



## Deliverable D2.1

### Requirements report and use cases

<b>Grant agreement nr</b>	688382
<b>Project full title</b>	Audio Commons: An Ecosystem for Creative Reuse of Audio Content
<b>Project acronym</b>	AudioCommons
<b>Project duration</b>	36 Months (February 2016 - January 2019)
<b>Work package</b>	WP2
<b>Due date</b>	15 June 2016 (M3)
<b>Submission date</b>	15 June 2016 (M3)
<b>Report availability</b>	Public (X), Confidential ( )
<b>Task leader</b>	QMUL
<b>Authors</b>	Mathieu Barthet, Gyorgy Fazekas, Damir Juric, Johan Pauwels, Mark Sandler and Lasse Vetter
<b>Document status</b>	Final





# Table of contents

<a href="#">Executive Summary</a>	4
<a href="#">Background</a>	6
<a href="#">1 Introduction</a>	7
<a href="#">1.1 Main objectives and goals</a>	7
<a href="#">1.2 Terminology</a>	7
<a href="#">2 Requirements and use cases</a>	9
<a href="#">2.1 Demographic</a>	9
<a href="#">2.2 Workflow</a>	13
<a href="#">2.3 User groups</a>	19
<a href="#">2.4 Search and metadata</a>	25
<a href="#">2.4.1 Search strategies</a>	25
<a href="#">2.4.2 Building Audio Commons Vocabulary</a>	30
<a href="#">2.5 Current problems and User Wishlist</a>	32
<a href="#">2.5.1 Current problems with audio content browsing and retrieving</a>	32
<a href="#">2.5.2 User Wishlist</a>	33
<a href="#">3 Ontology and Metadata</a>	34
<a href="#">3.1 Metadata</a>	34
<a href="#">3.2 Ontology</a>	35
<a href="#">4 Conclusion</a>	39
<a href="#">5 References</a>	40
<a href="#">APPENDIX 1: User requirements for metadata</a>	41





## **AudioCommons**

D2.1 Requirements report and use cases

<a href="#">APPENDIX 2: Most common terms</a>	54
<a href="#">APPENDIX 3: Audio Commons Initiative Survey</a>	56



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 688382



## Executive Summary

This deliverable is part of WP2, a work package that will describe the Audio Commons Ontology and Audio Commons API specification. Building the Audio Commons Ontology and Audio Commons API specification will require gathering the knowledge from existing datasets in the music domain and knowledge about the workflows and tasks that people who work in our target domains use in their everyday work. By creating the survey that targeted different groups of people from the music industry we collected the user requirements considering the vocabulary (metadata) and workflows.

Requirement analysis should uncover the current state-of-the-art of our target domains and direct us towards the path for improvement. Our goal is to successfully integrate Audio Commons content in the existing workflows and collect and formalize the knowledge about information seeking activities and learn more about the use of audio and music content in creative processes.

The task will inform the definition of the Audio Commons Ontology and API, inform research on rights management and help focusing the work on sound and music analysis algorithms and end user prototypes.

The survey contained 24 questions asking creatives working in music industry about various subjects like demographics, workflows they use and metadata they would like to use when searching for new audio content on the Web. By analysing the demographic of the participants of the survey we discovered that most of them are professionals with more than 10 years of experience in the music domain, especially in the field of music production. Questions regarding workflows asked users about the characteristics of the machines they use (connected to Internet or not), about the amount of data they use in their daily tasks, source and type of the data and the time they invest in working with the data. The final set of questions in the survey were about the search strategies and metadata associated with the audio files. Answers to those questions allowed us to get the insight into how users would like to search for specific files and how such strategy would impact the design of the user interfaces. Also, data collected about metadata requirements (Musical Properties, Perceptual Properties, General Metadata) should be used in later deliverables concerning the ontology building (T.2.2).

For this deliverable we analysed which ontologies and vocabularies exist that describe the music domain. Our conclusion is that we should follow the guidelines of Semantic Web and reuse as much of existing data as possible (ontologies and vocabularies like Music ontology, Musical Instrument Taxonomies, Media Value Chain Ontology, Audio features ontology, Studio Ontology).

The analysis of the data collected from the survey raised the need for knowledge about different categories within Audio Commons project. Those categories range from Sound properties (e.g. Bass, Rhythm, timbre), Software (e.g. Effects, VST, audio format), Hardware (e.g. audio channels, equipment), various measures (e.g. bit-depth, tempo, bit rate, bpm, loudness, sample rate), Genres (e.g. classical music, dubstep, new age, vocal music), Instruments (e.g. drums, electric piano, guitar, hi-hat, percussion, piano, violin), Music theory, Musical texture (e.g. monophony, polyphony), Techniques (e.g. slide guitar, distortion), Sound character (e.g. urban, rural, cinematic, retro, etc), Music representation (e.g. audio histogram, atonality), Moods (e.g. dark, faded, bright, meditative, happy, etc.), Arrangement/Orchestration (like orchestra), Legal document (licence and rights management).

The survey showed that audio community faces multiple problems when retrieving audio content from the internet or just working with that kind of content in the studios or at home. Those problems are





## AudioCommons

D2.1 Requirements report and use cases

related with licensing (not clear enough, hard to understand the rules), syntax (problematic labeling of audio content), sparseness of metadata, lack of workflow integrations (easily retrieving the data into some part of the workflow), bad recording quality of audio sources, various interface problems (bad design, popups, redirections, etc.) and lack of quality curation/recommendation.

According to the results of our survey, what users want is clear and understandable licensing information, intelligent interfaces with drop down functionalities straight into their workflows, high quality recommendation, rich metadata describing the audio content and availability of services that are capable of conducting various tasks in the audio domain.





## Background

This deliverable is part of WP2 (Task T.2.1). One of the main goals of WP2 is the description of the Audio Commons Ontology. Building the Audio Commons Ontology (and ontologies in general) requires an extensive knowledge about the domain that ontology will describe. Knowledge about the domain includes the vocabulary of the domain and knowledge about the workflows (processes) that are being carried out by various roles involved in them. Creating a survey and asking relevant people (people who work in music domain) to answer on questions about the daily tasks can give us an insight into the complex network of entities and tasks that are being carried out in the music domain. Those entities and tasks have to be recognized when analyzing the data produced by the survey and mapped into ontology. One of the deliverables of work package WP2 is to provide the Audio Commons API specification, which will define how the different components of the Audio Commons Ecosystem will be technically interconnected (i.e., how production tools will be able to access Audio Commons content and users will be able to communicate for the licensing process) and will also include the orchestration of the different services or components integrating the Audio Commons Ecosystem, providing the required technology layer for the interconnection of the different components, and providing the required software packages and guidelines to facilitate the incorporation of new actors in the Audio Commons Ecosystem. Knowledge about which services are actually needed and how to orchestrate them in the Audio Commons Ecosystem will come from the analysis of the user requirements presented in this deliverable.





# 1 Introduction

## 1.1 Main objectives and goals

To successfully build the Audio Commons Ontology and Audio Commons API specification, extensive knowledge about the domain is required. Knowledge that needs to be collected can be grouped as:

- Knowledge about the vocabulary (metadata)
- Knowledge about the tasks in our target domains (such as game audio or sound design for cinema)
- Knowledge about the existing tools and service that are being used in our target domains

The main goal of this deliverable is to describe and analyze the knowledge collected from within the consortium (and partly from outside the consortium) that has been collected through an online survey.

Requirement analysis should uncover the current state-of-the-art of our target domains and direct us towards the path for improvement. Our goal is to successfully integrate Audio Commons content in the existing workflows and collect and formalize the knowledge about information seeking activities and learn more about the use of audio and music content in creative processes.

This deliverable will inform the definition of the Audio Commons Ontology and API, inform research on rights management and help focusing the work on sound and music analysis algorithms and end user prototypes.

## 1.2 Terminology

**AudioCommons:** reference to the EC H2020 funded project AudioCommons, with grant agreement nr 688382.

**Audio Commons Initiative:** understanding of the AudioCommons project core ideas beyond the lifetime and specific scope of the funded project. The term "Audio Commons Initiative" is used to imply i) our will to continue supporting the Audio Commons Ecosystem and its ideas after the lifetime of the funded project, and ii) our will to engage new stakeholders which are not officially part of the project consortium.

**Audio Commons:** generic reference to the Audio Commons core ideas, without distinguishing between the concept of the initiative and the actual funded project.

**Audio Commons Ecosystem (ACE):** series of technologies and actors involved in publishing and consuming Audio Commons content.





## AudioCommons

### D2.1 Requirements report and use cases

**Audio Commons content (AC):** audio content released under Creative Commons licenses and enhanced with meaningful contextual information (e.g., annotations, license information) that enables its publication in the ACE.

**Content creator:** individual users, industries or other actors that create audio content and publish in the ACE through content providers.

**Content provider:** services that expose content created by content creators to the ACE.

**Content user:** individual users, industries or other actors that use the content exposed by content providers and created by content creators in their creative workflows.

**Ontology:** In the context of computer and information sciences, ontology defines a set of representational primitives with which to model a domain of knowledge or discourse. The representational primitives are typically classes (or sets), attributes (or properties), and relationships (or relations among class members). The definitions of the representational primitives include information about their meaning and constraints on their logically consistent application. In the context of database systems, ontology can be viewed as a level of abstraction of data models, analogous to hierarchical and relational models, but intended for modelling knowledge about individuals, their attributes, and their relationships to other individuals. Ontologies are typically specified in languages that allow abstraction away from data structures and implementation strategies; in practice, the languages of ontologies are closer in expressive power to first-order logic than languages used to model databases. [Gruber]

**Tool developer:** individual users, industries or other actors that develop tools for consuming (and also potentially publishing) Audio Commons content.

**Embeddable tools:** tools for consuming Audio Commons content that can be embedded in existing production workflows of creative industries.





## 2 Requirements and use cases

The following sections will present the results of the survey. At the time of writing this section we had 193 responses. Survey contained 24 questions (15 questions with predefined answers and 9 open ended questions). The complete survey can be found in Appendix 3.

Some of the questions/themes that this survey can provide deeper insight are:

- User workflow, e.g., stand-alone tools or as plugins?
- How people wish to search for content?
- Audience: Pro/Amateurs?, Sound designer/musicians/gaming?
- Interested in: Sounds/Loops/Songs?
- Current methods and/or sources of content?
- What limitations they see: What are the problems with the current methods?
- What ideas they have: What are the most important features they can think of?

Audio Commons industry partners were given the task to ask their user base to fill in the survey.

### 2.1 Demographic

This section of the deliverable will inform about the target demographic of the potential Audio Commons Ecosystem users. Standard demographic data can contain the information about the age range, gender, affiliation or experience:

- **Experience level:** users were asked if they consider themselves professionals or amateurs. As Fig. 1 is showing we can expect to have more users who consider themselves professionals than amateurs (62.7% vs 37.3%).



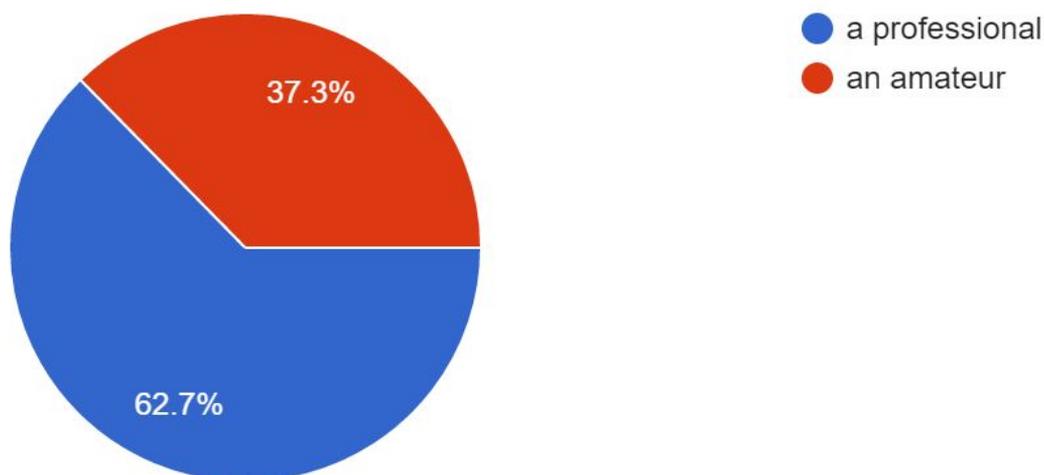


Fig. 1: Experience level of the potential users of Audio Commons Ecosystem

- **Years of experience:** the majority of people taking part in this survey have more than 10 years of experience of working with audio (55.8 %)

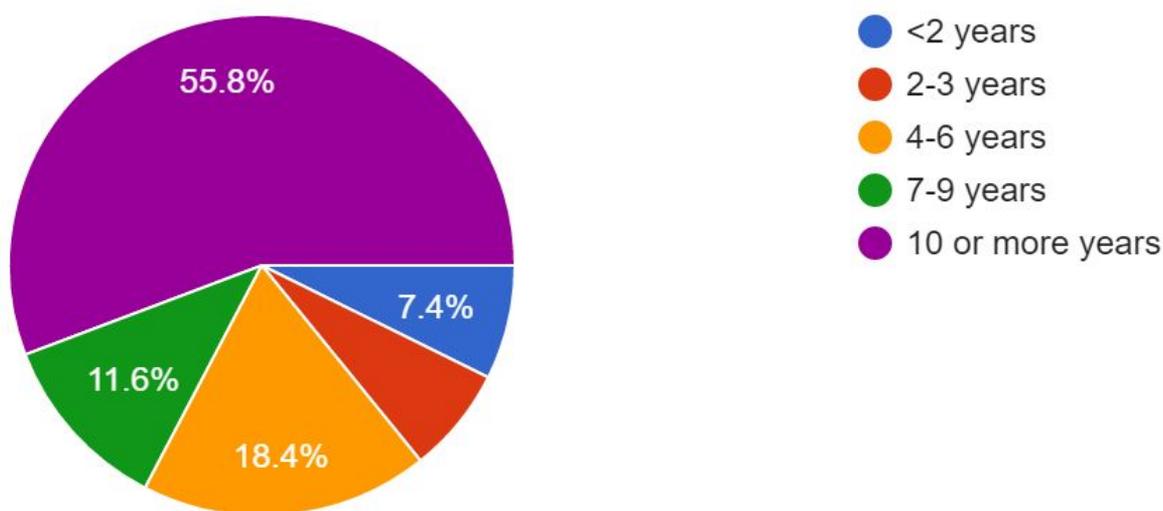


Fig. 2: Years of experience





Although a bit more than one third of the potential users are amateurs, the results show that 40% of those have 10 or more years of experience. This sums up to 75.0% of responders are either very experienced amateurs or they consider themselves professionals.

It is therefore a reasonable assumption that these have a good knowledge of their craft and that they're searching for high quality files.

Derived requirements:

1. Some measure of quality. This can be based on simple measures such as sample rate, clipping, and alike. Additionally it could be based on more sophisticated measures like dynamic range and similar perceptual descriptors.
2. Leveraging the experience: E.g., making the ratings for tracks a weighted sum of ratings from users. The weights could be determined by the users' experience and the track ratings she got for her own content, if she has any uploads. By that we'd push it more towards a professional platform, which reflect around 4/5 of our potential users (see above)
3. Point 2 above implies that the AC eco-system has its own user base. This would be necessary to keep track of user ratings.

- **Affiliation:** responders to our survey are working in the fields of: (Fig. 3):
  - Music production (mixing, mastering)
  - Music production (composition, recording)
  - Audio for games
  - Audio for movies and/or TV
  - Music performance
  - Other

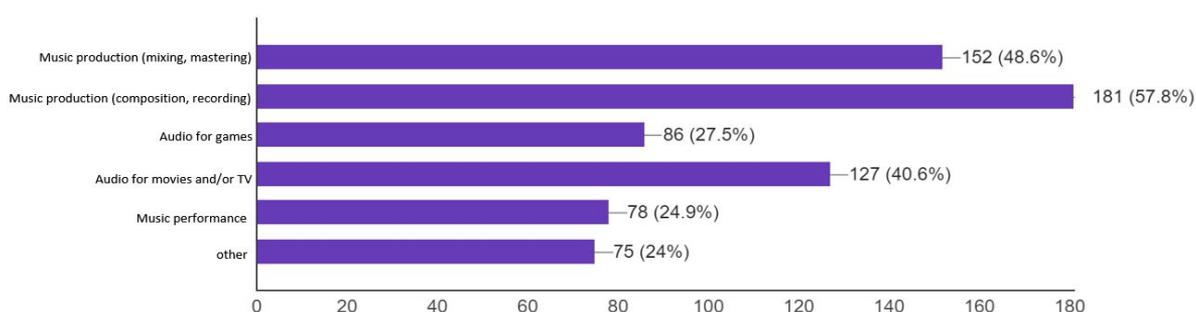


Fig. 3: Affiliation of potential users





Most of users that responded to our survey work in music production (tasks as composition, recording, mixing and mastering). Significant amount of users works with tasks connected with audio and gaming. Around one third of the users conduct tasks connected with music performance.

- **Gender:** results are showing that participants of the survey were mostly males (93.1%)

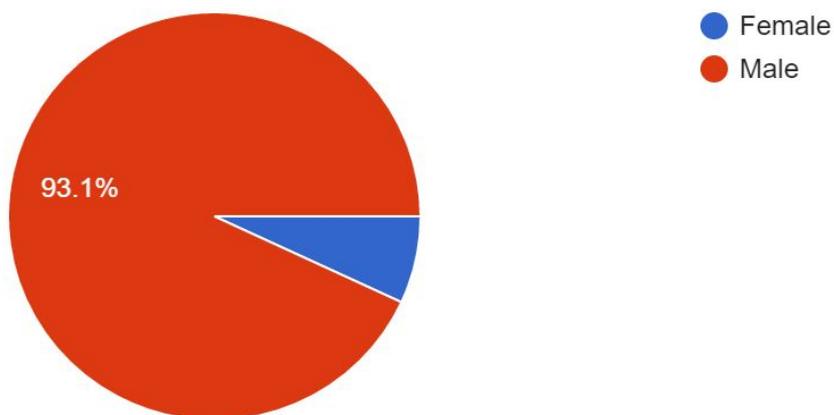


Fig. 4: Gender distribution

- **Age:** Results on Fig. 5 are showing the percentage of potential users in certain age ranges.

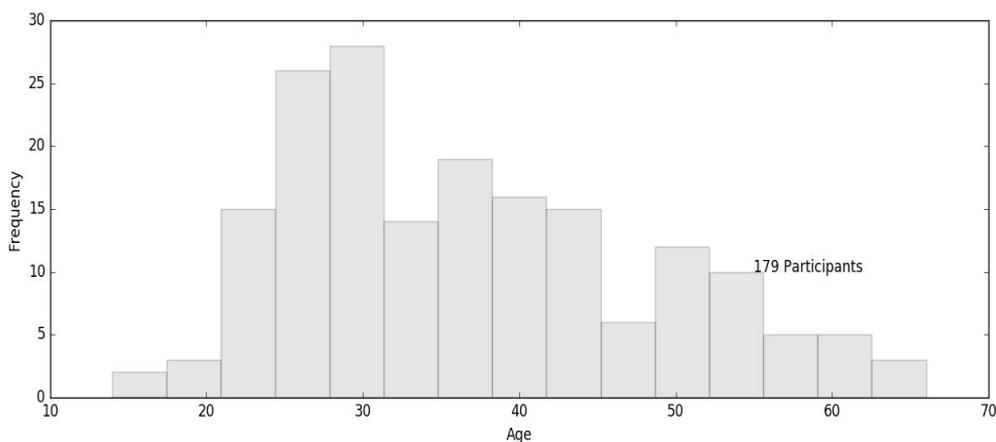


Fig. 5: Age histogram





**Conclusion:** Users that answered on our survey are mostly male (93.1 %) and in the age range between 20 - 40. They are mostly professionals and most of them are involved in the task of music production.

## 2.2 Workflow

In this section we analyse the answers from the survey regarding users' workflows (in the studio or at home). We asked users about the gear/equipment they use, about the amount of data they use in their daily tasks, source and type of the data and the time they invest in working with the data.

- **Connection:**

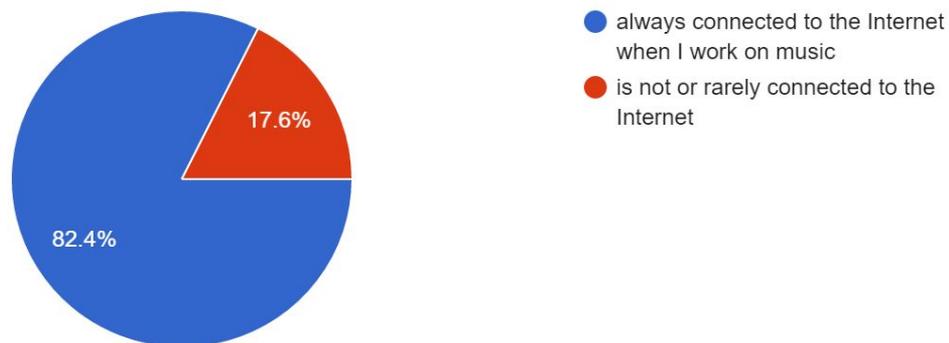


Fig. 6: Availability of the Internet connection

For approximately 18% of our potential users we can expect that they will work offline. 28.5% of those further reported to never use content from web databases. This means that almost 53% of users don't have Internet connection or never used the Internet to search for music. One potential way to get offline working people engaged would be to build a product, which lets one configure a custom sound library with a web interface, which is then downloadable as a whole library.

Requirements:

- Intelligence that can use a broad query of multiple keywords into finding a set of audio files. E.g., user enters: I need 100 audio files. Genre: Blues. Only Drums, Bass and Guitars.
- An entity that buffers all the selected files, compresses them and makes them downloadable as a whole package, so users can conveniently transfer them to their workstation.





- **Data size:** users were asked about the amount of the data they usually use in their daily tasks.

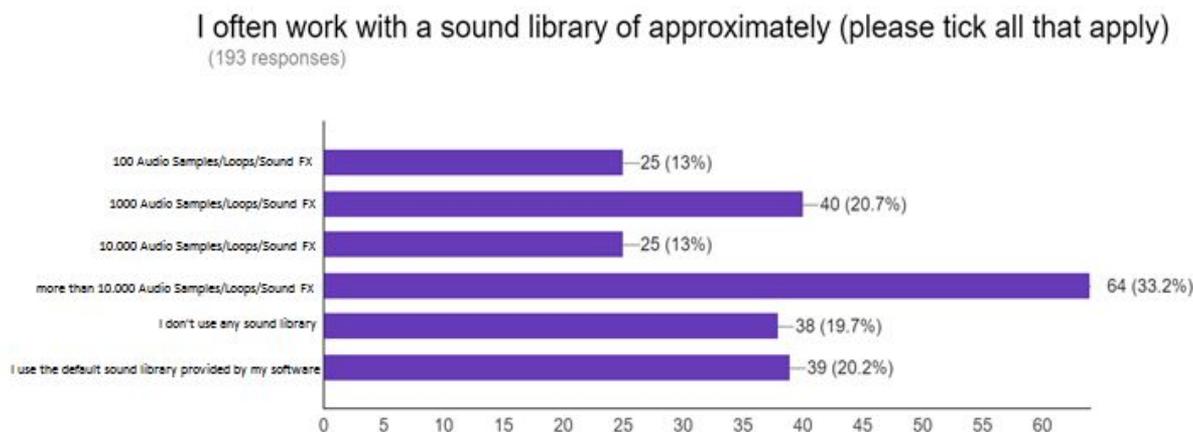


Fig. 7: The size of the libraries that user use in their tasks

From Fig. 7 we can see that most users work with quite large datasets containing more than 10,000 audio samples/loops/sound FX.

- **Audio from the Web:** when asked if they use audio files from web databases, a significant number of users (28.5 %) answered that they never use web databases.

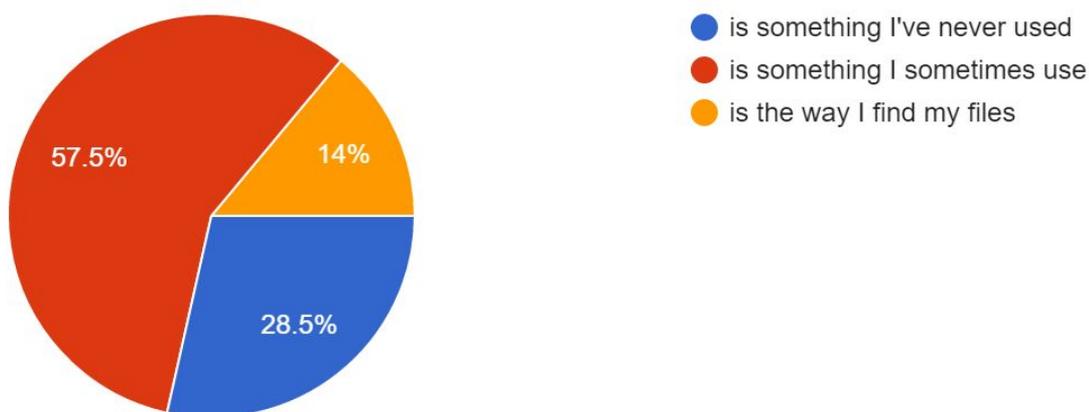


Fig. 8: Usage of web databases

- **Workload:** We asked users what they spend most of their time on. As shown in Fig 9, users spend a lot of time modifying the files with audio processors. Nevertheless, almost 40% of





users spend most of their time searching for files. Therefore, improving search workflows can have a great impact on production.

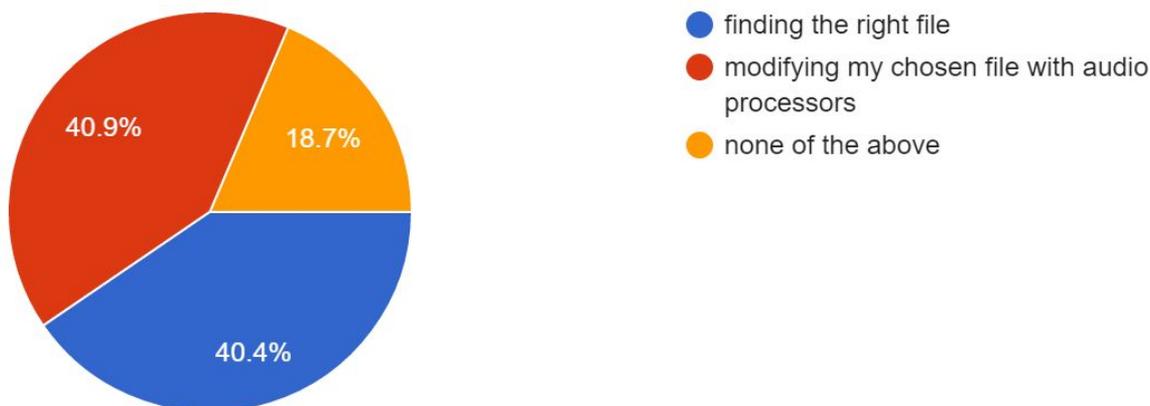


Fig. 9: Workload

- **Processing:** as established previously, 41% of users spend most of their time processing and transforming chosen audio files. The large majority of them find that task a part of their creative work, which was expected. A unique selling point could therefore be to enable users to not only find files, but also let them process them on the fly in terms of EQ, pitch, and dynamics and then download a rendered version.

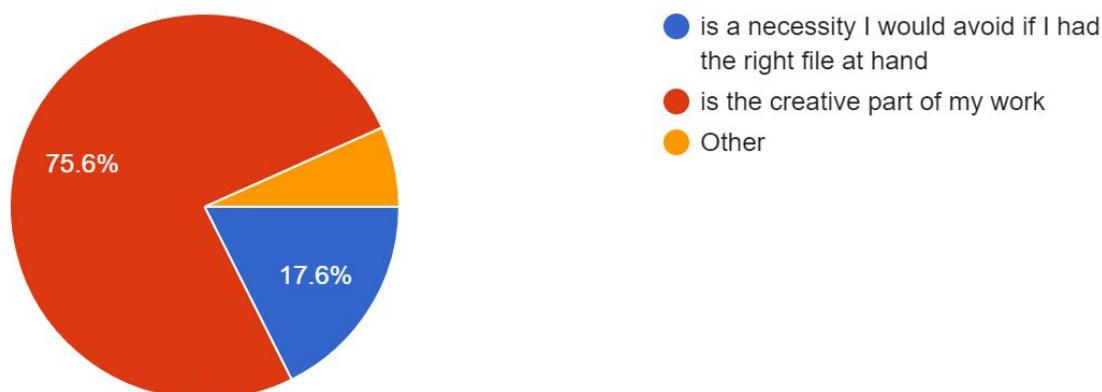


Fig. 10: How do users feel about processing of audio files





- **Type of data:** when asked what kind of files they mostly use in their daily tasks, the largest percentage of users chose sound effects (67.4 %) and audio loops (44 %). Only 19.7 % of them use the whole songs.

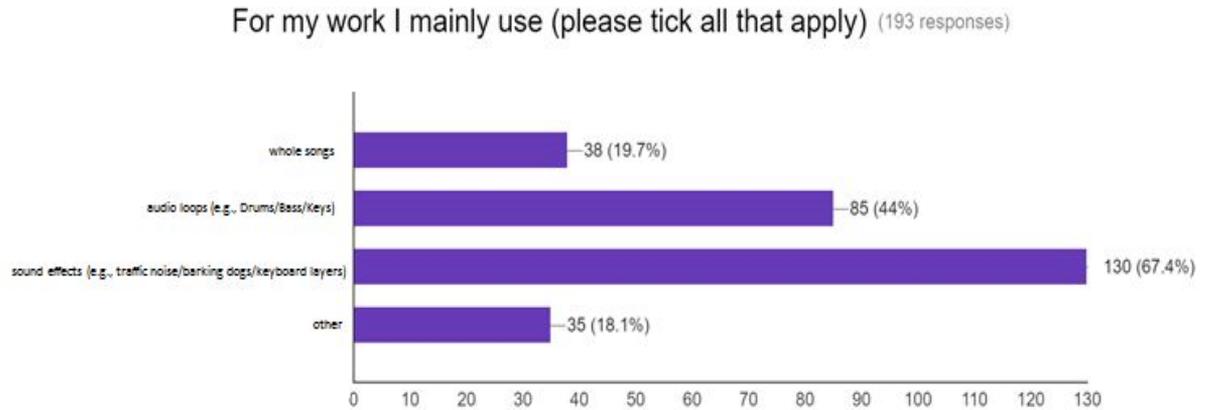


Fig. 11: Types of files users mostly use

**Scenario:** Typical use case is shown on Fig 12. Users mostly work on machines with an internet connection. They usually use software that comes with the default music library but sometimes they acquire audio from a music web database. Audio files they most often use are sound effects. Task that takes most of the time in the process workflow is modifying the file with audio processors.



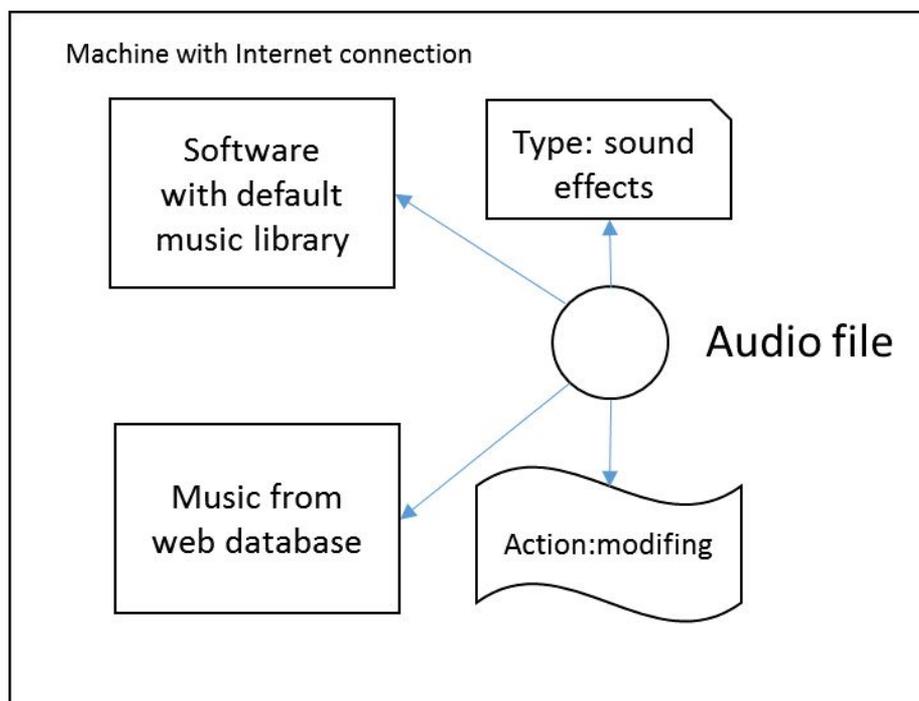


Fig. 12: User typical use case scenario based on survey

From the analysis done above, we now derive four user stories. These user stories aim to describe potential users and their requirements in a compact and easy to understand way.

### User story 1

As a cafe owner I would like to search for whole songs, which are free of any licensing fees. I would like to search via a browser and search with free text search. For an example, I'd like to search for: "Slow funk track without vocals". Once I found something I like, I would like to find tracks that play well together.

### User story 2

As an audio producer, I would like to have access to a 10,000+ set of high-quality audio loops from within my DAW. I imagine a plugin-based search interface like the one depicted in Fig. 13. I want to search by instrument type, genre, key, tempo. I only want high-quality files.

### User story 3

As a game sound designer, I would like to have access to an unlimited set of high-quality audio files from within my DAW. I want to search by effect type, mood, and some perceptual features like "warm", "bright", etc. Low quality audio files are sometimes exactly what I am after as a special effect but in general I need high quality files. The system could look like the one outlined in Fig. 13.

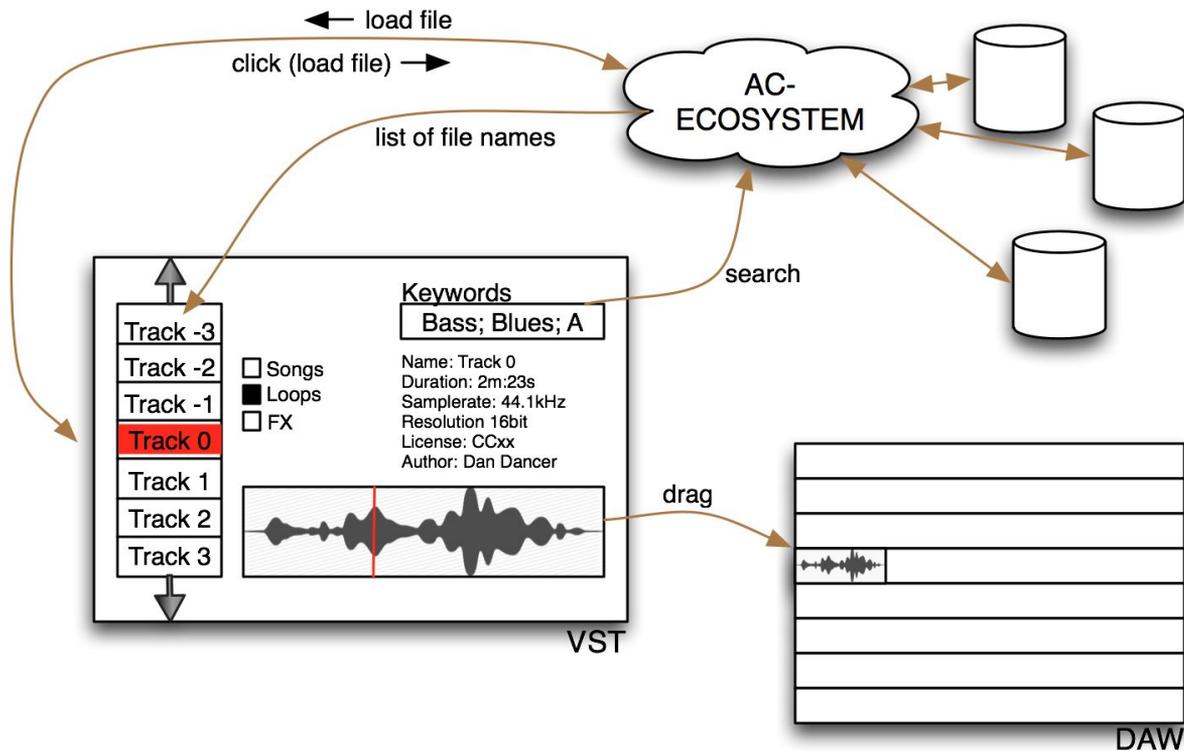


Fig. 13: Schematical overview of a an audio plugin, which enables browsing audio files from within a DAW.

#### User story 4

As a professional I never have my workstation connected to the internet. I still would like to use CC licensed material, but most web-based interfaces only allow me to search for and download single files, which is unpractical for me. I would therefore like to have a standalone app, which lets me define a set of things I am generally interested in, such as, “genre”, “mood”, production style, and a number of files I would like to have. I would then like to download the results in form of a library and use the same standalone app on my offline workstation to search for single files out of the library I’ve downloaded.





## 2.3 User groups

In this section we will present some specificities of the user groups identified by what they are working on. Large number of participants of this survey are working on a combination of things (e.g. music production and audio for games and movies/TV). That makes it hard to point to some specific requirement of a specific group because we couldn't know if that requirement is connected with music production of audio for gaming domain for example. That's why we will focus only on groups of people doing in specific field like:

- Audio for games
- Audio for movies/TV
- Music production (composition, recording)
- Music production (mixing, mastering)

Also we identified the groups that are working with a:

- Whole songs
- Sound effects
- Audio loops

### ***What I mainly work on is (please tick all that apply)***

Audio for games - 14

Audio for games;Audio for movies and/or TV - 14

Audio for games;Music performance - 2

Audio for movies and/or TV - 21

Music performance - 3

Music production (composition, recording) - 16

Music production (composition, recording);Audio for games - 3

Music production (composition, recording);Audio for games;Audio for movies and/or TV - 4

Music production (composition, recording);Audio for movies and/or TV - 11

Music production (composition, recording);Music performance - 4

Music production (mixing, mastering) - 11

Music production (mixing, mastering);Audio for movies and/or TV - 3

Music production (mixing, mastering);Music production (composition, recording) - 33

Music production (mixing, mastering);Music production (composition, recording);Audio for games - 5

Music production (mixing, mastering);Music production (composition, recording);Audio for games;

Audio for movies and/or TV - 6

Music production (mixing, mastering);Music production (composition, recording);Audio for games;Audio for movies and/or TV;Music performance - 9

Music production (mixing, mastering);Music production (composition, recording);Audio for games;Music performance - 3

Music production (mixing, mastering);Music production (composition, recording);Audio for movies and/or TV - 19

Music production (mixing, mastering);Music production (composition, recording);Audio for movies and/or TV;

Music performance - 11

Music production (mixing, mastering);Music production (composition, recording);Music performance - 25

### Audio for games:





## AudioCommons

### D2.1 Requirements report and use cases

There are 14 participants in this group. People that are part of this group are mostly using the computers that are connected to the Internet. They definitely prefer to search and retrieve audio through the web browser using google style search (keywords). Also they mostly work with sound effects.

#### Special requirements:

- preview samples across every increment of the waveform
- favourite system ( a way to keep track of locally stored files)
- more filters for finding the sounds I want
- an advanced search form
- fast database scan
- intelligent metadata editor

***My computer for making sound and/or music is always connected to the Internet when I work on music - 12***  
is not or rarely connected to the Internet - 2

---

#### ***I would prefer to search for and retrieve audio (please tick all that apply)***

No preferences: my DAW (Reaper) has this feature - 1  
from within my Digital Audio Workstation - 2  
no preference - 1  
through a standalone app - 1  
**through a web browser - 9**

---

#### ***When I search for music-related data I want to (please tick all that apply)***

query using suggested keywords from the interface(Drop Down Lists or similar);use keywords(Google style) - 2  
query using suggested keywords from the interface (Drop Down Lists or similar);use keywords (Google style);  
use natural language - 1  
use a query language (like programming) - 1  
**use keywords (Google style) - 7**  
use keywords (Google style);use a query language (like programming) - 2  
use natural language (for instance, whole sentences) - 1

---

#### ***For my work I mainly use (please tick all that apply)***

audio loops (e.g., Drums/Bass/Keys);sound effects (e.g., traffic noise/barking dogs/keyboard layers) - 3  
**sound effects (e.g., traffic noise/barking dogs/keyboard layers) - 8**  
whole songs;audio loops (e.g., Drums/Bass/Keys);sound effects (e.g., traffic noise/barking dogs/keyboard layers) - 3

### Audio for movies/TV:

There are 21 participants in this group. People that are part of this group are mostly using the computers that are connected to the Internet. Large number of participants would prefer to search and retrieve audio through a standalone app. Also they mostly work with sound effects and use audio content from web databases.

#### Special requirements:

- similarity based on an analysis waveforms
- a good search system
- favourites system
- Quick preview feature
- autocorrect, autocomplete
- search both online and local content.





## AudioCommons

### D2.1 Requirements report and use cases

- drag and drop to DAW

***My computer for making sound and/or music is always connected to the Internet when I work on music - 17***

is not or rarely connected to the Internet - 4

---

***For my work I mainly use (please tick all that apply)***

my imagination - 1

**sound effects (e.g., traffic noise/barking dogs/keyboard layers) - 17**

whole songs;audio loops (e.g., Drums/Bass/Keys);sound effects (e.g., traffic noise/barking dogs/keyboard layers) - 2

whole songs;sound effects (e.g., traffic noise/barking dogs/keyboard layers) - 1

---

***Audio content from web databases***

**is something I sometimes use - 13**

is something I've never used - 4

**is the way I find my files - 4**

---

***I would prefer to search for and retrieve audio (please tick all that apply)***

I think that real music is not pattern of audio samples - 1

from within my Digital Audio Workstation - 2

from within my Digital Audio Workstation;through a standalone app - 2

from within my Digital Audio Workstation;through a standalone app;through a web browser - 2

**through a standalone app - 6**

through a standalone app;through a web browser - 1

through a web browser - 7

### Music production (composition, recording):

There are 16 participants in this group. There is a significant number of participant in this group who are using computers that are not connected to the Internet. Large number of participants would prefer to search and retrieve audio through a Digital Work Station. Also they mostly work with audio loops.

Special requirements:

- Audio preview
- Scrobber
- Similarity recommendation
- categories (rich metadata)
- quality detection

---

***My computer for making sound and/or music is***

always connected to the Internet when I work on music - 10

**is not or rarely connected to the Internet - 6**

---

***For my work I mainly use (please tick all that apply)***

**audio loops (e.g., Drums/Bass/Keys) - 11**

audio loops (e.g., Drums/Bass/Keys);sound effects (e.g., traffic noise/barking dogs/keyboard layers) - 3

sound effects (e.g., traffic noise/barking dogs/keyboard layers) - 2

---

***I would prefer to search for and retrieve audio (please tick all that apply)***

No - 1

**from within my Digital Audio Workstation - 6**

from within my Digital Audio Workstation;through a standalone app - 2

from within my Digital Audio Workstation;through a standalone app;through a web browser - 2





## AudioCommons

### D2.1 Requirements report and use cases

from within my Digital Audio Workstation;through a web browser - 1  
through a standalone app - 1  
through a web browser - 3

### Music production (mixing, mastering):

There are 11 participants in this group. People that are part of this group are mostly using the computers that are connected to the Internet. They usually don't use sound libraries and they prefer to search and retrieve audio from within their Digital Audio Workstation. Also they mostly work with sound effects.

#### Special requirements:

- a tool that learns what user like or use most often
- auto suggest according to my past selections
- self organization (e.g. basses are displayed depending on how often I use them, similar sounds are recommended - youtube style)
- smart search, rapid preview, fast download, multiple sample rates
- easy preview that can be used within the DAW

---

***My computer for making sound and/or music is always connected to the Internet when I work on music - 9***  
is not or rarely connected to the Internet - 2

---

***I often work with a sound library of approximately (please tick all that apply)***  
10.000 Audio Samples/Loops/Sound FX - 1  
100 Audio Samples/Loops/Sound FX - 1  
1000 Audio Samples/Loops/Sound FX - 2  
1000 Audio Samples/Loops/Sound FX;I use the default sound library provided by my software - 1  
***I don't use any sound library - 5***  
more than 10.000 Audio Samples/Loops/Sound FX;I use the default sound library provided by my software - 1

---

***I would prefer to search for and retrieve audio (please tick all that apply) from within my Digital Audio Workstation - 5***  
from within my Digital Audio Workstation;through a standalone app;through a web browser - 1  
from within my Digital Audio Workstation;through a web browser - 1  
through a standalone app - 1  
through a standalone app;through a web browser - 1  
through a web browser - 2

### Whole songs:

There are 16 participants in this group (participants in this group use only whole songs in their work). People that are part of this group are mostly using the computers that are connected to the Internet. Large number of participants don't use sound libraries and would prefer to search and retrieve audio from within their Digital Audio Workstation.

#### Special requirements:

- Auto suggest according to history





## AudioCommons

### D2.1 Requirements report and use cases

- audio recognition similar to google images
- fast tagging feature

**My computer for making sound and/or music is always connected to the Internet when I work on music - 13**  
is not or rarely connected to the Internet - 3

---

**I often work with a sound library of approximately (please tick all that apply)**

10.000 Audio Samples/Loops/Sound FX - 2  
100 Audio Samples/Loops/Sound FX - 1  
100 Audio Samples/Loops/Sound FX;I don't use any sound library - 1  
1000 Audio Samples/Loops/Sound FX - 1  
1000 Audio Samples/Loops/Sound FX;I use the default sound library provided by my software - 1

**I don't use any sound library - 6**

I don't use any sound library;I use the default sound library provided by my software - 1  
I use the default sound library provided by my software - 1  
more than 10.000 Audio Samples/Loops/Sound FX - 1  
more than 10.000 Audio Samples/Loops/Sound FX;I don't use any sound library - 1

---

**I would prefer to search for and retrieve audio (please tick all that apply)**

**from within my Digital Audio Workstation - 9**

from within my Digital Audio Workstation;through a standalone app;through a web browser - 1  
from within my Digital Audio Workstation;through a web browser - 1  
through a standalone app - 2  
through a standalone app;through a web browser - 2  
through a web browser - 1

## Sound effects:

There are 110 participants in this group (participants in this group use only sound effects in their work). People that are part of this group are mostly using the computers that are connected to the Internet. Most of the participants belonging to this group work in audio for movies/TV domain. They prefer to search and retrieve audio from within their Digital Audio Workstation or web browser.

### Special requirements:

- audio preview
- autocomplete
- workflow integration into daw
- Tagging service
- quality detection
- metadata editor
- Similarity recommendation based on a specified attribute
- favourite system

**My computer for making sound and/or music is always connected to the Internet when I work on music - 93**  
is not or rarely connected to the Internet - 17

---

**What I mainly work on is (please tick all that apply)**

...

Audio for games - 8  
Audio for games;Audio for app development in general - 1  
Audio for games;Audio for movies and/or TV - 9





## AudioCommons

### D2.1 Requirements report and use cases

Audio for games;Audio for movies and/or TV;Audio for live performance - 1  
Audio for games;Voice work - 1  
Audio for games;YT videos - 1  
**Audio for movies and/or TV - 17**  
Audio for movies and/or TV;Podcast/Podfic (fan fiction audiobooks) - 1  
Audiobooks/audio plays - 1  
Music performance - 2  
Music production (composition, recording) - 2  
Music production (composition, recording);Audio for games - 1  
Music production (composition, recording);Audio for games;Audio for movies and/or TV - 1  
Music production (composition, recording);Audio for movies and/or TV - 3  
Music production (composition, recording);Music performance - 2  
Music production (mixing, mastering) - 4  
Music production (mixing, mastering);Audio for games - 1  
Music production (mixing, mastering);Audio for games;Audio for movies and/or TV - 1  
Music production (mixing, mastering);Audio for games;Audio for movies and/or TV;Music performance - 1  
Music production (mixing, mastering);Audio for games;Audio for movies and/or TV;Youtube - 1  
Music production (mixing, mastering);Audio for movies and/or TV - 1  
Music production (mixing, mastering);Music production (composition, recording) - 11  
...

---

I would prefer to search for and retrieve audio (please tick all that apply)

No - 1

No preferences: my DAW (Reaper) has this feature - 1

**from within my Digital Audio Workstation - 23**

from within my Digital Audio Workstation;cloud-based user-friendly unit - 1

from within my Digital Audio Workstation;through a standalone app - 11

from within my Digital Audio Workstation;through a standalone app;through a web browser - 2

from within my Digital Audio Workstation;through a web browser - 10

no preference - 1

sample things didn't made for it - 1

through a standalone app - 12

through a standalone app;through a web browser - 3

**through a web browser - 41**

through a web browser;I record some sound effects on my Dictaphone - 1

through a web browser;from microphone - 1

through an API - 1

### Audio loops:

There are 28 participants belonging to this group (participants in this group use only audio loops in their work). People that are part of this group are mostly working in the music production domain and they would prefer to search and retrieve audio from within their Digital Audio Workstation. Significant number of participants belonging to this group would like to search audio related data by using suggestions from the interface.

Special requirements:

- multi-search and preview retrieved sounds one after another
- Audio preview
- Easy querying
- Favorites section that still has keywords
- Similarity recommendation
- Tempo and key changes on the fly
- different clouds with continues filtering over the result set
- genre search/metadata





## AudioCommons

### D2.1 Requirements report and use cases

- integration across both local libraries, online libraries and products

#### *What I mainly work on is (please tick all that apply)*

##### **Music production (composition, recording) - 11**

Music production (composition, recording);Audio for games;Music performance - 1

##### **Music production (mixing, mastering);Music production (composition, recording) - 7**

Music production (mixing, mastering);Music production (composition, recording);Audio for games;Audio for movies - 1

Music production (mixing, mastering);Music production (composition, recording);Audio for movies and/or TV - 3

Music production (mixing, mastering);Music production (composition, recording);Music performance - 4

---

#### *I would prefer to search for and retrieve audio (please tick all that apply)*

##### **from within my Digital Audio Workstation - 18**

from within my Digital Audio Workstation;through a standalone app - 3

from within my Digital Audio Workstation;through a standalone app;through a web browser - 3

through a standalone app - 1

through a web browser - 3

---

#### *When I search for music-related data I want to (please tick all that apply)*

query using suggested keywords from the interface (Drop Down Lists or similar) - 2

##### **query using suggested keywords from the interface (Drop Down Lists or similar);use keywords (Google style) - 7**

query using suggested keywords from the interface (Drop Down Lists or similar);use keywords (Google style);use natural language (for instance, whole sentences) - 2

##### **use keywords (Google style) - 9**

use natural language (for instance, whole sentences) - 2

## 2.4 Search and metadata

In this section we will analyse the answers on questions regarding the search strategies and metadata associated with the audio files.

Insight into how users would like to search for specific files should mostly impact the design of the user interface. Metadata requirements should be used in later deliverables concerning the ontology building.

### 2.4.1 Search strategies

- **Interface:** users were asked about their preferred way of accessing the query endpoint. Four predefined solutions were offered: accessing the search results from within the Digital Audio Workstation (DAW), from a standalone application, web browser or other. Results on Fig. 14 show that most users prefer to have a query endpoint inside their Digital Audio Workstation (63.7%) which hence suggests to develop a DAW plugin that is capable of accepting the queries and presenting the results inside DAW interface.





### I would prefer to search for and retrieve audio (please tick all that apply) (193 responses)

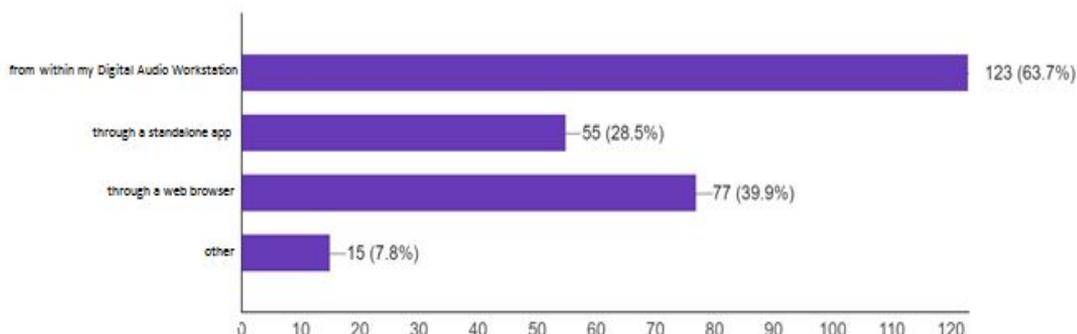


Fig. 14: Preferred way for accessing the query endpoint

- **Audio files labelling:** search engines can index textual fields describing the audio files and use those labels to match with keywords from user query (even if the name of the audio file is a generic one). Results can be grouped in many different categories. In case that additional metadata is not available, search engines rely only on the label of the file. Intelligent labels (semantically and syntactically well formed labels) on the retrieved files are important but abundance of metadata is even more (Fig. 15).

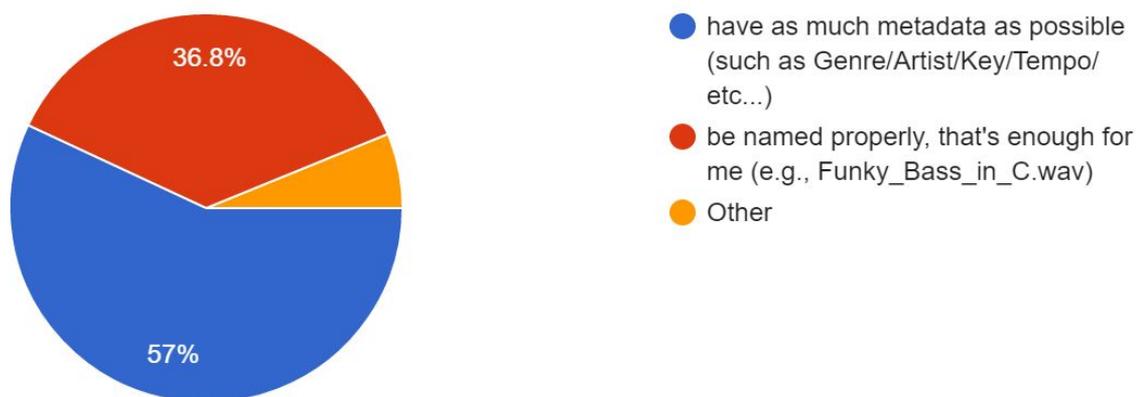


Fig. 15: Preferred style of labelling

- **Recall:** When presenting the search results to the user, system needs to be designed to conform to the limitations of the interface and users' wishes/needs and capabilities to digest certain





amounts of data. Fig. 16 suggests that most users want to search from an unlimited pool of data, but only if that data can be queried on some smart way. Also large percentage of users prefer to have curated content which is generally smaller in quantity but much higher in quality (something similar to a catalogue).

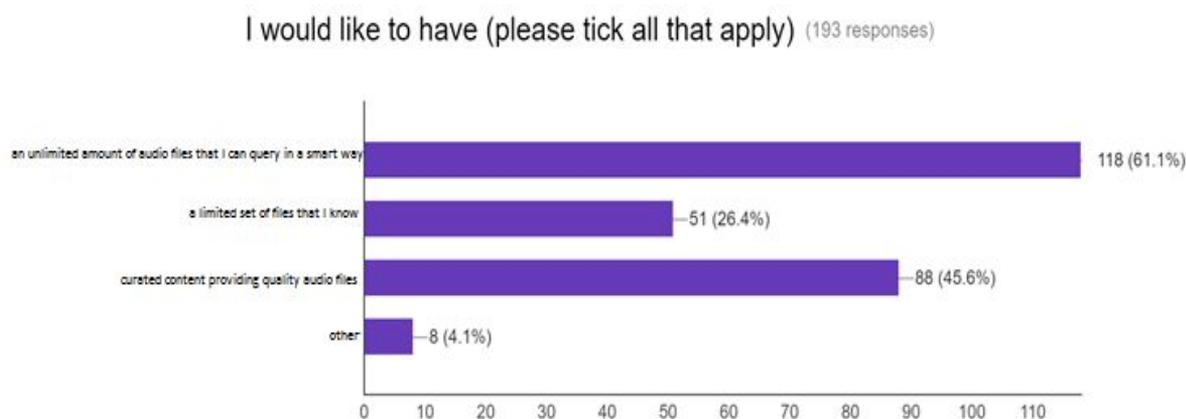


Fig. 16: Presentation of retrieved files

- **Query language:** search can be conducted using various strategies and languages:
  - One Approach is using a structured language that is similar to programming (SPARQL, SQL, etc.). While this approach allows building very powerful queries and works very well with linked data and RDF, mastering those languages can be hard for many non or less technical users (14% of responses favoring this approach).
  - An approach that is using some interface elements (like drop down lists) to help users build an expansive and structured query without using structured language (49.2% of responses favoring this approach).
  - Searching by using a simple list of relevant terms (keywords) proves to be the most popular approach with 81.3% responses favoring this approach (most people use it daily on popular search engines like Google).
  - Using natural language is one option that is appealing for many users because it represents the exact mapping of our thought, but this approach can be hard to





## AudioCommons

### D2.1 Requirements report and use cases

implement because of the complexity and ambiguity of human language (21.8% of responses favoring this approach).

- o Using the non-verbal graphical interface (13.5% of responses favoring this approach).

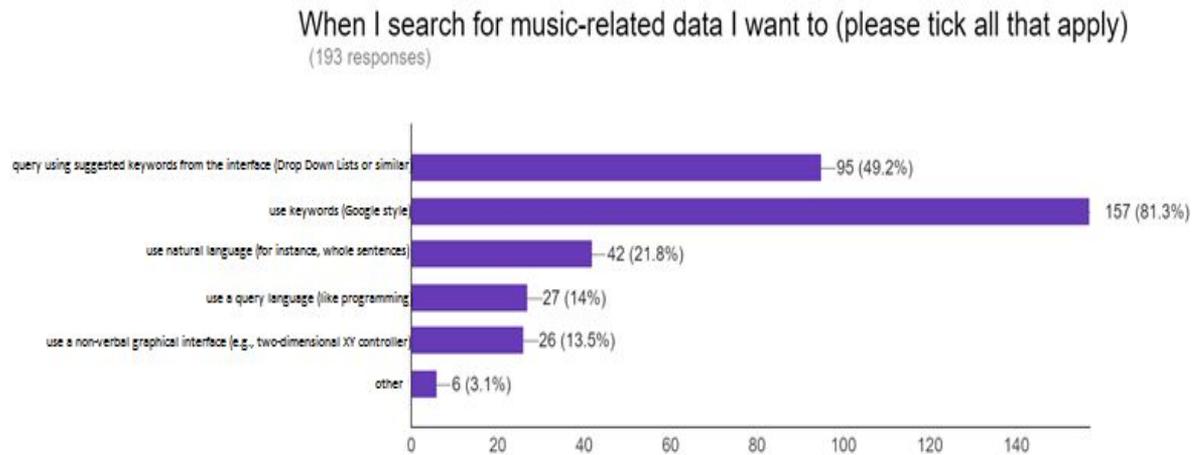


Fig. 17: Type of query language to search for files

#### Scenario:

- Plugin for searching the Web and Audio Commons Ecosystem should be incorporated into a DAW.
- The Search interface could be comprised of the following elements:
  - Keyword field - accepts the user query a list of keywords
  - Dropdown lists - creation of a complex query (list are thematically organised)
  - Word cloud - searching by curated or suggested tags



## DAW Interface

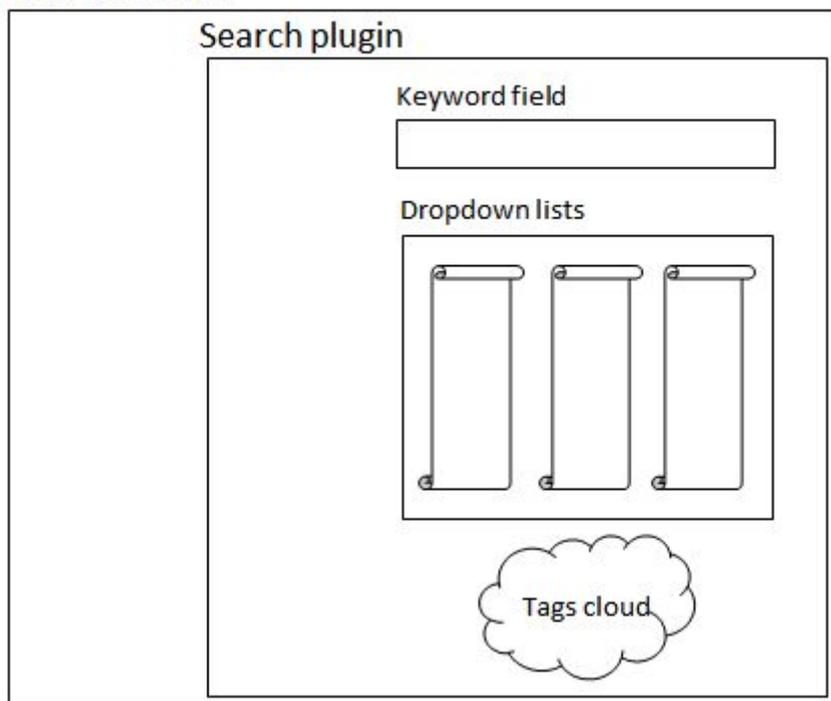


Fig. 18: Proposed elements of a search interface



## 2.4.2 Building Audio Commons Vocabulary

In our survey we asked people from our target domains to suggest the terms (metadata) that they would like to use for searching or filtering the audio content.

Questions asked about:

- Musical Properties
- Perceptual Properties
- General Metadata

Collected terms can be seen in Tables 1, 2 and 3 in the appendix 1. For each term we tried to extract its accompanying description (gloss) from Wordnet [Fellbaum] morphological dictionary (in case where description exist). Data set is also augmented with the URIs retrieved from DBpedia [DBpedia].

For data augmentation task we used Python Wordnet library (nlk.corpus) and DBpedia spotlight API [Spotlight].

After completion of this task the following categories were distinguished:

- Software - Effects, VST, audio format
- Hardware - audio channels, equipment
- Measures - bit-depth, bit rate, bpm, loudness, sample rate
- Genres- classical music, dubstep, new age, vocal music
- Instruments - drums, electric piano, guitar, hi-hat, percussion, piano, violin
- Musical attribute - key, monophony, polyphony, rhythm
- Techniques - slide guitar
- Effects - distortion
- Sound character - urban, rural, cinematic, retro, etc.
- Music representation - audio histogram
- Moods - dark, faded, bright, meditative, happy, etc.
- Instrumentation - orchestra
- Legal document - license

Each of these categories will require the creation of it own ontology/light taxonomy. A certain amount of entities is already defined in existing ontologies and should be reused (software and hardware entities describing music production is already defined in Studio ontology, measures and sound properties in music and audio features ontology, instrument taxonomy in musical instruments taxonomies).

The Audio Commons ontology should contain knowledge defined by service providers' APIs. Audio Commons partners Freesound and Jamendo provide their users with an API with large amount of parameters that are describing the audio content. API parameter descriptions provide a valuable knowledge and should be used in the task of building the Audio Commons ontology:

- Freesound: <https://www.freesound.org/docs/api/>
- Jamendo: <https://developer.jamendo.com/v3.0/docs>

Fig. 19. shows Freesound metadata related with specific sound.





Name	Type
<u>id</u>	number
<u>url</u>	URI
<u>name</u>	string
<u>tags</u>	array(strings)
<u>description</u>	string
<u>geotag</u>	string
<u>created</u>	string
<u>license</u>	string
<u>w3s</u>	string
<u>channels</u>	number
<u>filesize</u>	number
<u>bitrate</u>	number
<u>bitdepth</u>	number
<u>duration</u>	number
<u>samplerate</u>	number
<u>username</u>	string
<u>pack</u>	URI
<u>download</u>	URI
<u>bookmark</u>	URI
<u>previews</u>	object
<u>images</u>	object
<u>num_downloads</u>	number
<u>avg_rating</u>	number
<u>num_ratings</u>	number
<u>rate</u>	URI
<u>comments</u>	URI
<u>num_comments</u>	number
<u>comment</u>	URI
<u>similar_sounds</u>	URI
<u>analysis</u>	object
<u>analysis_stats</u>	URI
<u>analysis_frames</u>	URI

Fig 19: Freesound Metadata (sound entity)





## 2.5 Current problems and User Wishlist

### 2.5.1 Current problems with audio content browsing and retrieving

Question number 17 from the survey asked from participants to write about their biggest frustrations they face now when searching for audio. We analysed each response and identified each theme/problem that appeared in the answers. Those themes/problems can be grouped in the following categories:

- **Licensing** - licensing information is not always clearly presented and it is often hard to understand
- **Syntax** - problematic labeling of retrieved audio content. File names have generic names or spelling mistakes or simply the label does not represent the content correctly. The same problem is applied to tags associated with the audio content
- **Sparse metadata** - lack of rich metadata that describes the audio content can make that content less visible
- **Workflow integration** - (semi)automated way of retrieving relevant data into a specific part of workflow is lacking
- **Bad source** - retrieved data contains audio of bad recording quality
- **Interface problems** - interface for searching and retrieving audio content often suffers from bad design (lots of redirections, unwanted pop-ups, etc.)
- **Lack of recommendation/curation** - some users would prefer to get some interesting cues that could help them to start





## 2.5.2 User Wishlist

Naturally, users expressed their wishes concerning the task of searching and retrieving audio content that are connected with the list of problems they are facing today while conducting the task. By analysing answers from the question 18 from the survey we grouped the answers into following themes:

- **Licensing** - all retrieved files should contain licensing information
- **Interface** - ability to use drop down functionality to move retrieved files directly into their workflows (inside the application)
- **Recommendation** - show similar content
- **Metadata** - more metadata concerning the genres and other attributes (for full list refer to Appendix 2)
- **Services** - retrieved audio content can be enriched with metadata produced by the services that should be available on demand (retrieve the tempo and pitch information through analysis, provide time-stretched versions of the original file on the fly, fast tagging, etc. )





## 3 Ontology and Metadata

### 3.1 Metadata

Metadata is the data that describes other data or information about information (labeling, cataloging and descriptive information structured in such a way that allows Web pages to be properly searched and processed in particular *by computer*)[Metadata].

According to W3C there are many practical usages of metadata (expressed as RDF - Resource Description Framework [RDF]). These are the usages that are relevant to Audio Commons Project:

- **Thesauri and library classification schemes** - These are well-known examples of hierarchical systems for representing subject taxonomies in terms of the relationships between named concepts.
- **Description of the contents of Web pages.** This is one of the basic functions of the Dublin Core initiative. The Dublin Core is a set of 15 properties associated with bibliographic information. These can be used to describe items on the Web sufficiently well that search engines and other software can work much more efficiently.
- **Descriptions of device capabilities** - RDF provides a way to describe the capabilities and preferences associated with users and the hardware and software they are using to access the Web. This will permit Web content to be tailored to the specific needs of the user.
- **Rating systems** - offering a way of labeling resources so that people (or computers) can filter information.

When audio went from analogue to digital, it became possible to label audio files with more information than could be contained in just the filename. That descriptive information is called audio tag or audio metadata in general. Computer programs specialized in adding or modifying this information are called tag editors.

Metadata can be used to name, describe, catalogue and indicate ownership or copyright for a digital audio file, and its presence makes it much easier to locate a specific audio file within a group – through use of a search engine that accesses the metadata. As different digital audio formats were developed, it





was agreed that a standardized and specific location would be set aside within the digital files where this information could be stored.

## 3.2 Ontology

An ontology defines a set of representational primitives with which we can model a domain of knowledge or discourse. Those representational primitives are typically classes (or sets), attributes (or properties), and relationships (or relations among class members). The definitions of the representational primitives include information about their meaning and constraints on their logically consistent application.

In the Audio Commons project the ontology should play the central role in the ecosystem. The Audio Commons ontology should contain enough knowledge to be capable of dealing with all the requirements of the users collected through the survey and any other means.

Ontologies should not be built entirely from scratch because one of the main imperative of Semantic Web and Linked Data method is reuse. Outside the knowledge collected from the survey, large amount of knowledge should come from already established ontologies that are relevant for Audio Commons project such as:

- **Music ontology** - provides a vocabulary for publishing and linking a wide range of music-related data on the Web [Raimond].
- **Musical Instrument Taxonomies** - describing the concepts related to musical instruments, taking into account two different classification systems: i) one proposed by Horbostel & Sachs which is denoted taxonomy 'A', and ii) one proposed by Jeremy Montagu & John Burton [Kolozali]
- **Media Value Chain Ontology** - represents the Intellectual Property (IP) along the Value Chain. There are different kinds of objects of the Intellectual Property (we call them IP Entities) and different actions that are performed on them, what defines the different roles that users can play regarding these IP Entities [MediaValueChain].
- **Audio features ontology** - This ontology expresses some common concepts to represent some features of audio signals. It mainly relies on the Event ontology, in order to classify particular parts of the timeline backing an audio signal. It also supports dense features, such as chromagrams, onset detection function, etc. [AudioFeatures]
- **Studio Ontology** - can be used for describing and sharing detailed information about music production in the recording studio. It facilitates capturing the nuances of record production, by





Example on Fig. 21 shows the possible model for an ontology that is aware of how certain actions can have effect on licensing permissions. For example user can conduct different tasks with the audio content like creating the file, making the copy of the file, adapting the file or just distributing it.

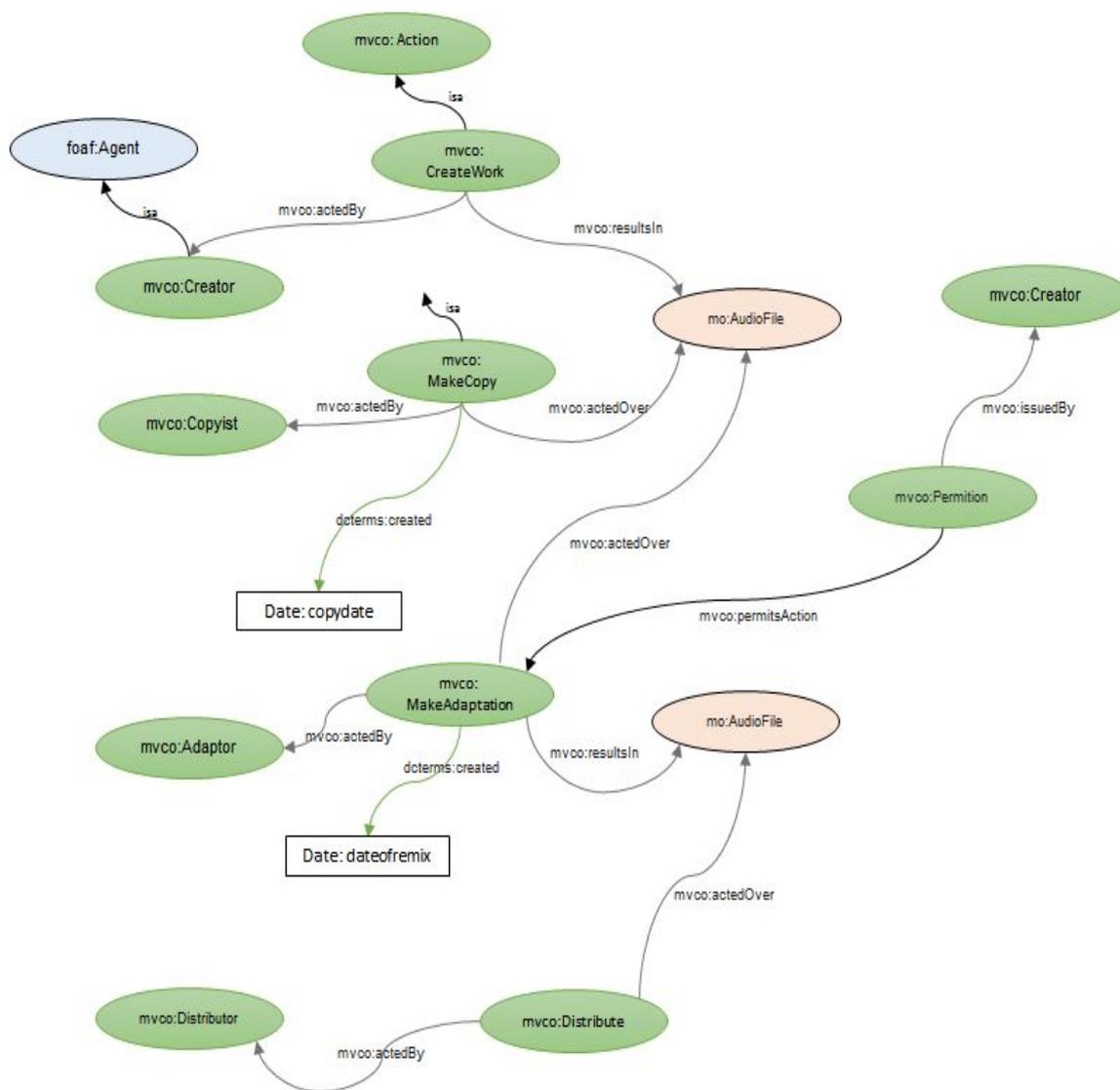


Fig. 21: Actions and permissions

It is necessary to build a vocabulary of actions that can have an impact on licensing permissions. For example Fig 22. is showing the textual description of different licenses defined by Creative Commons [CreativeCommons]. Those descriptions need to be assessed and modeled inside the ontology. Fig. 23 shows how semantic role labeling (a task in natural language processing consisting of the detection of the semantic arguments associated with the predicate or verb of a sentence and their classification into their specific roles) can be used to distinguish actions and entities that are implied in text through verbs.



[CC0 Public Domain Dedication](#) - By using CC0, you waive all copyright and related rights to a work to the extent possible under the law.

[Attribution](#) - This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.

[Attribution-ShareAlike](#) - This license lets others remix, tweak, and build upon your work even for commercial purposes, as long as they credit you and license their new creations under the identical terms. All new works based on yours will carry the same license, so any derivatives will also allow commercial use.

[Attribution-NoDerivs](#) - This license allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to you.

[Attribution-NonCommercial](#) - This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don't have to license their derivative works on the same terms.

[Attribution-NonCommercial-ShareAlike](#) - This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under identical terms.

[Attribution-NonCommercial-NoDerivs](#) - This license is the most restrictive of our six licenses, only allowing others to download your works and share them with others as long as they credit you, but they can't change them in any way or use them commercially.

Fig. 22: Creative Commons licences

**Arg0-PAG: *causer of transformation*** (vrole: 26.6.1-1-Agent, 45.4-Agent)  
**Arg1-PPT: *thing changing*** (vrole: 26.6.1-1-Patient, 45.4-Patient)  
**Arg2-PRD: *end state*** (vrole: 26.6.1-1-Result)  
**Arg3-VSP: *start state*** (vrole: 26.6.1-1-Material)

**Arg0-PAG: *distributor*** (vrole: 13.2-1-Agent)  
**Arg1-PPT: *thing distributed*** (vrole: 13.2-1-Theme)  
**Arg2-GOL: *distributed to*** (vrole: 13.2-1-Recipient)

**Arg0-PAG: *copyist, agent*** (vrole: 25.2-agent, 25.4-agent)  
**Arg1-PPT: *original*** (vrole: 25.2-theme, 25.4-theme)  
**Arg2-GOL: *copy*** (vrole: 25.2-destination, 25.4-destination)

**Arg0-PAG: *agent, uploader***  
**Arg1-PPT: *data, thing uploaded***  
**Arg2-GOL: *patient, uploaded to what***

Fig 23: Semantic roles automatically fetched from lexical resource PropBank [PropBank]





## 4 Conclusion

In this deliverable we presented the user requirements for the Audio Commons Ecosystem collected through the online survey. We also presented some initial work on the future Audio Commons ontology and existing ontologies relevant to the Audio Commons project. The survey contained 24 questions asking creatives working in music industry about various subjects like demographics, workflows they use and metadata they would like to use when searching for a new content on the Web. By analysing the demographic of the participants of the survey we discovered that most of them are professionals with more than 10 years of experience in the music domain, especially in the field of music production. Questions regarding workflows asked users about the tools they use, about the amount of data they use in their daily tasks, source and type of the data and the time they invest in working with that data. Final set of questions in the survey were about the search strategies and metadata associated with the audio files. Answers on those questions allowed us to gain insight into how users would like to search for specific files and how that strategy would impact the design of the user interface. Also, data collected on the metadata requirements (Musical Properties, Perceptual Properties, General Metadata) should be used in later deliverables concerning the ontology building (WP2.2.2).

For this deliverable we analysed which existing ontologies and vocabularies that are describing the music domain already exist. Conclusion concerning existing data in the audio content domain is to follow the guidelines of Semantic Web and reuse as much of existing data as possible (ontologies and vocabularies like **Music ontology**, Musical Instrument Taxonomies, Media Value Chain Ontology, Audio features ontology, Studio Ontology)

This deliverable is part of the WP2 work package. Other deliverables that will be created under this work package will be dealing with description of the Audio Commons Ontology and specification of the Audio Commons API (specification on how the different components of the Audio Commons Ecosystem will be technically interconnected (i.e., how production tools will be able to access Audio Commons content and users will be able to communicate for the licensing process)). This deliverable covered some aspects that will be dealt with more details in the future deliverables. In this deliverable we set some ideas and guidelines for building the Audio Commons ontology and Audio Commons API.





## 5 References

[Gruber] *Encyclopedia of Database Systems*, Ling Liu and M. Tamer Özsu (Eds.), Springer-Verlag, 2009

[Metadata] <https://www.w3.org/Metadata/Activity.html>

[RDF] <https://www.w3.org/RDF/>

[Raimond] Y. Raimond, S. Abdallah, M. Sandler, F. Giasson. The Music Ontology, Proceedings of the International Conference on Music Information Retrieval (ISMIR), 2007

[Kolozali] Sefki Kolozali, George Fazekas, Mathieu Barthet, Mark Sandler. Knowledge Representation Issues in Musical Instrument Ontology Design. 12th International Society for Music Information Retrieval Conference (ISMIR 2011), 2011.

[AudioFeatures] [http://motools.sourceforge.net/doc/audio\\_features.html](http://motools.sourceforge.net/doc/audio_features.html)

[MediaValueChain] <http://dmag.ac.upc.edu/ontologies/mvco/>

[Fazekas] G. Fazekas, M. Sandler. The Studio Ontology Framework. The 12th International Society for Music Information Retrieval, (ISMIR-11) Miami, Florida (USA) October 24–28, 2011.

[CreativeCommons] <https://creativecommons.org/>

[PropBank] M. Palmer, D. Gildea, P. Kingsbury. The Proposition Bank: A Corpus Annotated with Semantic Roles Computational Linguistics Journal, 31:1, 2005.

[Fellbaum] C. Fellbaum (1998, ed.) WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press.

[Spotlight] J. Daiber, M. Jakob, C. Hokamp, P. N. Mendes. Improving Efficiency and Accuracy in Multilingual Entity Extraction. Proceedings of the 9th International Conference on Semantic Systems (I-Semantics). 2013.





# APPENDIX 1: User requirements for Metadata

Table 1: Terms of musical properties (collected from the survey)

Term	Term description	Term hypernym	URI
Bass register	the lowest part of the musical range	[Synset('pitch.n.01')] / sound property	
Effects	processors for audio files	software	
Rhythm	the basic rhythmic unit in a piece of music	[Synset('musical_time.n.01')] / sound property	<a href="http://dbpedia.org/resource/Rhythm">http://dbpedia.org/resource/Rhythm</a>
VST	is a software interface that integrates software audio synthesizer and effect plugins with audio editors and recording systems	software	<a href="http://dbpedia.org/resource/Virtual_Studio_Technology">http://dbpedia.org/resource/Virtual_Studio_Technology</a>
audio channels	is an audio signal communications channel in a storage device	hardware	<a href="http://dbpedia.org/resource/Audio_channel">http://dbpedia.org/resource/Audio_channel</a>
audio format	is a file format for storing digital audio data on a computer system	software	<a href="http://dbpedia.org/resource/Audio_file_format">http://dbpedia.org/resource/Audio_file_format</a>
bit-depth	is the number of bits of information in each sample, and it directly corresponds to the resolution of each sample	measure	
clap	a sudden very loud noise	[Synset('noise.n.01')]	





classical music	is art music produced or rooted in the traditions of Western music, including both liturgical (religious) and secular music.	genre	<a href="http://dbpedia.org/resource/Classical_music">http://dbpedia.org/resource/Classical_music</a>
drums	is a collection of drums and other percussion instruments	instrument	<a href="http://dbpedia.org/resource/Drum_kit">http://dbpedia.org/resource/Drum_kit</a>
dubstep	is a genre of electronic dance music	genre	<a href="http://dbpedia.org/resource/Dubstep">http://dbpedia.org/resource/Dubstep</a>
electric piano	is a modern electronic musical instrument; a variation of electronic keyboard or synthesizer	instrument	<a href="http://dbpedia.org/resource/Electric_piano">http://dbpedia.org/resource/Electric_piano</a>
fx	the processing of sound using digital software	software	
genre	a kind of literary or artistic work	[Synset('kind.n.01')]	<a href="http://dbpedia.org/resource/Genre">http://dbpedia.org/resource/Genre</a>
guitar	a stringed instrument usually having six strings; played by strumming or plucking	[Synset('stringed_instrument.n.01')]	<a href="http://dbpedia.org/resource/Guitar">http://dbpedia.org/resource/Guitar</a>
hi-hat	a type of cymbal and stand used as a typical part of a drum kit by percussionists	instrument	<a href="http://dbpedia.org/resource/Hi-hat">http://dbpedia.org/resource/Hi-hat</a>
instrument	any of various devices or contrivances that can be used to produce musical tones or sounds	[Synset('device.n.01')]	<a href="http://dbpedia.org/resource/Musical_instrument">http://dbpedia.org/resource/Musical_instrument</a>





key	is a group of pitches, or <b>scale</b> upon which a music composition is created	music theory	
monophony	music consisting of a single vocal part (usually with accompaniment)	musical texture	<a href="http://dbpedia.org/resource/Monophony">http://dbpedia.org/resource/Monophony</a>
mood	a characteristic (habitual or relatively temporary) state of feeling	[Synset('feeling.n.01')]	<a href="http://dbpedia.org/resource/Mood_(psychology)">http://dbpedia.org/resource/Mood_(psychology)</a>
new age	is a genre of music intended to create artistic inspiration, relaxation, and optimism	genre	
orchestra	a musical organization consisting of a group of instrumentalists including string players	[Synset('musical_organization.n.01')]	<a href="http://dbpedia.org/resource/Orchestra">http://dbpedia.org/resource/Orchestra</a>
percussion	is a musical instrument that is sounded by being struck or scraped by a beater struck, scraped or rubbed by hand; or struck against	instrument	<a href="http://dbpedia.org/resource/Percussion_instrument">http://dbpedia.org/resource/Percussion_instrument</a>
piano	a keyboard instrument that is played by depressing keys that cause hammers to strike tuned strings and produce sounds	[Synset('keyboard_instrument.n.01'), Synset('percussion_instrument.n.01'), Synset('stringed_instrument.n.01')]	<a href="http://dbpedia.org/resource/Piano">http://dbpedia.org/resource/Piano</a>





plug-in	is a software component that adds a specific feature to an existing computer program	software	<a href="http://dbpedia.org/resource/Plug-in_(computing)">http://dbpedia.org/resource/Plug-in_(computing)</a>
polyphony	music arranged in parts for several voices or instruments	musical texture	<a href="http://dbpedia.org/resource/Polyphony">http://dbpedia.org/resource/Polyphony</a>
scale	is any set of musical notes ordered by fundamental frequency or pitch	[Synset('standard.n.01')] / representation	<a href="http://dbpedia.org/resource/Scale_(music)">http://dbpedia.org/resource/Scale_(music)</a>
slide guitar	is a particular method or technique for playing the guitar	technique	<a href="http://dbpedia.org/resource/Slide_guitar">http://dbpedia.org/resource/Slide_guitar</a>
snare	something (often something deceptively attractive) that catches you unawares	[Synset('design.n.02')]	<a href="http://dbpedia.org/resource/Snare_drum">http://dbpedia.org/resource/Snare_drum</a>
sound effects	the processing of sound using digital software	software	<a href="http://dbpedia.org/resource/Sound_effect">http://dbpedia.org/resource/Sound_effect</a>
sound quality	is typically an assessment of the accuracy, enjoyability, or intelligibility of audio output from an device	measure	<a href="http://dbpedia.org/resource/Sound">http://dbpedia.org/resource/Sound</a>
sounds	is a vibration that propagates as a typically audible mechanical wave of pressure and displacement, through a medium such as air or water	property	<a href="http://dbpedia.org/resource/Sound">http://dbpedia.org/resource/Sound</a>





tempo	(music) the speed at which a composition is to be played	[Synset('musical_time.n.01')]	<a href="http://dbpedia.org/resource/Tempo">http://dbpedia.org/resource/Tempo</a>
timbre	(music) the distinctive property of a complex sound (a voice or noise or musical sound)	[Synset('sound_property.n.01')]	<a href="http://dbpedia.org/resource/Timbre">http://dbpedia.org/resource/Timbre</a>
violin	bowed stringed instrument that is the highest member of the violin family; this instrument has four strings and a hollow body and an unfretted fingerboard and is played with a bow	[Synset('bowed_stringed_instrument.n.01')]	<a href="http://dbpedia.org/resource/Violin">http://dbpedia.org/resource/Violin</a>
vocal music	is a type of music performed by one or more singers, with or without instrumental accompaniment (a cappella), in which singing provides the main focus of the piece	genre	<a href="http://dbpedia.org/resource/Vocal_music">http://dbpedia.org/resource/Vocal_music</a>

Table 2: Perceptual properties (Wordnet)

Term	Term description	URI
Dark	devoid of or deficient in light or brightness; shadowed or black	<a href="http://dbpedia.org/resource/Darkness">http://dbpedia.org/resource/Darkness</a>
Faded	become less clearly visible or distinguishable; disappear gradually or seemingly;	
Noise	sound of any kind (especially unintelligible or dissonant sound)	<a href="http://dbpedia.org/resource/Noise">http://dbpedia.org/resource/Noise</a>





Powerful	having great power or force or potency or effect	<a href="http://dbpedia.org/resource/Power_(social_and_political)">http://dbpedia.org/resource/Power_(social_and_political)</a>
acid	harsh or corrosive in tone	<a href="http://dbpedia.org/resource/Acid">http://dbpedia.org/resource/Acid</a>
ambient	completely enveloping	<a href="http://dbpedia.org/resource/Ambient_music">http://dbpedia.org/resource/Ambient_music</a>
atonal	characterized by avoidance of traditional western tonality	<a href="http://dbpedia.org/resource/Atonality">http://dbpedia.org/resource/Atonality</a>
bright	emitting or reflecting light readily or in large amounts	<a href="http://dbpedia.org/resource/Apparent_magnitude">http://dbpedia.org/resource/Apparent_magnitude</a>
chilled	depress or discourage	<a href="http://dbpedia.org/resource/Chill-out_music">http://dbpedia.org/resource/Chill-out_music</a>
cinematic	of or pertaining to or characteristic of the cinema	<a href="http://dbpedia.org/resource/Film">http://dbpedia.org/resource/Film</a>
clear	readily apparent to the mind	
crunchy	the sound of something crunching	
cutting	(of speech) harsh or hurtful in tone or character	<a href="http://dbpedia.org/resource/Cut_(earthmoving)">http://dbpedia.org/resource/Cut_(earthmoving)</a>
dark	devoid of or deficient in light or brightness; shadowed or black	
deep	relatively deep or strong; affecting one deeply	
dry	free from liquid or moisture; lacking natural or normal moisture or depleted of water; or no longer wet	
dull	lacking in liveliness or animation	<a href="http://dbpedia.org/resource/Boredom">http://dbpedia.org/resource/Boredom</a>
emotional	determined or actuated by emotion rather than reason	<a href="http://dbpedia.org/resource/Emotion">http://dbpedia.org/resource/Emotion</a>
ethnic	denoting or deriving from or distinctive of the ways of living built up by a group of people	<a href="http://dbpedia.org/resource/Ethnic_group">http://dbpedia.org/resource/Ethnic_group</a>
experimental	relating to or based on experiment	<a href="http://dbpedia.org/resource/Experimental_music">http://dbpedia.org/resource/Experimental_music</a>





funky	having the soulful feeling of early blues	<a href="http://dbpedia.org/resource/Funk">http://dbpedia.org/resource/Funk</a>
happy	enjoying or showing or marked by joy or pleasure	<a href="http://dbpedia.org/resource/Happiness">http://dbpedia.org/resource/Happiness</a>
high	greater than normal in degree or intensity or amount	
light	of comparatively little physical weight or density	<a href="http://dbpedia.org/resource/Light">http://dbpedia.org/resource/Light</a>
long	primarily temporal sense; being or indicating a relatively great or greater than average duration or passage of time or a duration as specified	
meditative	deeply or seriously thoughtful;	<a href="http://dbpedia.org/resource/Meditation">http://dbpedia.org/resource/Meditation</a>
melodic	containing or constituting or characterized by pleasing melody	<a href="http://dbpedia.org/resource/Melody">http://dbpedia.org/resource/Melody</a>
minimal	the least possible	<a href="http://dbpedia.org/resource/Minimalism">http://dbpedia.org/resource/Minimalism</a>
modern	characteristic of present-day art and music and literature and architecture	
moody	showing a brooding ill humor	
motivational	of or relating to motivation	<a href="http://dbpedia.org/resource/Motivation">http://dbpedia.org/resource/Motivation</a>
natural	in accordance with nature; relating to or concerning nature	<a href="http://dbpedia.org/resource/Mood_(psychology)">http://dbpedia.org/resource/Mood_(psychology)</a>
noisy	full of or characterized by loud and nonmusical sounds	
powerful	having great power or force or potency or effect	<a href="http://dbpedia.org/resource/Power_(social_and_political)">http://dbpedia.org/resource/Power_(social_and_political)</a>
psychedelic	producing distorted sensory perceptions and feelings or altered states of awareness or	<a href="http://dbpedia.org/resource/Psychedelia">http://dbpedia.org/resource/Psychedelia</a>





	sometimes states resembling psychosis	
rasping	unpleasantly harsh or grating in sound	<a href="http://dbpedia.org/resource/Rasp">http://dbpedia.org/resource/Rasp</a>
relaxed	without strain or anxiety	
retro	affecting things past	<a href="http://dbpedia.org/resource/Retro_style">http://dbpedia.org/resource/Retro_style</a>
sharp	(of something seen or heard) clearly defined	<a href="http://dbpedia.org/resource/Sharp_(music)">http://dbpedia.org/resource/Sharp_(music)</a>
slick	made slick by e.g. ice or grease	
slow	not moving quickly; taking a comparatively long time	
smooth	having a surface free from roughness or bumps or ridges or irregularities	
smooth	having a surface free from roughness or bumps or ridges or irregularities	
soft	yielding readily to pressure or weight	
strong	having strength or power greater than average or expected	
upbeat	pleasantly (even unrealistically) optimistic	
warm	having or producing a comfortable and agreeable degree of heat or imparting or maintaining heat	

Table 3: Metadata for search (Wordnet)

Term	Term description	Term hypernyms	URI
Genre	a kind of literary or artistic work	[Synset('kind.n.01')] / style	<a href="http://dbpedia.org/resource/Genre">http://dbpedia.org/resource/Genre</a>





SFX	the processing of sound using digital software	software	
Style	how something is done or how it happens	[Synset('property.n.02')]	
arrangement	a musical reconceptualization of a previously composed work	[Synset('planning.n.03')]	<a href="http://dbpedia.org/resource/Arrangement">http://dbpedia.org/resource/Arrangement</a>
atonality	the absence of a key; alternative to the diatonic system	[Synset('musical_notation.n.01')]	<a href="http://dbpedia.org/resource/Atonality">http://dbpedia.org/resource/Atonality</a>
audio format	is a file format for storing digital audio data on a computer system	software	<a href="http://dbpedia.org/resource/Audio_file_format">http://dbpedia.org/resource/Audio_file_format</a>
audio histogram	is a graphical representation of the distribution of numerical data	representation	
bit rate	is the number of bits that are conveyed or processed per unit of time	measure	<a href="http://dbpedia.org/resource/Bit_rate">http://dbpedia.org/resource/Bit_rate</a>
bit-depth	is the number of bits of information in each sample, and it directly corresponds to the resolution of each sample	measure	
bpm	the pace of music measured by the number of beats occurring in 60 seconds	[Synset('tempo.n.02')]	<a href="http://dbpedia.org/resource/Tempo">http://dbpedia.org/resource/Tempo</a>





character	a characteristic property that defines the apparent individual nature of something	property	
description	a statement that represents something in words	statement	Nothing
distortion	is the alteration of the original shape (or other characteristic) of something, such as an object, image, sound or waveform	technique	<a href="http://dbpedia.org/resource/Distortion">http://dbpedia.org/resource/Distortion</a>
electronic instrument	a musical instrument that generates sounds electronically	instrument	<a href="http://dbpedia.org/resource/Electronic_musical_instrument">http://dbpedia.org/resource/Electronic_musical_instrument</a>
equipment	an instrumentality needed for an undertaking or to perform a service	hardware	
indoors	within a building	character	
instrument	a device that requires skill for proper use	[Synset('device.n.01')]	<a href="http://dbpedia.org/resource/Musical_instrument">http://dbpedia.org/resource/Musical_instrument</a>
jingle	a metallic sound	[Synset('sound.n.04')]	<a href="http://dbpedia.org/resource/Jingle">http://dbpedia.org/resource/Jingle</a>
license	a legal document giving official permission to do something	[Synset('legal_document.n.01')]	<a href="http://dbpedia.org/resource/License">http://dbpedia.org/resource/License</a>
loop	a loop is a repeating section of sound material	technique	<a href="http://dbpedia.org/resource/Control_flow">http://dbpedia.org/resource/Control_flow</a>





loudness	the magnitude of sound (usually in a specified direction)	[Synset('sound_property.n.01')]	<a href="http://dbpedia.org/resource/Loudness">http://dbpedia.org/resource/Loudness</a>
lyrics	the text of a popular song or musical-comedy number	written communication	
monophonic	music consisting of a single vocal part (usually with accompaniment)	musical texture	<a href="http://dbpedia.org/resource/Monophony">http://dbpedia.org/resource/Monophony</a>
mood	a characteristic (habitual or relatively temporary) state of feeling	[Synset('feeling.n.01')]	<a href="http://dbpedia.org/resource/Mood_(psychology)">http://dbpedia.org/resource/Mood_(psychology)</a>
polyphonic	music arranged in parts for several voices or instruments	musical texture	<a href="http://dbpedia.org/resource/Polyphony">http://dbpedia.org/resource/Polyphony</a>
quality	an essential and distinguishing attribute of something or someone	attribute	
rural	relating to rural area	character	<a href="http://dbpedia.org/resource/Rural_area">http://dbpedia.org/resource/Rural_area</a>
sample rate	is the number of samples of audio carried per second	measure	<a href="http://dbpedia.org/resource/Sampling_rate">http://dbpedia.org/resource/Sampling_rate</a>
saturation	the process of totally saturating something with a substance	[Synset('permeation.n.01')]	<a href="http://dbpedia.org/resource/Colorfulness">http://dbpedia.org/resource/Colorfulness</a>
scale	an ordered reference standard	[Synset('standard.n.01')]	<a href="http://dbpedia.org/resource/Scale_(music)">http://dbpedia.org/resource/Scale_(music)</a>





## AudioCommons

### D2.1 Requirements report and use cases

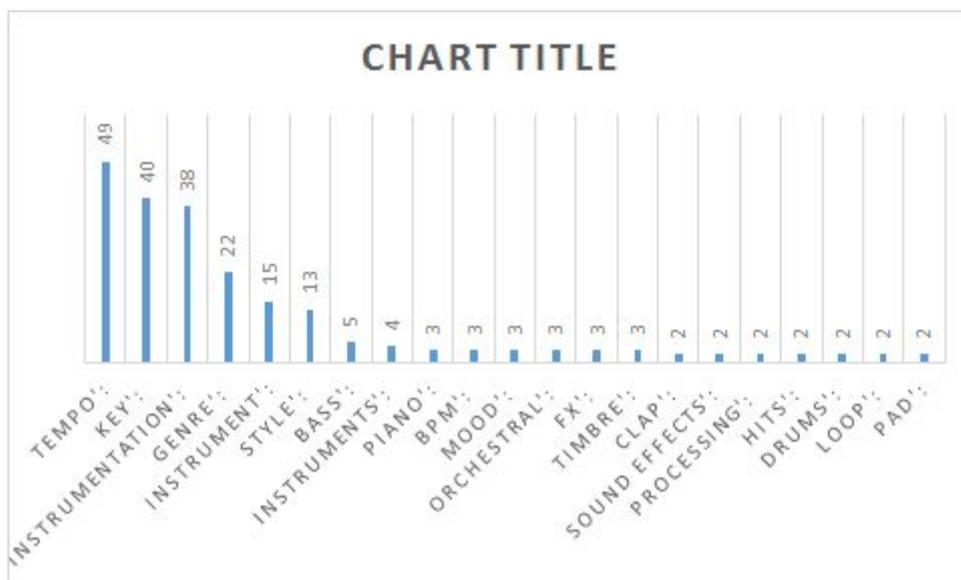
song	a short musical composition with words	[Synset('musical_composition.n.01')]	<a href="http://dbpedia.org/resource/Song">http://dbpedia.org/resource/Song</a>
soundscape	a piece of music considered in terms of its component sounds	music	<a href="http://dbpedia.org/resource/Soundscape">http://dbpedia.org/resource/Soundscape</a>
speed	distance travelled per unit time	[Synset('rate.n.01')]	<a href="http://dbpedia.org/resource/Speed">http://dbpedia.org/resource/Speed</a>
tempo	(music) the speed at which a composition is to be played	[Synset('musical_time.n.01')]	<a href="http://dbpedia.org/resource/Tempo">http://dbpedia.org/resource/Tempo</a>
timbre	(music) the distinctive property of a complex sound (a voice or noise or musical sound)	[Synset('sound_property.n.01')]	<a href="http://dbpedia.org/resource/Timbre">http://dbpedia.org/resource/Timbre</a>
tonality	any of 24 major or minor diatonic scales that provide the tonal framework for a piece of music	[Synset('musical_notation.n.01')]	<a href="http://dbpedia.org/resource/Tonality">http://dbpedia.org/resource/Tonality</a>
urban	located in or characteristic of a city or city life	character	<a href="http://dbpedia.org/resource/Urban_area">http://dbpedia.org/resource/Urban_area</a>
year	a period of time occupying a regular part of a calendar year that is used for some particular activity	period	





# APPENDIX 2: Most common terms

Most common music attributes ( by 129 responses)



Most common perceptual attributes (by 120 responses)







## APPENDIX 3: Audio Commons Initiative Survey

### Audio Commons Initiative Survey on Creative Interaction with Audio Content

We would like to thank you for your participation in this survey which is conducted as part of our European research and innovation project Audio Commons Initiative. The Audio Commons Initiative aims to bring Creative Commons audio content to the creative industries (Creative Commons copyright licenses provide a standardised way to give the public permission to share and use creative work on conditions defined by the content creators). For more information, please visit: [www.audiocommons.org](http://www.audiocommons.org)

#### 1. Survey goals

Through this survey we would like to gain insights into how creative users interact or wish to be able to interact with audio content whether e.g. samples, loops, sound effects or entire songs. The results will help us defining strategies and technologies to improve workflows for music artists, sound designers, audio engineers and producers. The overall aim is to develop a new ecosystem which will allow users to fetch audio content in novel ways and to establish a unified communication endpoint to access audio content from various online resources.

The survey consists in a set of multiple choice, checkbox and free form (open ended) questions. It should take about 5-10 minutes to complete. Participation is entirely voluntary and you have the right to withdraw from the study without any disadvantages. Your inputs and suggestions would however be extremely valuable to us and we would like to encourage you to complete the entire survey.

#### 2. Anonymity and data protection

Survey data use and storage will comply the standards of the UK Data Protection Act (1998). The confidentiality of personal information and the anonymity of all volunteers will be preserved. All responses will remain anonymous and cannot be linked to participants. The responses and their analyses may be used in academic publications in an anonymised way.

#### 3. Contact details

If you have questions, comments or suggestions about this research, do not hesitate to contact Lasse Vetter ([l.vetter@qmul.ac.uk](mailto:l.vetter@qmul.ac.uk)), George Fazekas ([g.fazekas@qmul.ac.uk](mailto:g.fazekas@qmul.ac.uk)) and Mathieu Barthet ([m.barthet@qmul.ac.uk](mailto:m.barthet@qmul.ac.uk)).

If you would like to receive news about our project and the tools we will develop or consider taking part in future surveys, please feel welcome to join the audiocommons-friends mailing list by following the instructions at:

<https://groups.google.com/a/llista.upf.edu/forum/#!forum/audiocommons-friends>

The Audio Commons Initiative team

\*Required





**1. I am \***

*Mark only one oval.*

- a professional
- an amateur

**2. What I mainly work on is (please tick all that apply) \***

*Tick all that apply.*

- Music production (mixing, mastering)
- Music production (composition, recording)
- Audio for games
- Audio for movies and/or TV
- Music performance
- Other: .....

**3. My computer for making sound and/or music is \***

*Mark only one oval.*

- always connected to the Internet when I work on music
- is not or rarely connected to the Internet

**4. I often work with a sound library of approximately (please tick all that apply) \***

*Tick all that apply.*

- 100 Audio Samples/Loops/Sound FX
- 1000 Audio Samples/Loops/Sound FX
- 10.000 Audio Samples/Loops/Sound FX
- more than 10.000 Audio Samples/Loops/Sound FX
- I don't use any sound library
- I use the default sound library provided by my software

**5. Audio content from web databases \***

*Mark only one oval.*

- is something I've never used
- is something I sometimes use
- is the way I find my files





**6. I spend more time \***

*Mark only one oval.*

- finding the right file
- modifying my chosen file with audio processors
- none of the above

**7. For me, processing the audio files \***

*Mark only one oval.*

- is a necessity I would avoid if I had the right file at hand
- is the creative part of my work
- Other: .....

**8. For my work I mainly use (please tick all that apply) \***

*Tick all that apply.*

- whole songs
- audio loops (e.g., Drums/Bass/Keys)
- sound effects (e.g., traffic noise/barking dogs/keyboard layers)
- Other: .....

**9. I would prefer to search for and retrieve audio (please tick all that apply) \***

*Tick all that apply.*

- from within my Digital Audio Workstation
- through a standalone app
- through a web browser
- Other: .....

**10. My audio files should \***

*Mark only one oval.*

- have as much metadata as possible (such as Genre/Artist/Key/Tempo/etc...)
- be named properly, that's enough for me (e.g., Funky\_Bass\_in\_C.wav)
- Other: .....

**11. I would like to have (please tick all that apply) \***

*Tick all that apply.*

- an unlimited amount of audio files that I can query in a smart way
- a limited set of files that I know
- curated content providing quality audio files
- Other: .....





12. When I search for music-related data I want to (please tick all that apply) \*

Tick all that apply.

- query using suggested keywords from the interface (Drop Down Lists or similar)
- use keywords (Google style)
- use natural language (for instance, whole sentences)
- use a query language (like programming)
- use a non-verbal graphical interface (e.g., two-dimensional XY controller)
- Other: .....

13. When I search for new audio files (please tick all that apply) \*

Tick all that apply.

- I use musical attributes (e.g., related to Key, Tempo or Instrumentation)
- I use perceptual attributes (e.g., "Punchy", "Bright", "Powerful")
- Other: .....

14. Please indicate the musical attributes you use the most to search for audio files, separated by commas (e.g., key, tempo, instrumentation).

.....

.....

.....

.....

.....

15. Please indicate the perceptual attributes you use the most to search for audio files, separated by commas (e.g., "punchy", "warm", "powerful").

.....

.....

.....

.....

.....

16. Please indicate the combinations of musical and perceptual attributes you use the most to search for audio files, separated by commas (e.g., "punchy bass in E", "soft drums 100 bpm").

.....

.....

.....

.....

.....





**17. Please indicate your biggest frustrations about current methods for browsing and retrieving audio content, if any.**

.....  
.....  
.....  
.....  
.....

**18. Please describe the number one feature you would like to have in a tool for browsing and retrieving audio content.**

.....

**19. Please describe how you would ideally search/filter audio content from a large database (indicate audio properties, workflow or other relevant aspects that come to mind).**

.....  
.....  
.....  
.....  
.....

**20. Please indicate as many metadata type as possible you would like to use to search/filter audio content, separated by commas (e.g., genre, mood, bpm, types of samples, timbral properties, production equipment, etc.).**

.....  
.....  
.....  
.....  
.....

**21. Gender**

*Mark only one oval.*

- Female
- Male

**22. My age**

.....





**23. Country of residence**

---

**24. I have worked with audio for approximately**

*Mark only one oval.*

- <2 years
  - 2-3 years
  - 4-6 years
  - 7-9 years
  - 10 or more years
- 

