not paid. It is instead “a balance between anarchy and control” (Lessig 2004 p. xvi).

As noted in Merkley (2015), CC-licensed work has nearly tripled between 2010 (400 millions CC-licensed works) and 2015 (over 1 billion CC-licensed works). CC content type includes images (photos, artworks), videos, research (journal articles), open educational resources, texts (articles, stories, documents), audio tracks (4 millions reported in 2015 from 16 platforms) and other (multimedia, 3D) (Merkley 2015). The various licenses are described in the CC organization website.[8] Table 1 outlines the four main components in CC licenses that are typically combined. Figure 8 shows the combination of these four components resulting in six different license options that span, in a continuum, from less to more permissive licenses from the point of view of free culture.[9] These licenses are providing various degrees of freedom when releasing a creative work. A more detailed explanation of the licenses illustrated with examples can be found on the CC webpage[10] with a service that helps to choose the most suitable license.[11]

CC Elements	***Description***
BY	Attribution or the need to credit the original creation.
NC	NonCommercial or building upon the original work noncommercially.
ND	NonDerivatives or keeping unchanged the original creation.
SA	ShareAlike or license the new creations under identical terms than the original creation.

Table 1: Description of the available CC elements.

[8] https://creativecommons.org/licenses/
[9] This flowchart by CC Australia is helpful for choosing the most suitable CC license: http://creativecommons.org.au/content/licensing-flowchart.pdf
[10] https://creativecommons.org/share-your-work/licensing-types-examples/licensing-examples/
[11] https://creativecommons.org/choose/

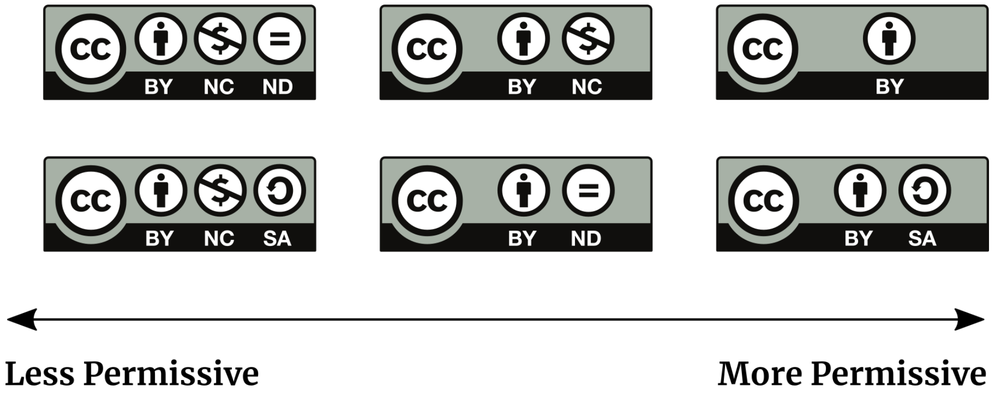

Figure 8: Continuum of the available CC licenses from less to the more permissive levels.

When using CC-licensed audio content, it is important to make sure that the planned application matches the license requirements. For instance, reusing a CC-licensed musical track for commercial purposes may require licensing fees (e.g. see the Jamendo Licensing stock music for commercial use).[12] Placing a work in the public domain or *CC0* (no rights reserved) is also possible. As explained in the Freesound website,[13] apart from CC0 licenses, CC-licensed audio content should be credited by citing the title of the sound/music, the author and a link to the resource. Figure 9 exemplifies how to attribute correctly a CC-licensed audio item. If the list of sounds is too long to be displayed in the credits section, alternatively it is also suitable to provide a link to a separate document with the list (e.g. "`for the full list see here: http://www.mysite.com/work-credits.html`").

This [video/performance/...] uses these sounds from Freesound:

- sound1 by user1 (http://freesound.org/people/user1)
- sound2, sound3 by user2 (http://freesound.org/people/user2)
- (...)

Figure 9: Example of how to attribute sounds from Freesound (example's source: Freesound.org).

[12] https://licensing.jamendo.com

[13] A detailed explanation on how to credit properly can be found in the Freesound's FAQ section: https://freesound.org/help/faq/#how-do-i-creditattribute

A r*emix* or the creation of an adaptation from existing CC-licensed content needs to be done carefully to comply with CC licensing. Typically, it is possible to remix CC content when it is licensed without the NonDerivative element. As a rule of thumb, it is possible to create an adaptation and release it under a similar, compatible license or a more restrictive license (unless it has the ShareAlike element), but never the other way around. Table 2 highlights the license restrictions when publishing new sounds or music that include, modify, or remix others' sounds or music.[14] [15]

Type of license for user A's work	**Type of license for user B's remix**	**Can user B remix A's work?**
CC0	CC0	Yes
CC0	BY	Yes (*)
CC0	BY-NC	Yes (*)
BY	CC0	No
BY	BY	Yes (**)
BY	BY-NC	Yes (**)
BY-NC	CC0	No
BY-NC	BY	No
BY-NC	BY-NC	Yes (**)

(*) If a third user C adapts the creation from user B, they must attribute it to user B.

(**) User B must attribute the creation to user A. If a third user C uses the creation from user B, they must attribute both users A and B.

Table 2: License restrictions when creating a remix (example's source: Freesound.org).

[14] A detailed explanation on how to remix or repurpose new sounds or music can be found in the Freesound's FAQ section: https://freesound.org/help/faq/#license-restrictions-when-publishing-new-sounds-that-includemodifyremix-other-sounds.

[15] A complete chart and further explanations on how to legally remix CC-licensed material can be found in the CC website: https://creativecommons.org/faq/

2.4 Transforming linear audio production using online and web applications and resources

Before venturing into how the Internet can transform linear audio production, some of the differences between online, web and cloud applications are worth mentioning. An *online application* or tool is a software installed locally on a computer and which relies on an Internet connection to access information (e.g. the Skype communication tool). A *web* (or *web-based*) *application* or *tool* is a software based on a client-server architecture for which the client side runs in a web browser (e.g. the Google Docs word processor). For *cloud-based applications*, the majority of the processing and data storage take place in remote servers and data centers. Cloud applications are designed so as to be operational most of the time for potentially large numbers of concurrent users. They can be operated from the web browser and/or be installed on desktops (e.g. the Dropbox file hosting software). These different architectures can lead to a varied collection of interaction models to leverage online audio content and services for media production. Next, we review how web technologies can be applied to tools used for sound design, music production and live performance, such as DAWs, plugins and live coding environments.

Internet-connected DAWs

Historically, digital audio workstations (DAWs) and digital musical interfaces have been conceived to operate with local resources and content (e.g. local databases of sounds or music) by being for the major part disconnected from the Internet (apart for updates and license authentication). DAWs have traditionally operated in isolation from the Internet both by design, for example for security and data protection reasons, and due to technological limitations, for example, due to the lack of mature web standards. However, several connected audio production tools have emerged recently, paving the way for the Internet to transform audio production similarly to other domains such as communication.

Turning a DAW into an online tool changes its status of a “closed box” to one where the amount of information and audio content available to end users can be made more open-ended. Here, connected DAWs refers to desktop DAWs connected to the

Internet, in contrast to browser-based DAWs, which are discussed in the following section "Web-based music production applications and resources". When looking for specific books, searching a personal local library collection may be limited compared to having access to the content from a well-curated public library. Likewise, when searching for specific sounds, personal audio collections or libraries of music samples and loops featured in some modern DAWs, may not provide content which is topical enough for a given creative task. Online audio collections provide an opportunity to diversify the audio content that one can access to support creative needs. Storing large collections of audio files locally also takes space compared to downloading online content only when it is needed.

In addition to gaining access to a wider range of audio content, a connected DAW may also leverage web services providing information facilitating music analysis and production. These services can expand the intrinsic capabilities of native DAWs programmed for specific platforms. For example, artificial intelligence services could be deployed to provide DAWs with the capability to analyse multitrack audio to infer knowledge about musical attributes, e.g., instruments, chords, structure (Fazekas et al. 2011) or to make synthetic renderings more expressive by modifying timbral patterns (Barthet et al. 2007). Cloud services can also be tailored at making recommendations to find similar-sounding tracks (Fazekas et al. 2013), which could be used to find reference tracks in a mastering session, or to find songs to learn for pedagogical purposes (Barthet et al. 2011). Web services may also be designed to share creative information generated within the DAW with online communities, for example metadata characterising associations between equalization (EQ) parameters and timbre to investigate various mixing techniques (Stables et al. 2016). Internet-connected DAWs also enable to switch the workflow model from individual to collaborative music production. Tools such as Avid Cloud Collaboration for Pro Tools[16] or Ohm Studio[17] let multiple users access and work together on a same production project.

In summary, from the perspective of linear media production, Internet-connected DAWs have the following (nonexhaustive) interesting features:

[16] http://www.avid.com/pro-tools/cloud-collaboration
[17] https://www.ohmforce.com/OhmStudio.do

- To support creativity by enabling access to audio content coming from a wide diversity of online audio content providers.
- To benefit from web services based on artificial intelligence algorithms that expand the capabilities of native DAWs by providing users with additional metadata.
- To share creative content with online user communities.
- To collaborate on a production with remote users.

Figure 10 illustrates how web technologies may benefit desktop DAWs by enriching features related to sound content, music analysis, synthesis, and distribution.

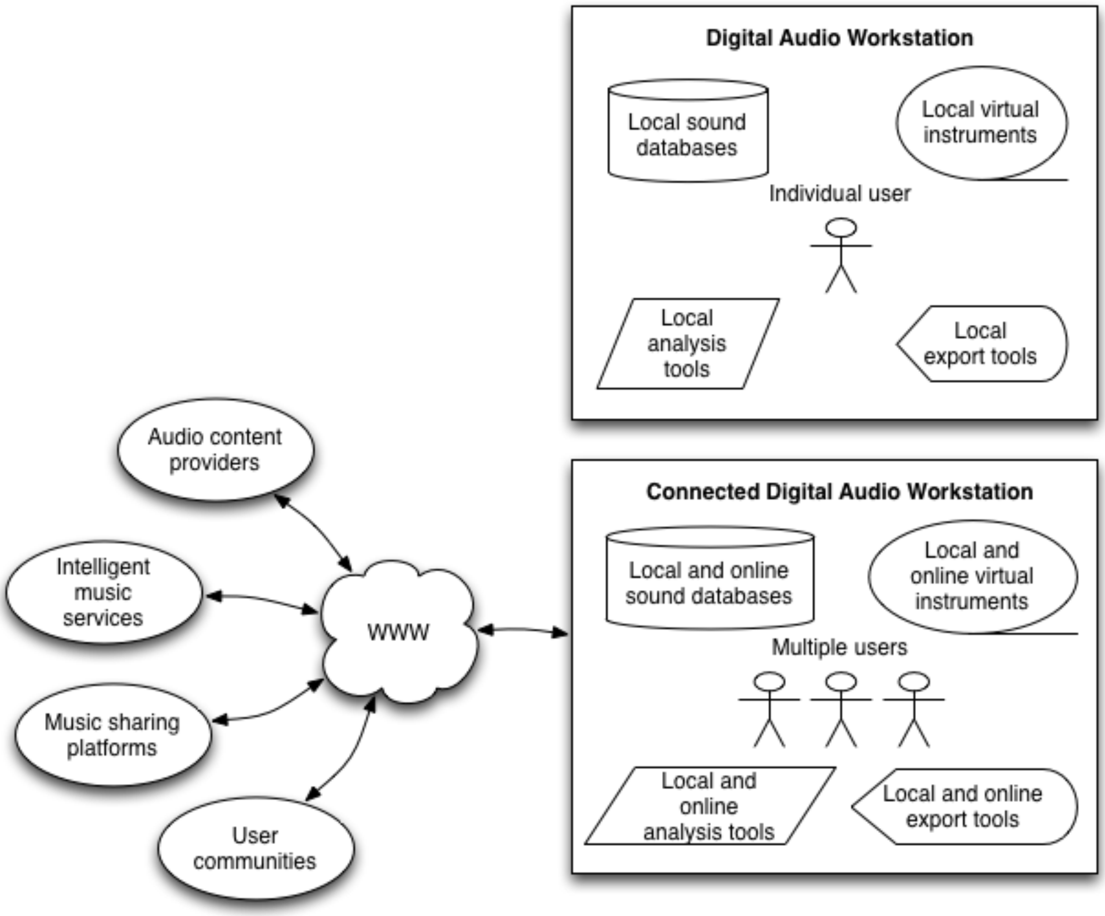

Figure 10: Shift from standard (top) to Internet-connected (bottom) digital audio workstations.

There are different ways in which desktop DAWs can benefit from features brought by the Internet. One approach is to have web functionalities integrated directly within the DAW by means of appropriate extensions that expose features through the graphical user interface (GUI). For example, some manufacturers enable this through their Software Development Kit, see e.g. Reaper's Extensions SDK.[18] Another approach consists in installing web-enabled plugins developed using interfaces compatible with the DAW (e.g. Virtual Studio Technology, Audio Units). We present in Section 3 two examples of web-enabled audio plugins allowing users to search for, retrieve and process samples from online audio content providers (Waves Audio's SampleSurfer and AudioGaming's AudioTexture). A third, less direct approach involving apps external to the DAW, is to use online applications bridging the gap between online audio content and standard DAWs. This is the case for the Splice Studio[19] application which enables users to search for sounds from an online database and import them to the DAW using drag and drop actions.

Web-based music production applications and resources

Contrary to online desktop applications, web-based applications run in the browser and generally do not require users to install additional software components locally. The evolution of audio frameworks and APIs for the web, such as Web Audio, a high-level JavaScript API for processing and synthesizing audio,[20] has made possible the development of web applications supporting major features from desktop DAWs such as audio signal routing, sample-accurate sound playback with low latency, high dynamic range, and mixing processing techniques. Chapter 13 in this series' 2nd volume discusses how Web Audio can be used to create interactive music applications on the web.

Soundtrap[21] is an example of browser-based collaborative music production platform exploiting Web Audio (Lind and MacPherson 2017). CCMixter[22] acts as a social music platform connecting instrumentalists, vocalists and producers to create music collaboratively using audio content licensed under CC (see Section 2.3). This platform facilitates the sharing of *stems*, which are digital audio files containing a group of instruments or vocal tracks serving a specific function in a musical

[18] https://www.reaper.fm/sdk/plugin/plugin.php
[19] https://splice.com/features/studio
[20] https://www.w3.org/TR/webaudio/
[21] https://www.soundtrap.com
[22] http://ccmixter.org

arrangement (e.g. bass, rhythmic, singing, accompaniment parts), and that can be used in combined ways to create novel music compositions (often called *remixes*). Playsound.space,[23] which we describe further in Section 3, is a web application designed to let users mix CC-audio content retrieved from the Freesound online provider by using semantic terms (Stolfi et al. 2018). The application can be used to compose soundscapes (e.g. Pigrem and Barthet 2017) or to play free music improvisations.

Cloud-based live coding

Live coding is a musical practice involving the use of computer programming to generate sounds by writing and executing code on the fly (Collins et al. 2003). If the practice has traditionally relied on sound synthesis and samples stored on local sound databases or more rarely, online databases, recent approaches have investigated how multiple online audio content could be repurposed during live coding (Xambó et al. 2018). Software modules can interface with a dedicated API, like the Audio Commons one, so that live coders can pull and further process online sounds during a live performance. This approach is examined further in Section 3.

2.5 Audio Commons Ecosystem

In the previous section, we reviewed several models with which online audio content and services could benefit linear media production. In order to bridge the existing gap between online audio content providers, music production software and end users, there is a need to establish a networked architecture and tools enabling the exchange of audio and licensing information at different points of the sound design and music production chain. The *Audio Commons Initiative* (Font et al. 2016) has started as a European-funded project investigating these issues with a particular focus on crowdsourced online audio content licensed under Creative Commons.

The *Audio Commons Ecosystem* (ACE), described in Figure 11, refers to the complex network made up of interconnected audio content, users (e.g. creators, consumers) and software systems, designed to support the aims of the Audio

[23] http://www.playsound.space

Commons Initiative. The ACE is designed so that content providers can expose CC audio content to amateurs and professionals from the creative industries alike, and to provide an infrastructure for users to seamlessly integrate such CC content in creative workflows. The use cases presented in Section 3 introduce technologies implementing this ecosystem.

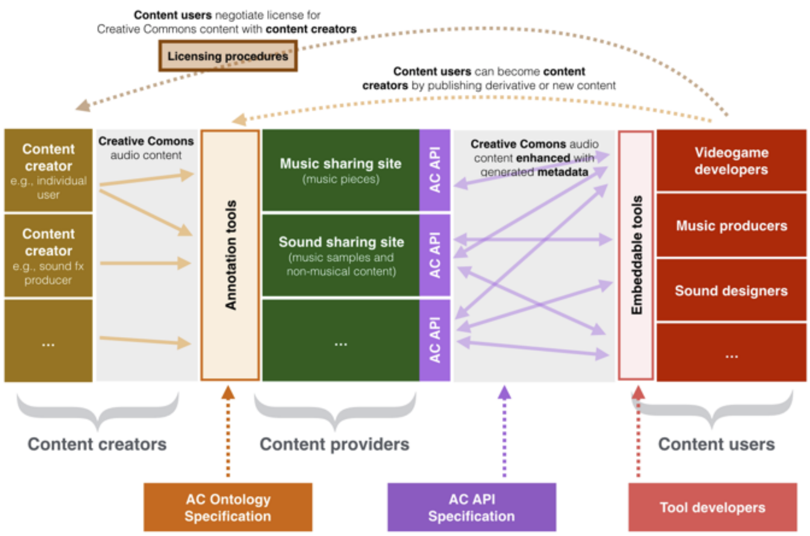

Figure 11: Conceptual diagram of the different components that are interconnected in the Audio Commons Ecosystem.

To date, the Audio Commons Ecosystem includes the following content providers:

- Freesound,[24] providing a crowdsourced audio collection of several hundreds of thousands of nonmusical and musical sounds released under CC licenses.
- Jamendo,[25] providing curated collections of several hundred of thousands of songs from independent artists for free streaming and downloads (Jamendo Music) and commercial use through licensing (Jamendo Licensing).
- Europeana, hosting a collection of music and sounds related to European cultural heritage under CC licenses.[26]

A portal to the ACE content can be found online (Audio Commons Music and Sound Search Tool, 2018).

[24] https://freesound.org
[25] https://www.jamendo.com
[26] https://www.europeana.eu/portal/en/collections/music

2.6 Integrating Audio Commons content using web APIs

The Audio Commons Ecosystem provides a web API that can be used to integrate Audio Commons content from the aforementioned sound and music providers in third party applications. The Audio Commons API is offered as a REST API[27] that acts as an intermediary for individual content provider APIs (e.g. Freesound, Jamendo and Europeana) and offers a unified interface to access all services. In the rest of this section we show simple "hello world" examples for using the Audio Commons API. A basic knowledge of web technologies is assumed. It is out of the scope of this chapter to describe how to build a platform for sound sharing and retrieval, which the reader can learn from other sources (e.g. see Font et al. 2017).

In order to use the Audio Commons API, an account needs to be created at the website of the Audio Commons Mediator.[28] Then, API credentials must be generated through the "Developers" page. The reader can find details about how to generate API credentials (and other Audio Commons API usage topics not covered in this section) in the Audio Commons API documentation.[29] Assuming API credentials have been obtained and are properly included in the requests sent to the API, we can use the "Search" URL to query all content providers at once and get all responses in a single JSON file:

```
REQUEST:
https://m.audiocommons.org/api/v1/search/text/?q=cars

RESPONSE:
{
  "meta": {
    "response_id": "8f64cbcd-47d2-4ac8-bebf-2c2d9f416cef",
    "status": "PR",
    "n_expected_responses": 3,
    "n_received_responses": 0,
    "sent_timestamp": "2018-07-03 13:19:20.091232",
    "collect_url": "https://m.audiocommons.org/api/v1/collect/?rid=8f64cbcd",
    "current_timestamp": "2018-07-03 13:19:20.101594"
  },
  "contents": {}
}
```

[27] https://en.wikipedia.org/wiki/Representational_state_transfer
[28] https://m.audiocommons.org
[29] https://m.audiocommons.org/docs/api.html

This will return a response containing the URL of where the search results can be retrieved as soon as individual content providers return a search result response. Accessing this URL will show the list of results retrieved so far:

```
REQUEST:
https://m.audiocommons.org/api/v1/collect/?rid=8f64cbcd

RESPONSE:
{
  "meta": { ... },
  "contents": {
    "Jamendo": {
      "num_results": 48,
      "results": [{
        "ac:id": "Jamendo:1317252",
        "ac:url": "http://www.jamendo.com/track/1317252",
        "ac:name": "1000miglia",
        "ac:author": "Naturalbodyartist",
        "ac:license": "BY-NC-SA",
        "ac:preview_url": "https://mp3d.jamendo.com/download/track/13172..."
      }, ... ]},
    "Freesound": {
     "num_results": 7153,
     "results": [{
        "ac:id": "Freesound:326146",
        "ac:url": "https://freesound.org/s/326146/",
        "ac:name": "Inside Car Ambience Next to School ...",
        "ac:author": "15050_Francois",
        "ac:license": "BY-NC",
        "ac:preview_url": "https://freesound.org/data/previews/326/32614..."
     }, ... ]},
    "Europeana": {
     "num_results": 25,
     "results": [{
        "ac:id": "Europeana:/916107/wws_object_2164",
        "ac:url": "http://www.europeana.eu/portal/record/91610...",
        "ac:name": "Brokindsleden - The sounds of traffic",
        "ac:author": null,
        "ac:license": "BY",
        "ac:preview_url": "http://www.workwithsounds.eu/soundfiles/5a7/5..."
     }, ... ]}
  }
}
```

Note that in each of the audio results above there is a `ac:preview_url` metadata field which points to a preview version of the actual sound file. This preview version can be played and downloaded. Queries can be narrowed down using filters. For example, the following query will return the same results as the one above, but will only include sounds with Creative Commons Attribution license.

```
REQUEST: https://m.audiocommons.org/api/v1/search/text/?q=cars&f=ac:license:BY
```

The Audio Commons API provides a unified access to all of the services of the Audio Commons Ecosystem, but sometimes these services offer specific functionalities which are not supported by the Audio Commons API and can only be used by directly accessing the service's own API. As an example, Freesound supports similarity-based queries which, at the time of this writing, are not supported by the Audio Commons API. To use such service, requests will need to be addressed to Freesound instead of Audio Commons. For example, assuming API credentials for Freesound have been obtained,[30] the following request will return a list of sounds which sound similar to the target Freesound sound with ID `291164`:

```
REQUEST:
https://freesound.org/apiv2/sounds/291164/similar/
```

3 Repurposing Audio Commons Content: Use Cases

In this section, we present five use cases from sound design to music production and live performance applications that exploit crowdsourced sounds differently: (1) sound texture generation using AudioGaming's AudioTexture plugin; (2) music production using Waves Audio's SampleSurfer plugin; (3) soundscape composition leveraging online audio content; (4) live coding with the MIRLC library for SuperCollider, and (5) compositions through semantic ideation using the web-based app Playsound.space. This section should provide the reader with the width and breadth of potential novel media production applications relying on cloud-based audio databases.

3.1 Sound Texture generation with AudioTexture

Examples of sound textures include the sound of rain, crowd, wind and applauses (Saint-Arnaud and Popat 1995; Strobl et al. 2006). Although there is no consensus

[30] Documentation for the Freesound API can be found in http://freesound.org/docs/api/

on the definition of *sound texture*, an agreed working definition includes two main features: constant long-term characteristics and short attention span (Saint-Arnaud and Popat 1995; Strobl et al. 2006). The *constant long-term characteristics* refer to a sound that emits similar characteristics over time (e.g. sustained pitch and rhythm) irrespective of the presence of local randomness and variation. This means that if two snippets are randomly picked from the same sound texture, they should sound similar. Therefore, in a sound texture, the sound is constantly sustained. The *attention span* refers to the time needed to characterize the texture, typically a few seconds. It is also worth mentioning that in conjunction with the high-level characteristics, looking closer, at a lower level, we find that sound textures are formed of *atoms* (basic sound snippets) that are repeated periodically, randomly, or both, a behavior that is defined by the high-level characteristics. Sound textures can have multiple applications, ranging from background music and game music, to audio synthesis and audio signal restoration (Lu et al. 2004).

Developed by AudioGaming, AudioTexture is a plugin prototype for sound texture synthesis that leverages Audio Commons by bringing CC-licensed audio content into the DAW. The AudioTexture plugin lets users generate sound textures from audio recordings from either online or local databases within a DAW environment, such as Logic Pro X, Ableton Live, or Reaper. In particular, the plugin integrates Audio Commons content for creative sonic/musical explorations in the form of sample-based synthesis or *concatenative synthesis* (see Schwarz (2007) for a discussion on concatenative synthesis), which refers to the computational generation of sounds using existing sound samples. The plugin is particularly suited for environmental sounds with short-term repetitive units or atoms (e.g. water drops, rock falls, construction work, and so on). It is however possible to use the plugin with musical sounds which can also lead to interesting textures.

To operate AudioTexture, first users need to select a sound from either an online content provider from the Audio Commons Ecosystem such as Freesound (see Figure 11) or from a local database. The sound is represented with a waveform providing overall temporal and envelope cues and a larger spectrogram for frequency and energy cues over time (see Figure 12). The plugin decomposes with automatic segmentation the audio signal into adaptively defined atoms that are not equal in size. Markers can be visually activated to display the positions of the atoms and understand the behavior of the sound synthesis algorithm. The unit size of the atoms can be changed ranging from multiple small atoms to a fewer large bigger atoms. It is possible to select a range of the full sound using the horizontal slider X, which can be moved to start at any position of the sound. The control of the sound synthesis (the logic of how the atoms are played) is mainly determined by the values of three semantic (low-level audio) descriptors grouped in the vertical slider Y: energy, noisiness and brightness. AudioTexture includes factory presets that can guide users with suitable values (e.g. rain, fire, footsteps, birds, mechanics, waterfall, wave, applause, music).

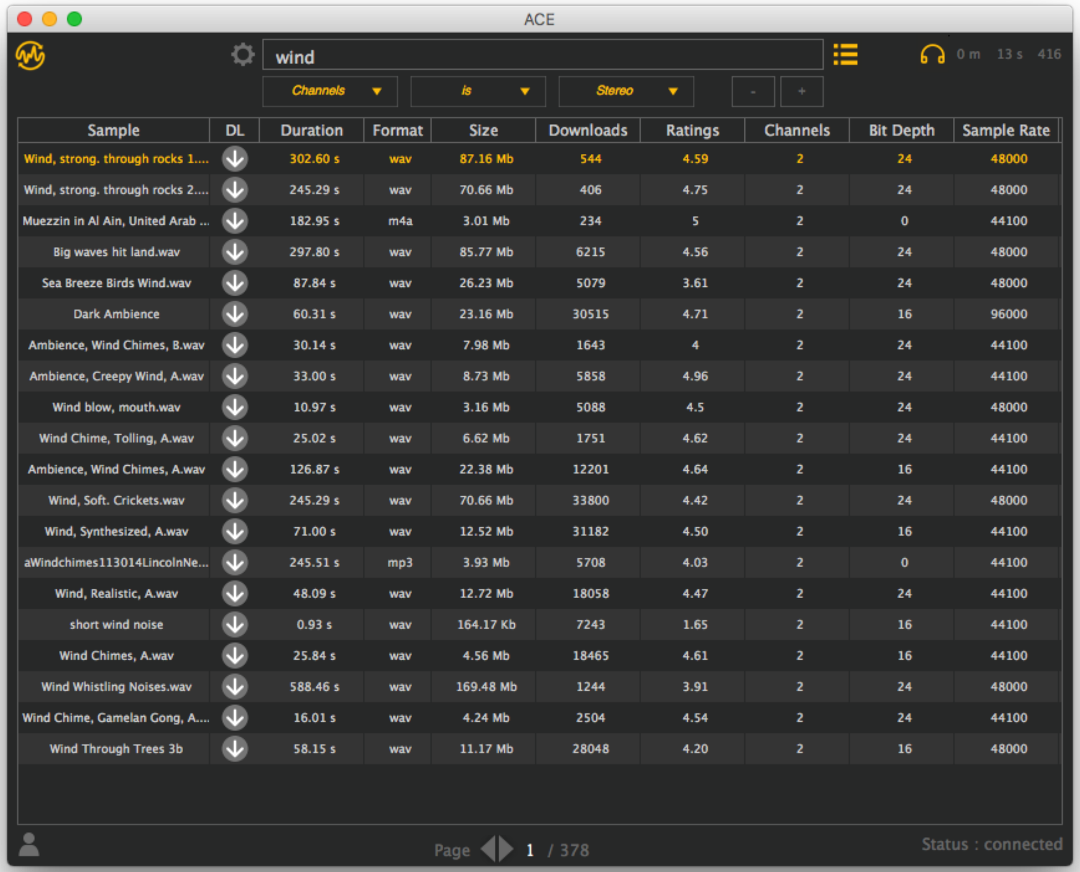

Figure 12: Screenshot of the Audio Commons retrieval interface in AudioGaming's AudioTexture plugin.

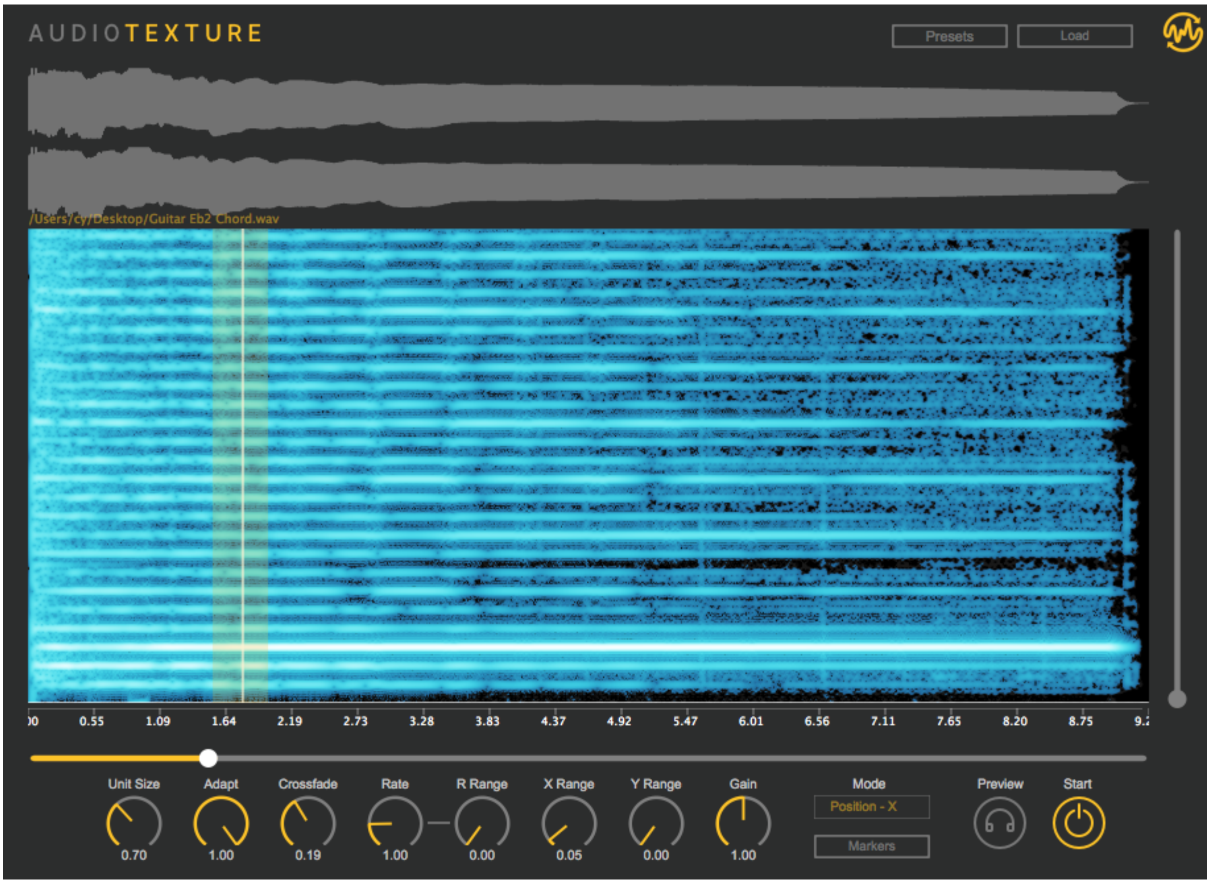

Figure 13: Screenshot of the sound editing interface in AudioGaming's AudioTexture plugin.

The resulting sound texture can be recorded in the DAW by routing the output of a track with the plugin to another track armed for recording. A set of examples are provided in the companion website, which showcases the sonic and musical possibilities of the plugin using crowdsourced sounds (Audio Commons Routledge Website 2018).

3.2 Music Production with SampleSurfer

Music production is a whole field within music technology, which includes sound recording (Huber and Runstein 2013), mixing (Owsinski 2013), and sometimes audio mastering (Katz 2003). Music production software is typically referred to as digital audio workstation (DAW), which we discussed in Section 2.4.

SampleSurfer has been developed by Waves Audio LTD and is another plugin for the ACE that serves as an audio content search engine based on semantic metadata and musical features. The plugin is designed to integrate Audio Commons sound and music samples in a DAW-based environment by providing basic editing capabilities (e.g. fades, trims) to optimize the music production workflow. It supports well established DAW applications, such as Logic Pro X and Ableton Live.

As shown in Figure 14, the plugin lets the user choose from a set of CC-licensed audio content providers (including Freesound, Jamendo), search by a set of filters including type of CC license, harmony (key, scale), beat (BPM, duration) and sample characteristics (sample format, sample rate, bit depth, number of channels). From a query, the user gets back a filtered list, from which sounds can be selected and downloaded. The user will be able to access the download history. The plugin offers edit tools, such as fade in and out, change of the amplitude level, or trim of the sound to a shorter range from the original audio clip. Another feature offered by the plugin is the possibility of connecting to a folder in the local drive, scan and analyze the content, and then be able to search the local material. The companion website presents a set of examples that illustrate the capabilities of the plugin (Audio Commons Routledge Website 2018).

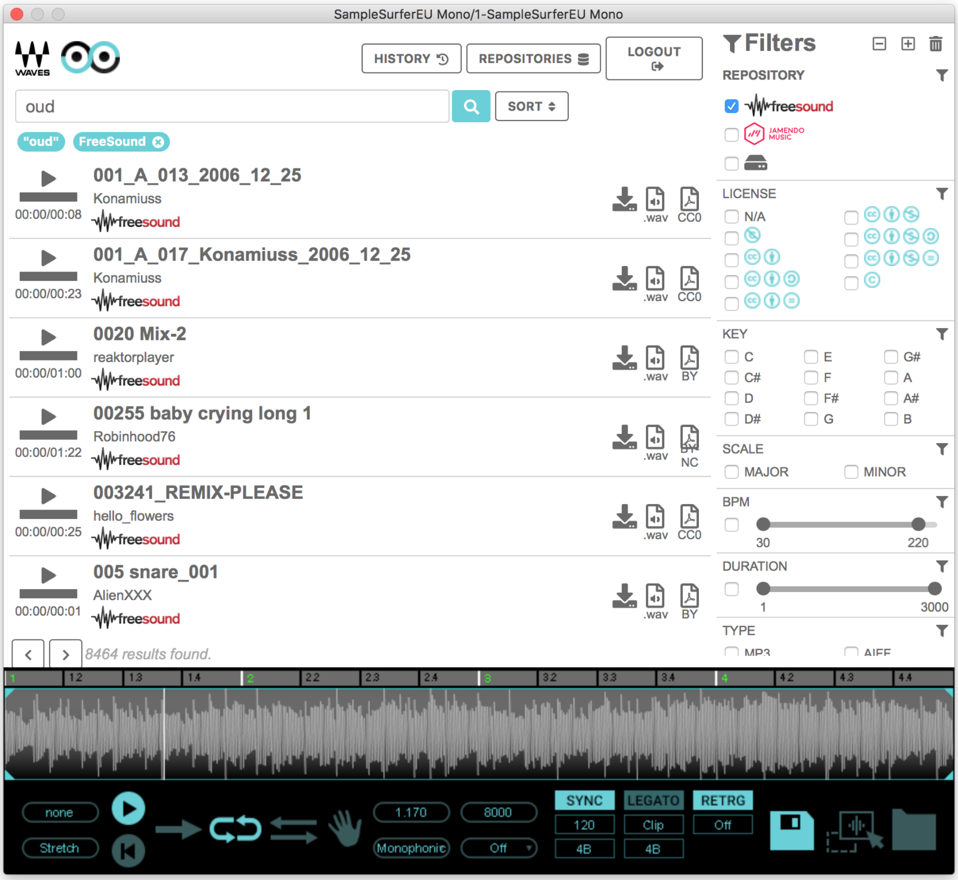

Figure 14: Screenshot of the search interface and edit tools from Waves Audio's SampleSurfer plugin.

3.3 Soundscape Composition using Online Audio Content

Soundscape composition has grown from acoustic ecology and soundscape studies, which are fields seeking to document, archive and analyze the evolving sounds of our world (Schafer 1993). Barry Truax defines *soundscape composition* as a form of electroacoustic music "characterised by the presence of recognizable environmental sounds and contexts, the purpose being to invoke the listener's associations, memories, and imagination related to the soundscape".[31] Reviews on soundscape composition approaches and their applications can be found in Truax (2002) and Pigrem and Barthet (2017), respectively. A large amount of crowdsourced online audio content is based on recordings of human- and nature-related environmental sounds. About two-thirds of Freesound's content falls within these categories, and cultural heritage resources such as Europeana also dispose of environmental

[31] https://www.sfu.ca/~truax/scomp.html

recordings.[32] Crowdsourced CC audio content can hence be a rich resource for soundscape composers looking to convey meanings about place and time through audio.

In the Fall of 2017, students from the *Sound Recording and Production Techniques* module led by Mathieu Barthet at Queen Mary University of London were invited to produce short soundscapes leveraging Audio Commons online audio content and tools. Soundscape themes were ideated in class inspired by the *bootlegging* participatory design technique introduced in (Holmquist, 2008). Students had to write down on post-its two ideas in each of the four following categories: character, place/environment, situation/action, and mood. Post-its were shuffled and students had to pick up randomly one idea per category (see Figure 15).

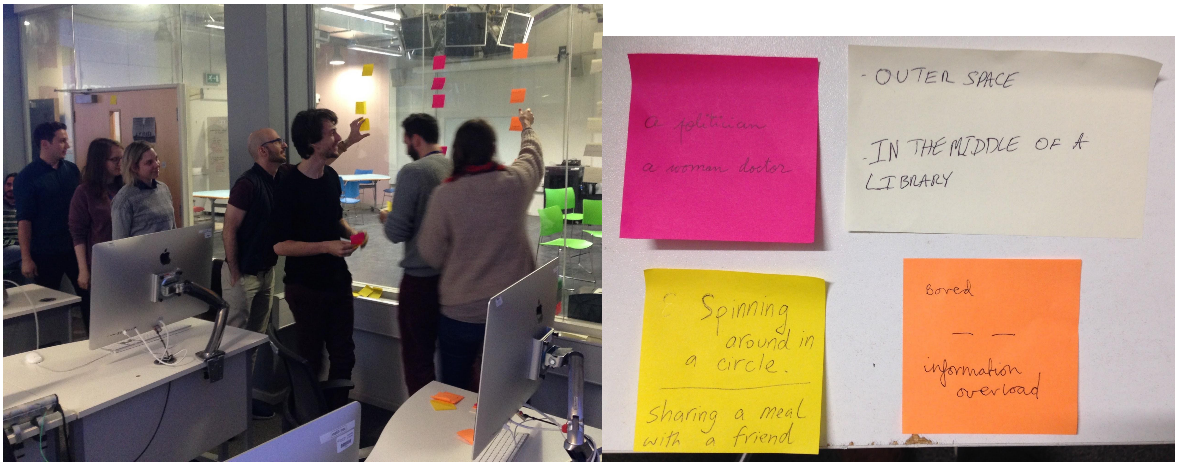

Figure 15: Participatory ideation of soundscape themes - students picking up post-its in character/place/situation/mood categories (left) and examples of categories (right).

After a brainstorming session, original soundscape themes emerged resulting from the combination of ideas from each category. Students were given three weeks to produce their short soundscapes. They could use an opened range of approaches, from figurative to abstract, and were given as a creative constraint to only use found sounds or loops retrieved/processed using Freesound and AudioTexture (previously introduced), or Apple Loops (a royalty free collection of pre-recorded musical patterns and sound effects). Some examples of soundscapes produced by the

[32] https://www.europeana.eu/portal/en/collections/music

students can be found on SoundCloud.[33] Although it mainly served as an academic purpose, this exercise demonstrated the ability of the ACE to respond to a wide range of audio production creative needs. More information can be found on our companion website (Audio Commons Routledge Website 2018).

3.4 Live Coding with MIRLC

As pointed out in Section 2.4, the musical practice of *live coding* is based on improvisation and generation of code in real time (Collins et al. 2003). This can be done with several live coding environments, such as SuperCollider (McCartney 2002), a platform for audio synthesis and algorithmic composition. In live coding, the integration of music information retrieval (MIR) techniques for sound retrieval and the use of Audio Commons content have been little explored. MIRLC (see Figure 16) is a library designed to repurpose audio samples from Freesound, which can also be applied to local databases, by providing human-like queries and real-time performance capabilities (Xambó et al. 2018). The system is built within the SuperCollider environment by leveraging the Freesound API.

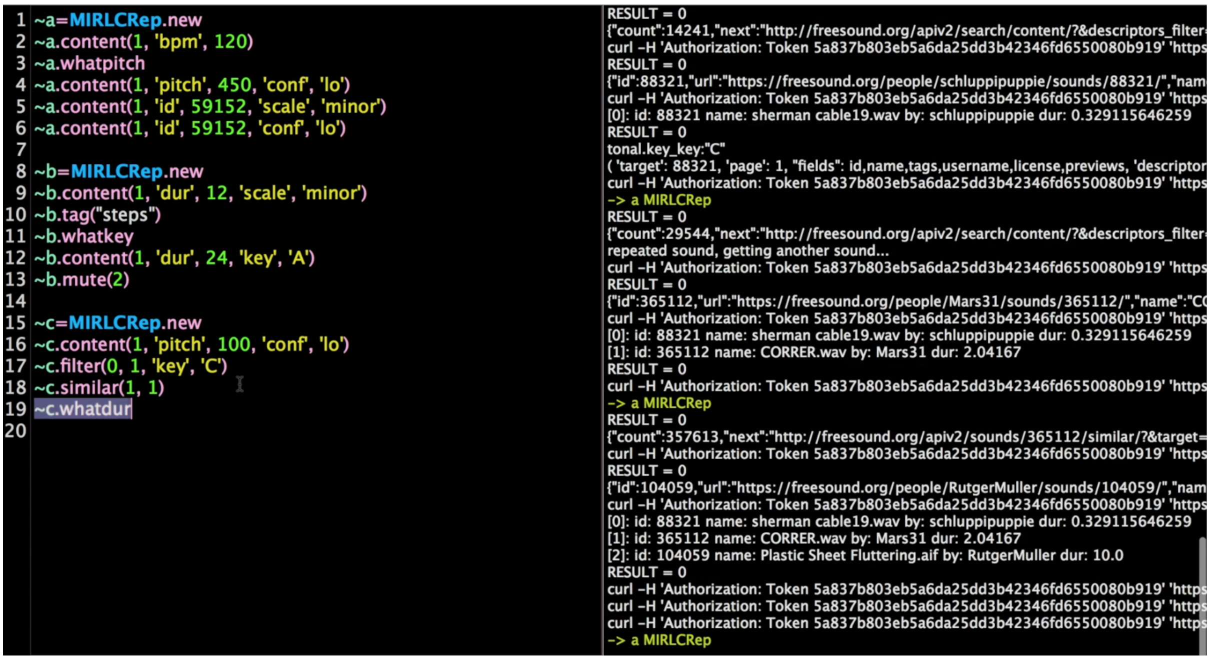

Figure 16: Screenshot illustrating how the MIRLC library can be used during a live coding session.

---

[33] Examples of soundscapes produced using CC sounds: https://soundcloud.com/qmulsrpt/sets/qmul-short-soundscapes-2017-18

The novelty of this approach lies in exploiting high-level MIR methods (e.g. query by pitch or rhythmic cues) and ACE content using live coding techniques. Both content-based (e.g. similarity) and text-based (e.g. tags) queries are possible. Sounds are loaded in user-defined groups and played in loop. Sounds of each group can be triggered either simultaneously or sequentially. This approach allows the user to load groups of related sounds and operate them with a higher level of control, which contrasts with operating single sounds. It is also possible to operate playback controls on a group of sounds (e.g. solo, mute, pause). Textual feedback is given of the different processes (e.g. queries, results).

The capacity of accessing a large amount of MIR parameters and sounds invites the live coder to explore and create subspaces of sounds through code. Feedback from four expert users reported that the tool has an interesting level of unpredictability and experimentation in the musical process, and that querying was perceived as a non-linear process, where sounds are retrieved organically following their own downloading times (Xambó et al. 2018). A credit list of the downloaded sounds is automatically created for each live coding session. A number of examples of live coding using crowdsourced online sounds are provided in the companion website (Audio Commons Routledge Website 2018).

3.5 Live Collaborative Music Making With Playsound

As discussed in Section 2.1, sounds can be described with high-level semantic attributes representing how they were generated or what they express, in lieu of or in addition to musical characteristics such as pitch or chords. The Playsound.space[34] platform (Stolfi et al. 2018) was designed using Web Audio to let users mix Audio Commons content using semantic searches without requiring specific musical knowledge. The ACE provides ways to query sounds using descriptive metadata through its API. Playsound provides a fast access to the Freesound audio content and allows users to play and loop multiple audio files with basic editing capabilities, including segment selection, panning, playback rate and volume controls. Figure 17 shows the GUI of Playsound, with the list of selected sounds to the left and the

[34] http://www.playsound.space

search interface to the right, displaying the search textbox, and the list of retrieved sound items with metadata and visual *spectrogram*[35] representations. Credits to authors of selected CC sounds are displayed at the bottom of the interface. Live interactions with Playsound can be recorded and exported.

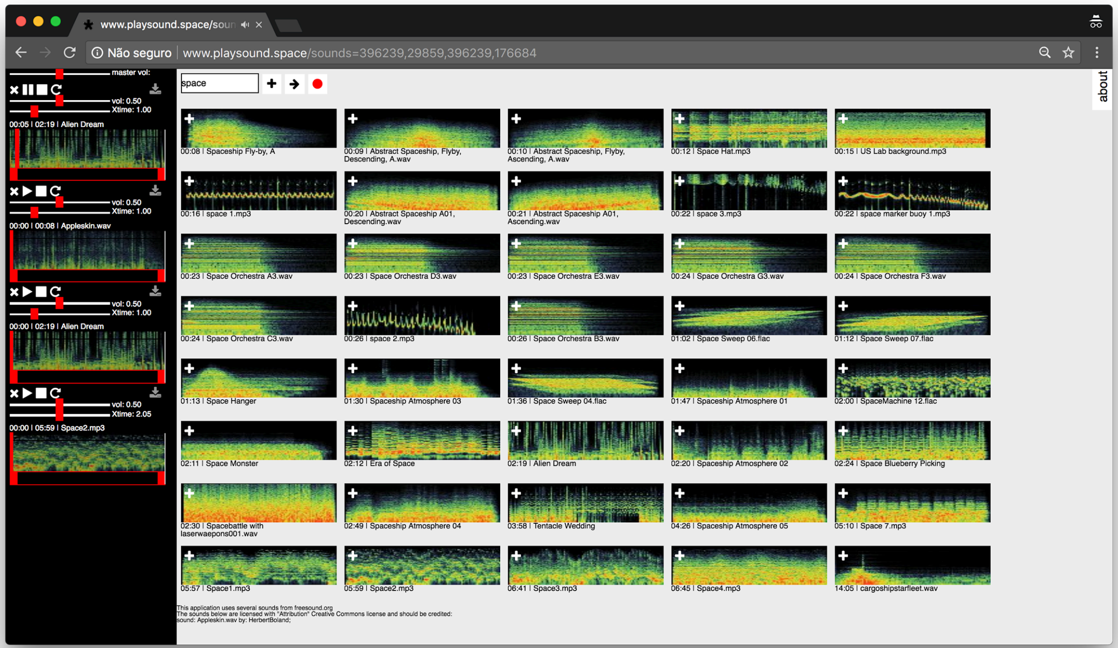

Figure 17: Graphical User Interface from the Playsound.space web application.

In Stolfi et al. (2018), the authors present a study where Playsound was used by small ensembles of laptop musicians to play free live music improvisations. Free music improvisations are not necessarily bound to predefined musical attributes and structure like in score-based compositions (e.g. key, chords, meter) and the activity emphasizes the performing process and the interaction between musicians (Bergstroem-Nielsen 2016). Results indicated that it was easy for both musicians and nonmusicians to play live collaboratively using the tool. The semantic sound search functionality facilitated verbal and nonverbal interaction between musicians and led to interesting musical situations through the use of similar or contrasting materials at different times, and rich variation of timbres and rhythms. Users were able to express “sound ideas” even without technical expertise and musical

[35] A spectrogram is a time-frequency representation of a sound indicating how the energy of frequency components evolve over time.

technique. The companion website includes several examples of productions made with Playsound (Audio Commons Routledge Website 2018).

4 Conclusion

In this chapter, we reviewed some of the salient opportunities and challenges related to the use of Audio Commons online content in linear media production. We first outlined key concepts related to the uploading, retrieval or consumption of online audio content. These key concepts include principles to describe the technical and creative characteristics of online audio content (metadata, folksonomies), technologies that enable to search for sounds and navigate collections (semantic audio, text-based queries, content-based queries, graphic-based interfaces), and licenses. We then described how Audio Commons content can be reused and repurposed either through Internet-connected DAWs or browser-based applications. We introduced the Audio Commons application programming interface (API) which enables developers to access resources from the Audio Commons Ecosystem. Finally, we illustrated through five use cases how online Audio Commons content can be leveraged by sound designers and musicians. We presented a set of audio plugins and web-based interfaces to generate novel sound textures, integrate CC sounds into compositions, create soundscapes collaboratively, and augment live music performance with audio content from the web.

As shown in the above sections, combining crowdsourced and cloud-based online audio resources with the traditional DAW is a promising approach that can enrich the creation of media content, both original or remixed. Sound designers inclined to develop creative coding skills may also learn how to design their own web-based tools leveraging CC audio ecosystems (e.g. web development, practice-based research). These new skills can be combined with the traditional skillset of the professional sound designer in innovative ways leading to novel technologies and workflows yet to be discovered. The aim of this chapter was to highlight existing methods, tools and techniques that can be seen as a starting point in this promising new domain.

References

Audio Commons Routledge Website (2018). Available at: https://www.audiocommons.org/routledge/ [Accessed 9 Jul. 2018].

Audio Commons Music and Sound Search Tool (2018). Available at: http://audiocommons.jamendo.com/ [Accessed 9 Jul. 2018].

Barthet, M., Kronland-Martinet, R. and Ystad, S. (2007). Improving musical expressiveness by time-varying brightness shaping. In *International Symposium on Computer Music Modeling and Retrieval*. Berlin/Heidelberg: Springer, pp. 313-336.

Barthet, M., Anglade, A., Fazekas, G., Kolozali, S. and Macrae, R. (2011). Music recommendation for music learning: Hotttabs, a multimedia guitar tutor. In *Workshop on Music Recommendation and Discovery*. pp. 7-13.

Barthet, M., Marston, D., Baume, C., Fazekas, G. and Sandler, M. (2013). Design and evaluation of semantic mood models for music recommendation. In: *Proceedings of the International Society for Music Information Retrieval Conference*.

Barthet, M., Fazekas, G., Allik, A., Thalmann, F. and Sandler, M. B. (2016). From interactive to adaptive mood--based music listening experiences in social or personal contexts, *Journal of Audio Engineering Society*, 64(9), pp. 673-682.

Bergstroem-Nielsen, C. (2016). Keywords in musical free improvisation. *Music and Arts in Action*, 5(1), pp. 11-18.

Bischoff, K., Firan, C. S., Nejdl, W. and Paiu, R. (2008). Can all tags be used for search? In: *ACM Conference on Information and Knowledge Management*. ACM, pp. 193–202.

Choi, K., Fazekas, G., Sandler, M. and Cho, K. (2018). The Effects of noisy labels on deep convolutional neural networks for music tagging. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(2), pp. 139-149.

Cole, S. J. (2011). The prosumer and the project studio: the battle for distinction in the field of music recording. *Sociology*, 45(3), 447-463.

Collins, N., McLean, A., Rohrhuber, J. and Ward, A. (2003). Live coding in laptop performance. *Organised Sound*, *8*(3), 321-330.

Fazekas, G., Raimond, Y., Jakobson, K. and Sandler, M. (2011). An overview of semantic web activities in the OMRAS2 project. *Journal of New Music Research special issue on Music Informatics and the OMRAS2 Project*, 39(4), pp. 295-311.

Fazekas, G., Barthet, M. and Sandler, M. (2013). The BBC Desktop Jukebox Music recommendation system: a large scale trial with professional users. In: *IEEE International Conference on Multimedia and Expo Workshops*, pp. 1-2.

Font, F. and Bandiera G. (2017). Freesound Explorer: make music while discovering Freesound! In: *Proceedings of the Web Audio Conference*. London.

Font, F., Brookes, T., Fazekas, G., Guerber, M., Burthe, A. L., Plans, D., Plumbley, M. D., Shaashua, M., Wang, W. and Serra, X. (2016). Audio Commons: bringing Creative Commons audio content to the creative industries. In: *Proceedings of the 61st AES International Conference: Audio for Games*.

Font, F., Roma, G. and Serra, X. (2017). Sound sharing and retrieval. In T. Virtanen, M. D. Plumbley and D. Ellis, eds., *Computational Analysis of Sound Scenes and Events*. Cham, Switzerland: Springer International Publishing, pp. 279–301.

Gruber, T. (2008). Collective knowledge systems: where the social web meets the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(1), pp. 4-13.

Holmquist, L. E. (2008). Bootlegging: multidisciplinary brainstorming with cut-ups. In: *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008*. Indiana University, pp. 158-161.

Huber, D. M. and Runstein, R. E. (2013). *Modern recording techniques*, New York: Focal Press.

Katz, B. (2003). *Mastering audio: the art and the science*, Oxford: Butterworth-Heinemann.

Lamere, P. (2008). Social tagging and music information retrieval. *Journal of New Music Research*, 37(2), pp. 101-114.

Lee, J. A. and Downie, J. A. (2004). Survey of music information needs, uses, and seeking behaviors: preliminary findings. In: *Proceedings of the 5th International Society for Music Information Retrieval Conference*.

Lessig L. (2000). Copyright's first amendment. *UCLA Law Review*, 48,pp. 1057-1073.

Lessig, L. (2004). *Free culture: how big media uses technology and the law to lock down culture and control creativity*, New York: The Penguin Press.

Lind, F. and MacPherson, A. (2017). Soundtrap: a collaborative music studio with web audio. In: *Proceedings of 3rd Web Audio Conference*. London.

Lu, L., Wenyin, L. and Zhang, H. J. (2004). Audio textures: theory and applications. *IEEE Transactions on Speech and Audio Processing*, 12(2), pp. 156-167.

Marcell, M. M., Borella, D., Greene, M., Kerr, E. and Rogers, S. (2001). Confrontation naming of environmental sounds. *Journal of Clinical and Experimental Neuropsychology*, 22(6), pp. 830-864.

McCartney, J. (2002). Rethinking the computer music language: SuperCollider. *Computer Music Journal*, *26*(4), pp. 61-68.

Merkley, R. (2015). State of the Commons. Available at:
https://apo.org.au/node/60681/ [Accessed 9 Jul. 2018].

Owsinski, B. (2013). *The mixing engineer's handbook*.
Boston, MA: Nelson Education.

Pigrem, J. and Barthet, M. (2017). Datascaping: data sonification as a narrative device in soundscape composition. In: *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences*. New York: ACM, pp. 43:1--43:8.

Ritzer, G. and Jurgenson, N. (2010). Production, consumption, prosumption: the nature of capitalism in the age of the digital ‘prosumer’. *Journal of Consumer Culture*, 10(1), pp. 13-36.

Russell, J. A. (1980). A circumplex model of affect. *Personality and Social Psychology*, 39:6, pp. 1161-1178.

Saint-Arnaud, N. and Popat, K. (1995). Analysis and synthesis of sound textures. In: *Readings in Computational Auditory Scene Analysis*. pp. 293-308.

Schafer, R. M. (1993). *The soundscape: Our sonic environment and the tuning of the world*. Rochester, VT: Inner Traditions/Bear & Co.

Schwarz, D. (2007). Corpus-based concatenative synthesis. *IEEE Signal Processing Magazine*, 24(2), pp. 92-104.

Shadbolt, N., Berners-Lee, T. and Hall, W. (2006). The semantic web revisited. *IEEE Intelligent Systems*, 21(3), pp. 96-101.

Smith, T. (2009). The social media revolution. *International Journal of Market Research*, 51(4), pp. 559-561.

Stables, R., De Man, B., Enderby, S., Reiss, J.D., Fazekas, G. and Wilmering, T., (2016). Semantic description of timbral transformations in music production. In: *Proceedings of the 2016 ACM on Multimedia Conference*. ACM, pp. 337-341.

Stallman, R. (2002). *Free software, free society:selected essays of Richard M. Stallman*, Boston: GNU Press.

Stolfi, A. de Souza, Ceriani, M., Turchet, L. and Barthet, M. (2018). Playsound.space: inclusive free music improvisations using Audio Commons. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. pp. 228-233.

Strobl, G., Eckel, G., Rocchesso, D. and Le Grazie, S. (2006). Sound texture modeling: a survey. In: *Proceedings of the 2006 Sound and Music Computing International Conference*.

Truax, B. 2002. Genres and techniques of soundscape composition as developed at Simon Fraser University. *Organised Sound*, 7(1), pp. 5-14.

Xambó, A., Roma, G., Lerch, A., Barthet, M. and Fakekas, G. (2018). Live repurposing of sounds: MIR explorations with personal and crowdsourced databases. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. pp. 364-369.

Author biographies

Anna Xambó is a postdoctoral researcher at the Centre for Digital Music (QMUL, UK) and was previously a postdoctoral researcher at the Center for Music Technology and Digital Media Program (Georgia Tech, USA). She has been co-founder and co-chair of Women in Music Tech at Georgia Tech (2016-2017). She was a researcher at the London Knowledge Lab (UCL, UK) during 2013-2014. Her PhD dissertation is entitled “Tabletop Tangible Interfaces for Music Performance: Design and Evaluation" (2015, The Open University, UK). Anna's research focuses on interactive real-time systems for music investigating live coding and generative computer music, collaborative and participatory interfaces, and real-time MIR.

Frederic Font is a postdoctoral researcher at the Music Technology Group of the Department of Information and Communication Technologies of Universitat Pompeu Fabra, Barcelona. His current research is focused on facilitating the reuse of audio content in music creation and audio production contexts. Alongside his research, Frederic is leading the development of the Freesound website and coordinating the EU funded Audio Commons Initiative.

George Fazekas is a Lecturer in Digital Media at the Centre for Digital Music, Queen Mary University of London, where he obtained a PhD in Semantic Audio. He is leading the QMUL team for the EU-funded project Audio Commons as Principal Investigator. He was Co-Investigator of five collaborative projects, programme co-chair of the International Workshop on Semantic Applications for Audio and Music (SAAM2018), general chair of the International Audio Mostly conference (ACM-SIGCHI, 2017) and paper co-chair of the AES 53rd International Conference on Semantic Audio (2014). He is guest editor of two JAES issues and published over 100 peer-reviewed papers related to Semantic Audio.

Dr Mathieu Barthet is a Lecturer in Digital Media at the Centre for Digital Music at Queen Mary University of London and Technical Director of the qMedia Studios. He was awarded a PhD in Acoustics, Signal Processing and Computer Science applied to Music from Aix-Marseille University and CNRS-LMA in 2008. His research lies at the confluence of several disciplines including Music Information Retrieval, Interaction and Perception. He is Principal Investigator of the EU-funded project ‘Towards the Internet of Musical Things’ and Co-Investigator of the EU-funded project ‘Audio Commons’. He was chair of several audio conferences (CMMR 2012, AM’17) and serves as guest editor for the Journal of the Audio Engineering Society.