Deliverable D2.3
Final Ontology Specification

Grant agreement nr	688382
Project full title	Audio Commons: An Ecosystem for Creative Reuse of Audio Content
Project acronym	AudioCommons
Project duration	36 Months (February 2016 - January 2019)
Work package	WP2
Due date	31 July 2017 (M18)
Submission date	31 July 2017 (M18)
Report availability	Public (X), Confidential ( )
Task leader	QMUL
Authors	Miguel Ceriani, György Fazekas, Johan Pauwels, Mathieu Barthet, Mark Sandler
Document status	Draft ( ), Final (X)

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

Table of contents

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

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 develops the Audio Commons Ontology and the Audio Commons API specification. Building the Audio Commons Ontology required gathering knowledge from real and anticipated user requirements and already existing ontologies in the music and audio domain, as well as knowledge about the workflows and tasks that people in the audio/music domain use in their everyday professional activities. The ontology presented here pushes the boundaries in communication and interoperability between stakeholders in the audio domain, for instance by introducing a layered conceptualisation in the broader field of audio metadata modelling as well as facilitating novel functionalities in our API and end-user tools.

This deliverable informs further development of the Audio Commons API specification. The Audio Commons API defines how the different components of the Audio Commons Ecosystem are technically interconnected (i.e., how production tools are able to access Audio Commons content and how users are able to communicate for the licensing process). The ontology is intended to be used primarily by the Audio Commons semantic mediator, an application layer that orchestrates different services and components comprising the Audio Commons Ecosystem. This provides the required technological layer for mediating between service components (e.g. sound, music or metadata download or search facilities), and provides software packages and guidelines to facilitate the incorporation of new actors in the ecosystem. It also informs research on rights management and help focus the work on sound and music analysis algorithms and end user prototypes.

The current version of the ontology is the final version before work commences on integrating it in our API layer. The version builds on the previous drafts (presented in D2.2) with a stronger focus on a smaller domain and practical considerations. It also builds on further analysis of existing ontologies and vocabularies that describe music and audio related domains, and on current development and use of the Audio Commons API and specific APIs (Jamendo, Freesound, Europeana Sounds).

This deliverable presents the first complete version of the Audio Commons ontology and demonstrates how it can be used specifically in the context of the Audio Commons mediator API. This version of the Audio Commons ontology is final in the sense that is mature enough to be published and used. This does not however prevent further changes in the ontology to incorporate new requirements and desiderata. The ontology itself already provides different possible extension points to incorporate as many as possible external requirements. It also provides for the integration of other models, such as specific taxonomies focussing on subdomains relevant to Audio Commons, for instance, the classification of environmental or urban sounds presented elsewhere.

Instead of providing a formal specification in this document, we focus on the theoretical and practical considerations in the design of the ontology, contrast the ontology with other related ontologies and provide rationale for design decisions. The ontology is available online for use (the current version is published via GitHub), evaluation and discussion at https://github.com/AudioCommons/ac-ontology It is due to be published for broader application in the context of the Semantic Web, possibly outside of the Audio Commons Ecosystem, using a persistent URL and corresponding namespace.

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

Background

This deliverable is part of task T2.2 from work package WP2. The purpose of WP2 is to develop and evaluate the Audio Commons Ontology and the mechanisms for API orchestration in the context of the Audio Commons ecosystem. Building the Audio Commons Ontology (and ontologies in general) requires a clear definition of scope and extensive knowledge about the domain that the ontology describes. 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. Deliverable D2.1 presented the results of the survey that we circulated among professionals and amateurs in the music creation domain. Participants were asked to answer questions about the daily tasks which 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 were used as a starting point for the Audio Commons ontology. Another important source of information for the Audio Commons ontology were the already existing music and audio related ontologies. Those ontologies were analysed and connection points were identified. The first Audio Commons ontology draft, presented in the previous related deliverable D2.2, has been developed from these connection points.

This deliverable presents the first complete definition of the Audio Commons ontology. A coherent specification that will allow the adoption of a common data model by the end user tools of the Audio Commons ecosystem.

The Audio Commons ontology, together with the Audio Commons API specification, defines how the different components of the Audio Commons Ecosystem are 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 enable the orchestration of the different services or components integrating the Audio Commons Ecosystem. A semantic mediator acts as the required technology layer for the interconnection of the different components.

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

1 Introduction

1.1 Main objectives and goals

The main goal of the Audio Commons ontology is to make the Web service orchestration possible by providing necessary knowledge about entities in that process. The previous deliverable related to the ontology, D 2.2 presented the first draft of the Audio Commons ontology. The aim of that deliverable was to explore how an ontology could coordinate the multiple requirements and desiderata related to the Audio Commons ecosystem. The design of the draft ontology and related deliverable was deliberately broader in scope, proposing generic classes and properties to cover multiple aspects, while leaving the detailed practical specification to be defined after a series of trials completed within the context of other work packages that refine the requirements and experiences gained from them

This deliverable aims instead at publishing a coherent and complete ontology that deal with the most stringent requirement of the Audio Commons ecosystem, that is having a common data model to search and interact with audio resources. A model to be used in the Audio Commons Web API. The scope of this deliverable is thus in some aspects narrower than D 2.2, in order to focus on the specification of a complete and usable.

In this sense, this deliverable is complementary to D 2.2. The broader analysis of actions and provenance that was described there remains still valid as input for potential future extensions of the model presented here. It should be stressed that, while the role of this deliverable in the project is to present the final version of the ontology, this does not mean that the ontology will not possibly evolve later on. Used ontologies necessarily evolve over time, especially if, as in the Audio Commons project, a lot of the use cases still have to be fully analysed.

This deliverable features four main chapters. Chapter 2 briefly describes the existing music and audio related ontologies and comments about the potential usefulness of each ontology for the Audio Commons ontology and within the Audio Commons Ecosystem (ACE). Chapter 3 presents the main entities of Audio Commons ontology explaining the rationale behind his design. Chapter 4 shows examples of an important feature of the Audio Commons ontology, that it how sounds may be classified using multiple sound/music taxonomies. Chapter 5 shows how the ontology can be used in the context of the Audio Commons API.

1.2 Terminology

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

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

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, an 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.

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.

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

2 Related Ontologies and Data Models

An ontology defines a set of representational primitives which we can use to 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.

One of the main imperatives of the Semantic Web[1] is that ontologies should not be built entirely from scratch and that previously used vocabulary should be reused as much as possible. For the Audio Commons ontology, outside of the knowledge collected from the survey, the source of knowledge comes from already existing ontologies and data models describing music and audio related domains.

In the rest of this chapter the most relevant existing models and data models are described. Some of these are extended and others integrated in the context of the Audio Commons ontology. For completeness, all relevant models are included, even the ones that were already discussed in deliverable D2.2.

[1] The reuse of existing vocabularies and ontologies are core to Semantic Web interoperability. This notion is in contrast with highly specialised, domain specific data models. Our approach in this project is to balance the requirements between large scale interoperability and domain specificity.

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

2.1 Music Ontology

The aim of the Music Ontology framework [Raimond] is to provide a comprehensive, yet easy to use and easily extended domain specific knowledge representation for describing music related information. Integration of music related resources (Web services and data repositories) on the Semantic Web, and facilitation of service integration and data communication in distributed music processing environments are among its existing applications. It has certain properties which make it particularly suitable as basis for a general semantic audio information management framework as well as data collection in recording and production. For instance, it relies on, and extends the full FRBR (Functional Requirements for Bibliographic Records [FRBR]) model, and provides an event based conceptualisation of music production workflows as shown in Figure 5.

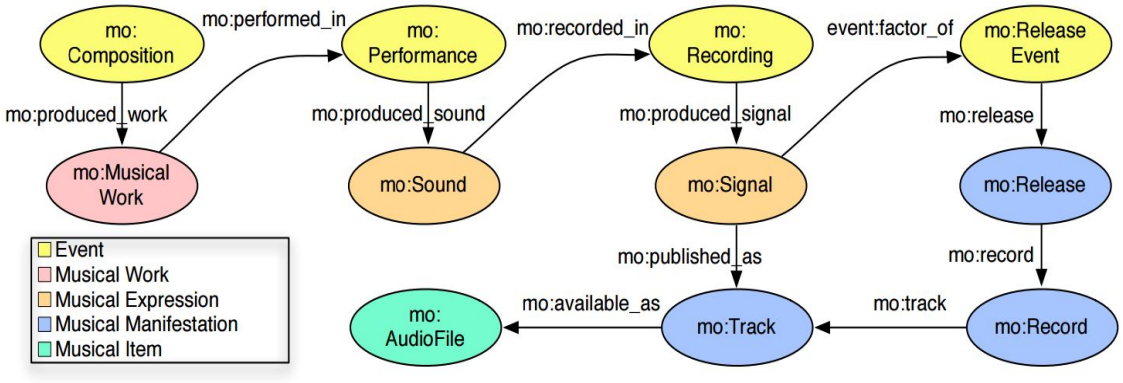

Figure 1. Music Production Workflow Model [Fazekas]

The Music ontology can represent the following things [Fazekas]:

- Editorial metadata: Concepts and relationships involving artists, bands, labels, albums, tracks, audio files or downloads and their identifiers in various databases.
- Music production workflow: The life cycle of musical works from composition through performance, to the produced sounds and recorded signals and their publication.
- Event decomposition: Further details about particular events in the production workflow such as individual performances by different musicians in a recording.
- Content annotation: Audio signals and temporal annotation of their content.

In the broader context (outside the domain music production workflow), the Music ontology lacks the connection with the concepts coming from the legal and intellectual property side of the audio

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

domain, as well as web and web services concepts or a generic sound model that can represent audio content other than music pieces.

2.2 Jamendo API Data Model

It is worth considering the underlying data model of Jamendo API, as is one of the Web APIs that are being integrated through the Audio Commons mediator.

Its main concept is the *track*, that a represented a published track, possibly from an *album*. An album has as author an *artist*, that may have a *location*. A track, album, or information of an artist is uploaded by a *user*. A user may also create *playlists*, that are collection of tracks. From a track, playlist, or album some multiple *files* with different formats or properties may be downloaded.

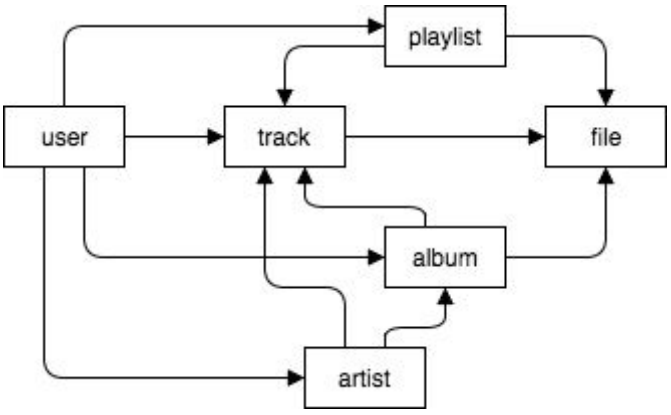

Figure 2. Representation of the main entities of the data model underlying Jamendo API and their relationships.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 688

Figure Jamendo API Read Methods

2.3 Freesound API Data Model

Freesound API is another Web API that is being integrated through the Audio Commons mediator.

Its main concept is the *sound*, that a represented an audio clip uploaded to the site by a *user*. Sounds can be grouped into *soundpacks*.

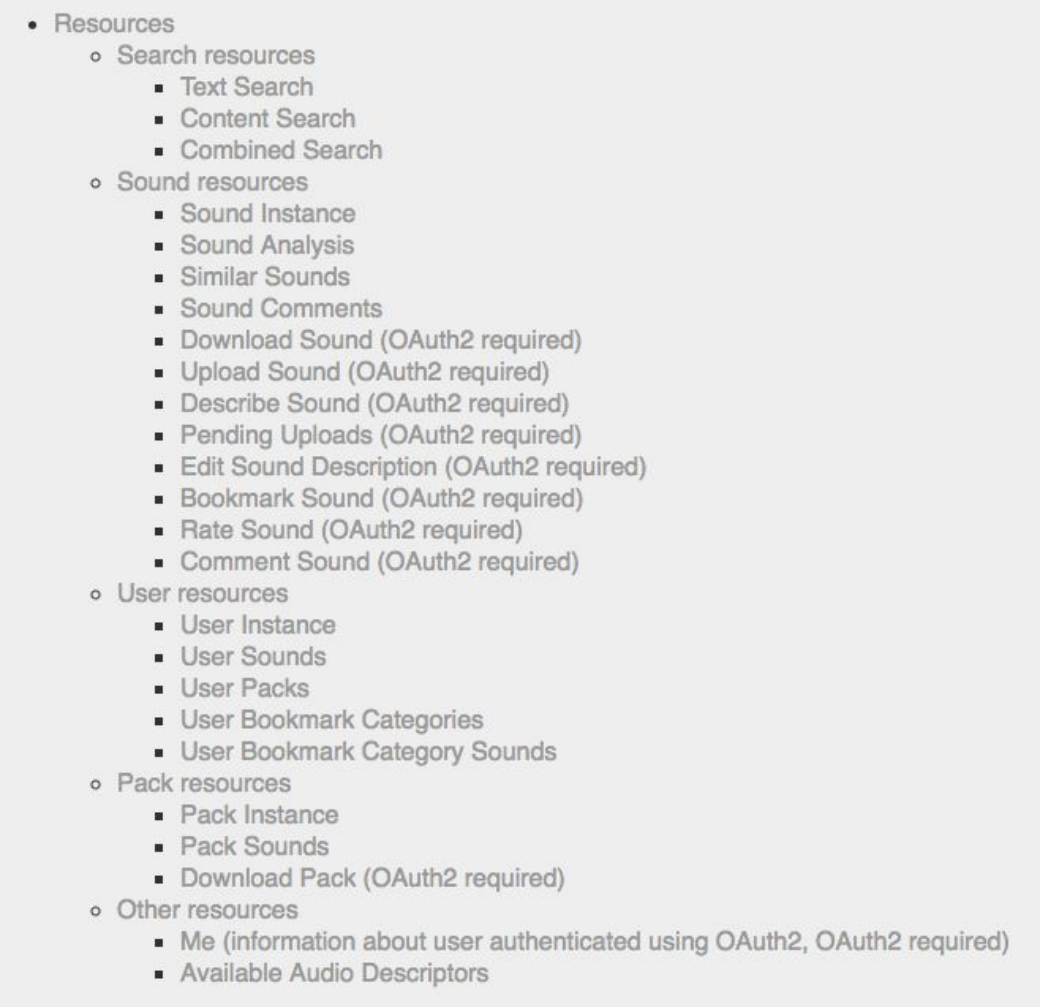

Figure 3. Screenshot of Freesound API documentation

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

2.4 Europeana Sounds sound data model

Europeana (http://www.europeana.eu) is an initiative co-financed by the European Union to incentive member states to digitise and make Europe's cultural heritage accessible through a digital platform. The portal provides free access to a wide array of digital content including audio. In the context of Europeana, Europeana Sounds (http://www.europeanasounds.eu) is a project funded by the European Commission whose aim is to increase the amount of audio content available via Europeana, to improve geographical and thematic coverage by aggregating recordings with widespread popular appeal, to enrich this content and to promote the creative reuse of the content. Europeana Sounds strives to add meaningful contextual knowledge and medium-specific metadata to 2 million items in Europeana's audio and audio-related collections, developing techniques for cross-media and cross-collection linking.

Work carried out in the context of Europeana Sounds consisted in identifying a set of requirements essential to the description of sound objects from the Europeana collection and extending the Europeana Data Model (EDM) with new properties to better accommodate audio and audio-related objects [Europeana]. Figure 8. shows the modelling of the concept of a Web resource in Europeana, where anything that is part of the Europeana ecosystem will "live" as a Web resource. Similar concepts will be present in the Audio Commons ontology.

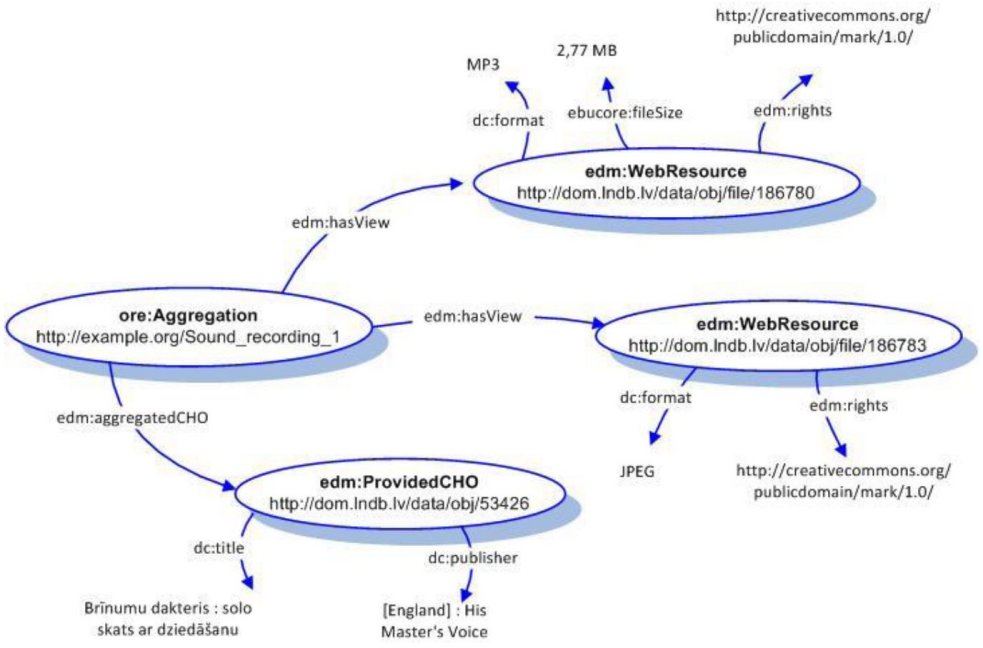

Figure 4. Web Resources in Europeana Sound [Europeana]

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

2.5 EBU Core Ontology

The European Broadcasting Union (EBU), developed EBU Core [EBUCore], a minimum and flexible list of attributes to describe audio and video resources for a wide range of broadcasting applications including archives, exchange and production in the context of a Service Oriented Architecture. It is useful as a metadata schema to describe audio resources and formats. There are aspects of the EBU ontology relevant to Audio Commons which are reused as described in Chapter 3.

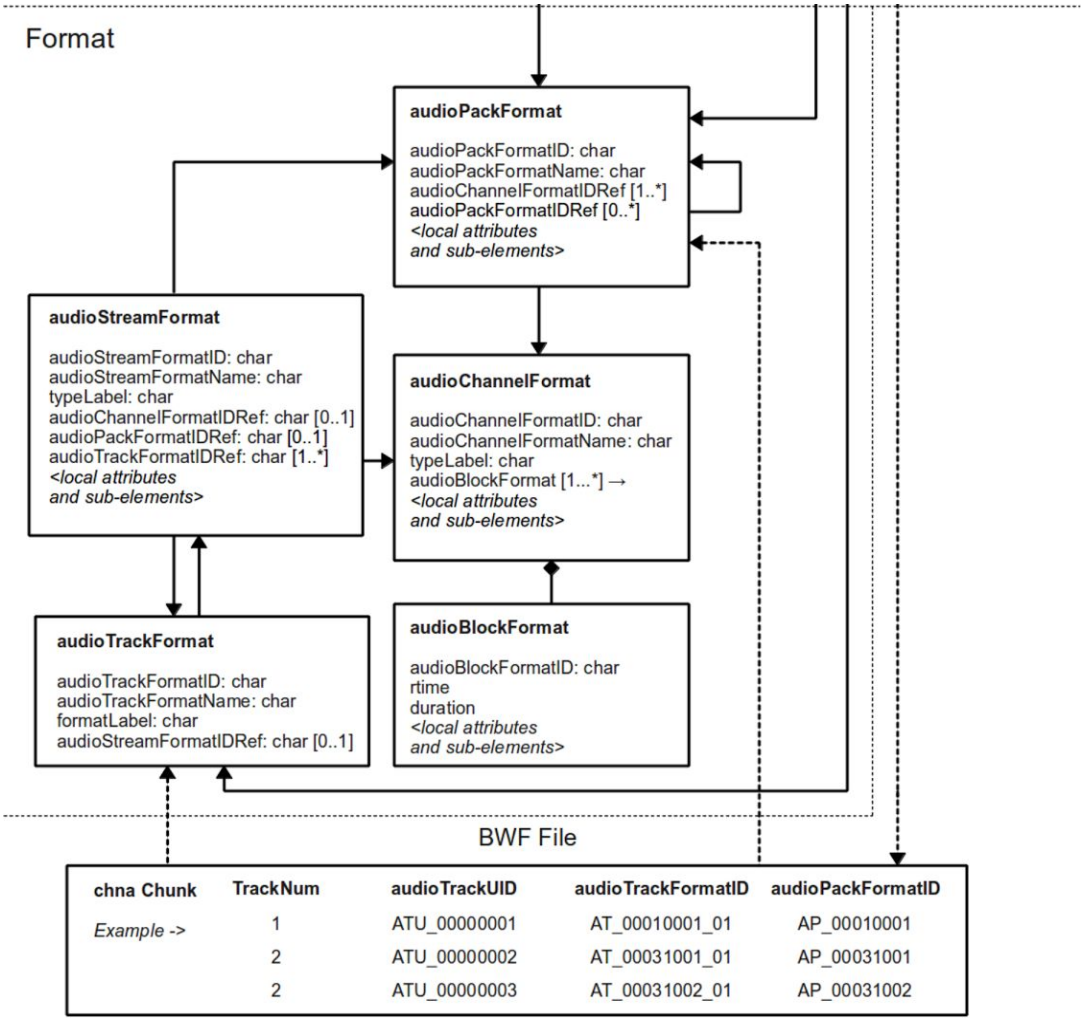

Figure 5. Entities and attributes used to represent formats in EBU Core

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

2.6 Taxonomies: SKOS Ontology

The Simple Knowledge Organization System (SKOS), is a data model for sharing and linking knowledge organization systems via the Web.

Many knowledge organization systems, such as thesauri, taxonomies, classification schemes and subject heading systems, share a similar structure, and are used in similar applications. SKOS captures much of this similarity and makes it explicit, to enable data and technology sharing across diverse applications [SKOS].

SKOS is used in connection with the Audio Commons ontology to represent multiple classifications of the concepts defined in the ontology, specifically the taxonomical organisation of content types that require the use of potentially several domain specific taxonomies in tandem.

With respect to hierarchies and classification directly defined through ontologies (e.g., through subclass hierarchies), SKOS allows to define less stringent semantic relationships. These kind of classifications are best suited when the hierarchies are meant for user categorization but not necessarily for full fledged inference. Furthermore, SKOS provides way to coordinate multiple taxonomies without necessarily combining them.

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

2.7 Creative Commons and Licensing Ontology

Creative Commons is a global nonprofit organization that enables sharing and reuse of creativity and knowledge through the provision of free legal tools. Creative Commons licenses provide an easy way to manage the copyright terms that attach automatically to all creative material under copyright. Those licenses allow that material to be shared and reused under terms that are flexible and legally sound. Creative Commons offers a core suite of six copyright licenses (Figure 9). The Audio Commons ontology will map relevant entities of Creative Commons to the concepts in the AudioComons ecosystem. Information about the Creative Commons licensing framework and its implications in the context of Audio Commons is provided in the publically available deliverables D3.1 and D3.2 [Del].

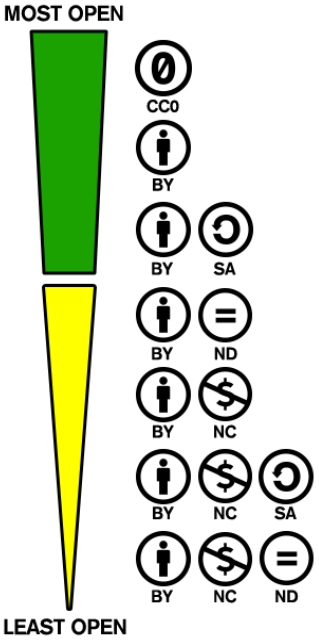

Figure 6. Core suite of six copyright licences [CreativeCommons]

Furthermore, Creative Commons designed Creative Commons Rights Expression Language (CC REL) [CC-REL], a basic ontology to describe license information and to attach licenses to works. The CC REL ontology is used by the Audio Commons ontology to add license information to audio content.

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

3 The Audio Commons Ontology

This section describes the Audio Commons ontology. Rather than providing a formal specification in this document, we focus on practical and theoretical considerations in the design of the ontology, contrast the Audio Commons ontology with other related ontologies and provide rationale for design decisions. The formal specification is provided as an online document using the Ontology Web Language (OWL)[2]. The ontology (as shown on Figure 1) builds upon existing audio and audio related (or relevant) ontologies. Building an ontology that would encompass the whole audio domain (and all other domains that are connected with the audio domain) in all its complexity would be a very significant task that is beyond the scope of this work. The Audio Commons ontology will, for this reason, be an implementation driven ontology that is evaluated and evolved in use. This means that the Audio Commons ontology will be growing depending on the demand for new services in the Audio Commons ecosystem. Nevertheless, this first version of the ontology will set the path for all future iterations of the ontology presented here, because it defines high-level concepts that are relevant and shown to be broadly accepted by domain experts in our previous analyses of the literature and practical implementations. The ontology also describes how these high-level concepts should be interconnected.

The core of the Audio Commons ontology is a generalization of the part of the Music Ontology to the substantially broader context of audio description and publishing (i.e. our domain is not restricted to music). At the same, the Audio Commons ontology is, as the Music Ontology, a specialization of the FRBR ontology. The main Audio Commons classes and properties sit thus “in the middle” between the corresponding more general elements of the FRBR ontology and the more specific ones of the Music Ontology.

The generalization with respect to the Music Ontology is necessary to represent resources and concepts that are not necessarily in the musical domain, for instance the recording or synthesis and publishing of non-musical sounds. The formal description of this generalization in the Audio Commons ontology preserves the interoperability with the Music Ontology for the cases in which musical resources and concepts are described.

The following diagram (Fig.17) shows the most general classes and properties of Audio Commons ontology and their relationship with elements of the FRBR and the Music ontologies. Following the FRBR model, the following classes have been defined:

- **ac:AudioExpression**. Represents the specific intellectual or artistic form that a work takes each time it is 'realized', in the audio domain. This includes a musical performance, the recording or synthesis of music or sounds.

[2] https://github.com/AudioCommons/ac-ontology

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

- **ac:AudioManifestation**. Represents the physical embodiment of an audio expression. This includes a published musical track, sound, album, …
- **ac:AudioItem**. Represents a single exemplar of an audio manifestation. This includes a copy of a CD, vinyl or (more interesting for the Audio Commons context) a specific media file available for download.

The FRBR class Work, representing a distinct intellectual or artistic creation on a more conceptual level, has not been specialized in Audio Commons because this does not seem to be directly relevant for the Audio Commons ecosystem at this stage. This class is used in the Music Ontology to represent the common creation act between different expressions, for instance, different drafts of a symphony, or its existence in the composer's mind at its most abstract level. If needed for musical resources, the mo:MusicalWork class can still be used.

In the Music Ontology some specific properties (e.g., mo:genre and mo:instrument) are used orthogonally to classify both musical works, expressions, manifestations, and items, attaching them to specific instances of some classification schema (e.g., instances of mo:Genre and mo:Instrument). In the Audio Commons ontology these properties are generalized by the ac:audioCategory property that associates any audio expression or manifestation (or item, but the practical use of the latter case seems limited) to some generic ac:AudioCategory. Chapter 4 describes how different specific taxonomies can be plugged in through this formalism. Specific subclasses and properties related to audio expressions, manifestations, and items will be described in the following sections of this chapter.

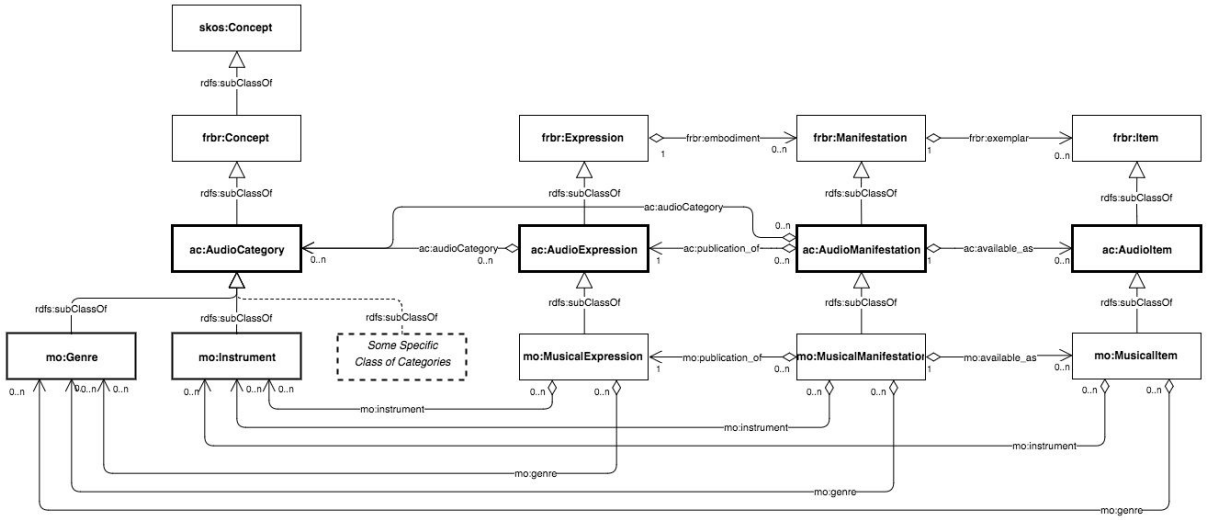

Figure 7. Generic classes of Audio Commons ontology representing important concepts in the conceptualisation of the audio domain. (Arrows represent exemplar relationships, numeric ranges represent cardinality constraints.)

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

3.1 Audio Manifestations: Clips and Collections

The class ac:AudioManifestation a generalisation of (superclass of) a central entity in the Audio Commons ecosystem, ac:AudioClip. An instance of ac:AudioClip is any audio segment that has been somehow published or uploaded for consumption, for example a track on Jamendo or a sound on Freesound.

Different models provide ways to represent collections of audio clips, as tracks organized records/albums in the Music Ontology and Jamendo or sounds organized in sound packs in Freesound. The Audio Commons ontology offers thus an abstraction called ac:AudioCollection, which is itself another subclass of ac:AudioManifestation. A third class, ac:AudioCollectionNode, represent a single node of the collection, offering local information like the index in the collection and pointers to the previous and next nodes. The separation between the collection node and its actual content (say, an ac:AudioClip) permits the same content to be shared in multiple collections (as sound packs of Freesound or playlists of Jamendo). Furthermore, the content of each node of a collection is not just an ac:AudioClip but any ac:AudioManifestation. Collection can thus contain other collections to support specific cases like the Music Ontology model, in which a mo:Release can contain multiple mo:Record(s) that can in turn contain multiple mo:Track(s). Finally, the property ac:principalContainerNode permits to attach a specific “principal” collection to an ac:AudioManifestation, for example the corresponding album to a music track.

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

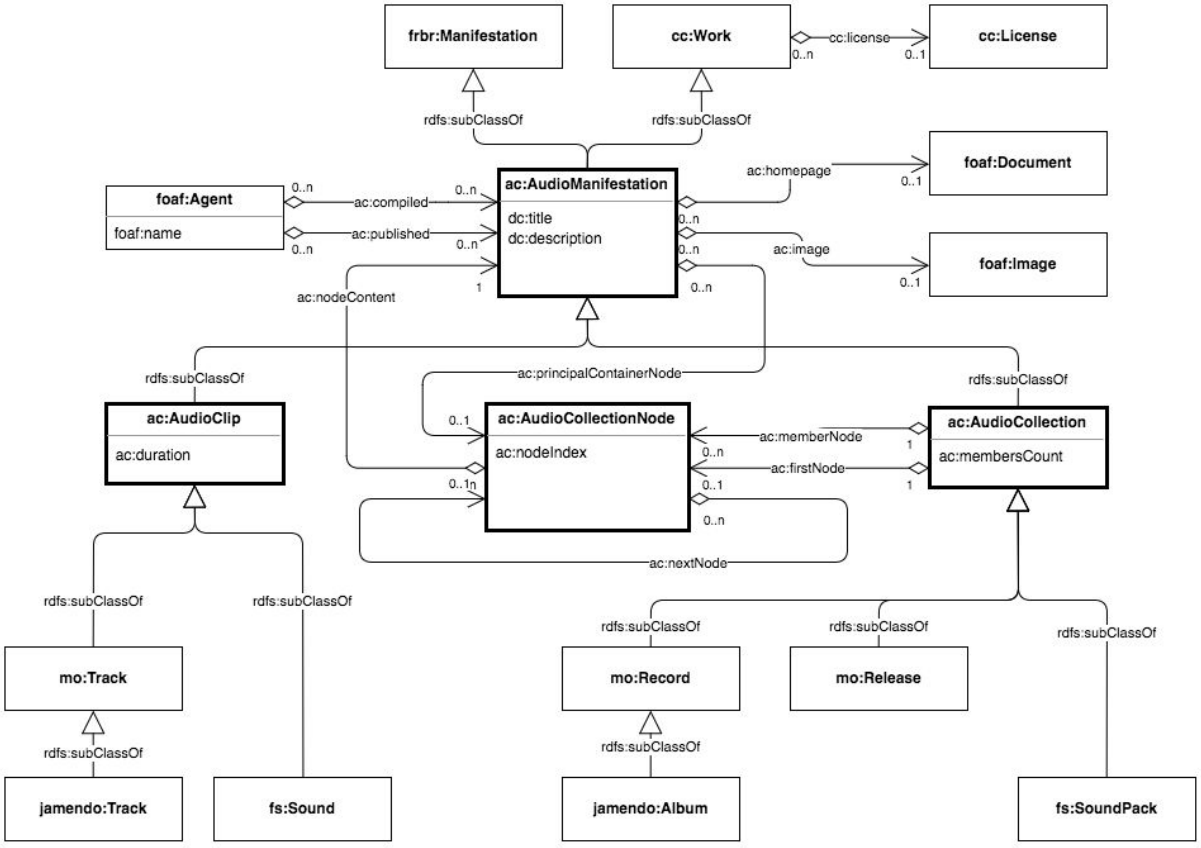

Figure 8. Representation of audio manifestations in the Audio Commons ontology.

For interoperability, the widely used Dublin Core vocabulary may be used for basic meta-data, including title and description of every described ac:AudioManifestation. Other information more specific to the domain is represented through properties that generalize audio specific predicates defined in the Music Ontology:

- ac:compiled and ac:published, generalization of mo:compiled and mo:published, that associate an agent with manifestations he/she/it respectively created or published;
- ac:homepage and ac:image, generalization of mo:homepage and mo:image, specialization of foaf:isPrimaryTopicOf and foaf:depiction, that associate a manifestation with its page on a site (e.g., Jamendo, Freesound, Europeana) or with its depiction (e.g., the cover art of an album);
- ac:duration, generalization of mo:duration, that associate an audio clip with its duration expressed in milliseconds.

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

To represent licensing information, the Creative Commons licensing ontology is reused. The ac:AudioManifestation class is also a subclass of cc:Work, so that the property cc:license can be used to attach a license to any audio manifestation.

3.2 Audio Items: Files and Audio Encoding and Container Formats

The class ac:AudioItem represents any concrete exemplar of an audio manifestation. In the Audio Commons ecosystem the main exemplars are the files that contain the audio manifestations serialized in some audio format. This case is represented by the class ac:AudioFile subclass of ac:AudioItem and superclass of mo:AudioFile. To represent the information related to the audio file and its format, part of the EBU core ontology is reused. The class ac:AudioFile is thus a subclass of ebu:MediaResource too. Specific formats can be represented as instances of ebu:EncodingFormat and ebu:ContainerFormat. There is a predefined set of instances with the most common formats, for each of these two classes.

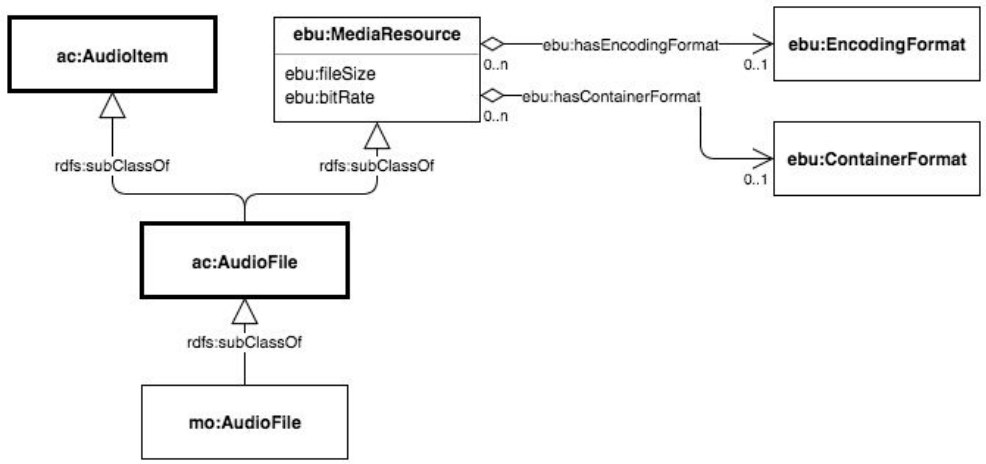

Figure 9. Representation of audio files and encoding as well as container formats in the Audio Commons ontology.

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

3.3 Audio Expressions: Digital Signals

Digital signals representing audio are of primary importance because they have become an almost exclusive distribution mechanism of of content, especially online. The class mo:MusicalExpression in the Music Ontology includes as subclass mo:Signal and more specifically mo:DigitalSignal. The class ac:DigitalSignal, subclass of ac:AudioExpression, is the generalization of mo:DigitalSignal to all the audio domain.

Here, ac:DigitalSignal is the representation of the digital signal corresponding to an audio clip. Digital signals are represented at the relatively abstract layer corresponding to expressions in FRBR. This conceptualisation was chosen because it pertains to the weakest ontological commitment with respect to how the signal is represented or encoded and where it is situated in a specific workflow. The properties ac:sample_rate, ac:bitsPerSample, and ac:channels, generalization of the homonymous properties in the Music Ontology, associate a signal with its basic indicators.

The property ac:publication_of can be used to associate an ac:AudioClip with the corresponding digital signal. The property ac:encodes instead, associates an ac:AudioFile with the encoded digital signal. The latter property works as a shortcut of traversing the inverse of ac:available_as and ac:publication_of.

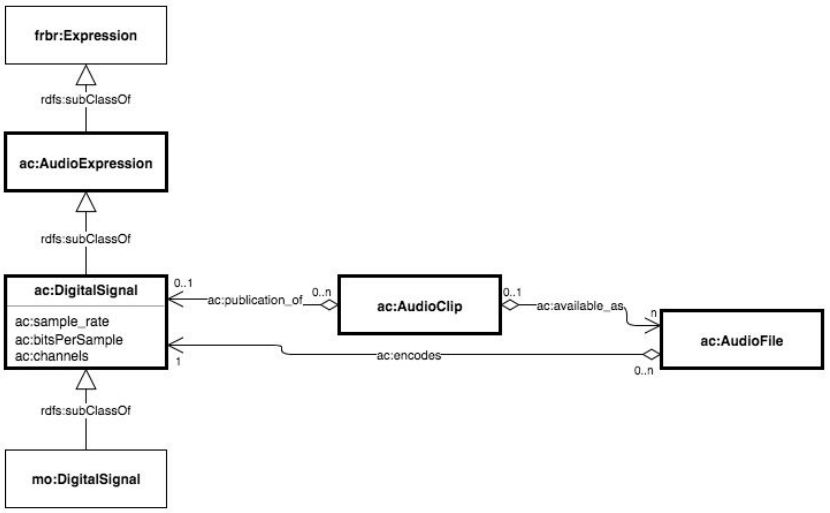

Figure 10. Representation of audio expressions in the Audio Commons ontology.

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

3.4 Events: Production and Publication

In the conceptualisation of the Music Ontology, the description of temporal events is crucial to describe transitions in the production workflow. For instance, the act of recording or publication enables the transition between expression and manifestation layers. The AC ontology adopts this fundamental conceptualisation. The Event ontology permits to model different aspects of temporal events (both instantaneous as well as events having a finite or infinite duration) that are part of the audio production and publication workflow. Using event:Event, details of the event such as its location in time and space, its factor, and its products may be explicitly described. Moreover, the events can be composed using the property event:sub_event, to build complex events.

The Music Ontology offer subclasses of event:Event for two specific actions that are interesting also for the audio domain in general:

- mo:Recording, the act of recording a mo:Sound (e.g., the sound created by a musical band that is playing) so that it can be represented as a mo:Signal (that could be either a mo:AnalogSignal or a mo:DigitalSignal);
- mo:ReleaseEvent, the event representing the public release of a piece of work (e.g., the release of a new album by a band).

The Audio Commons ontology generalize the class mo:Recording through defining the class ac:Recording. It also considers a the case in which the signal could had been synthesised rather than recorded, through the class ac:Synthesis. The common superclass ac:SignalProduction represents the abstract concept of the production of a signal.

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

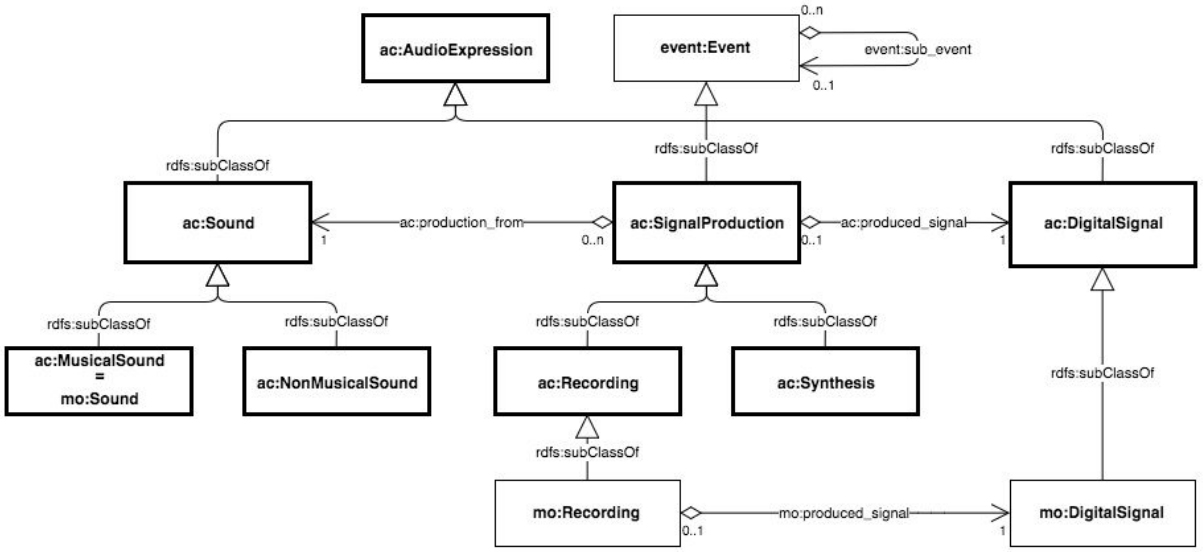

Figure 11. Representation of signal production in the Audio Commons ontology.

The mo:ReleaseEvent is generalized in Audio Commons ontology by the class ac:AudioPublication. This class provides a way to describe a publication event related to any audio manifestation (single audio clips as well as collections).

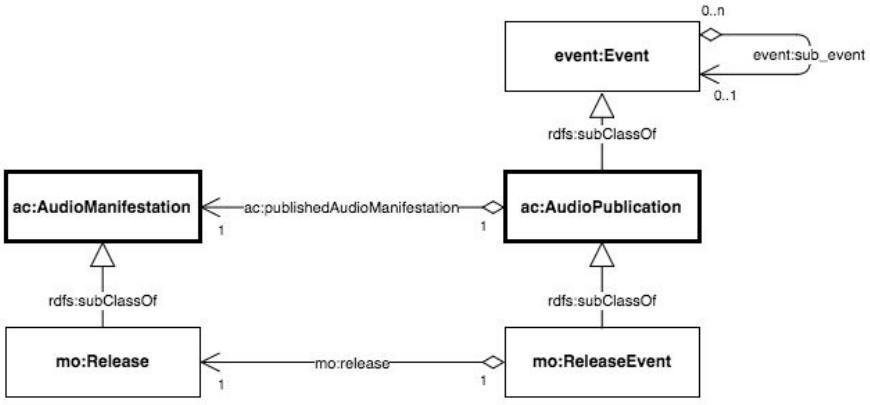

Figure 12. Representation of a publication event in the Audio Commons ontology.

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

3.5 Audio Collections as Lists

The ac:AudioCollection entity provides a mechanism to describe collections of audio content in a way that is coherent and integrated with the rest of the Audio Commons ontology. In the context of the Audio Commons ecosystem it is important to provide ways to convey information about instances in a concise way. However, the full serialization of an instance of ac:AudioCollection is an explicit representation of a linked list and tends to be quite convoluted no matter what specific RDF syntax is used.

For the standard list class rdf:List, instead, several RDF syntaxes provide ways to represent them in a concise way. For this reason, as well as for interoperability reasons, the ac:AudioCollection is related to the rdf:List.

Specifically:

- the ac:AudioCollectionNode is a subclass of rdf:List;
- the property rdf:rest associates every instance ac:AudioCollectionLastNode with rdf:nil;
- for every instance ac:AudioCollectionNonLastNode, the property rdf:rest is equal to ac:nextNode;
- for every instance ac:AudioCollection, the property rdf:first is equal to ac:nodeContent.

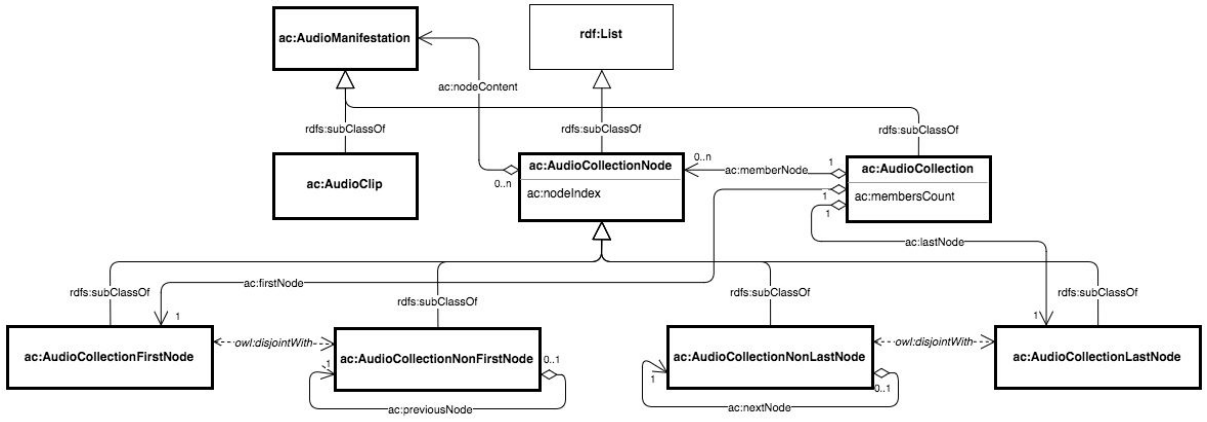

Figure 13. Audio Commons collection as a specialization of standard rdf lists.

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

The instances of ac:AudioCollection are thus formally related to standard RDF lists and can be thus represented as such. The actual conversion between the two different representations can be obtained either through OWL-based reasoning or through an ad-hoc transformation.

Furthermore, we define a property equivalent to ac:firstNode, called ac:collectionAsList. While this would not be strictly needed from a semantic point of view, this can help humans in understanding the meaning when using the rdf:List formalism reducing cognitive load from the perspective of the users of the ontology.

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

4 Integration with Specific Taxonomies

As described in Chapter 3, audio content can be classified by using instances of the ac:AudioCategory class or one of its subclasses (for instance, mo:Genre or mo:Instrument). Sets of these instances can be organized in taxonomies using the SKOS vocabulary.

To demonstrate how taxonomies may be represented and used in conjunction to each other in a relationship, two specific taxonomies are used here as specific examples:

- the AudioSet ontology by Google [AudioSet];
- the Music Instruments Taxonomy (used in MusicBrainz) by Ivan Herman [MIT].

The AudioSet ontology is a hierarchy for classifying a wide range of sounds from human and animal vocalisations, to natural and environmental sounds, to musical and miscellaneous sounds. It is thus a taxonomy of many different sound categories, somewhat improperly called an ontology (at least considering the established meaning that the term has in computer science).

It is defined simply as a tree of labeled concepts, that is available as JSON. Figure 14 below shows the first two levels of the hierarchy. The code snippet in Listing 1 is an excerpt of the the JSON file.

```
[
  {
    "id": "/m/068hy",
    "name": "Domestic animals, pets",
    "description": "Sounds of animals kept in close proximity to humans for ...",
    "citation_uri": "http://en.wikipedia.org/wiki/Pet",
    "positive_examples": ["youtu.be/3SQwwK4JU2I?start=50&end=60"],
    "child_ids": ["/m/0bt9lr", "/m/01yrx"],
    "restrictions": []
  },
  {
    "id": "/m/05zppz",
    "name": "Male speech, man speaking",
    "description": "Speech uttered by an adult male human.",
    "citation_uri": "",
    "positive_examples": ["youtu.be/6niRPYpLOpQ?start=30&end=40",
                          "youtu.be/fPgUDHwQOmM?start=320&end=330"],
    "child_ids": [],
    "restrictions": []
  },
  ...
]
```

Listing 1. An excerpt of the JSON representation of the AudioSet Ontology

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

Each concept has a unique identifier (*id*), a label (*name*) and a short *description*. It may have an associated external web page describing the concept (*citation_uri*), some example youtube video-clips with this kind of sound (*positive_examples*), and some narrower concepts (*child_ids*).

Finally the value of the key restrictions is an array that can include one of the two following values:

- "`abstract`", for concepts that are there to organize the hierarchy but not meant to classify directly a sound;
- "`blacklist`", a concept whose use for tagging has been temporarily disabled.

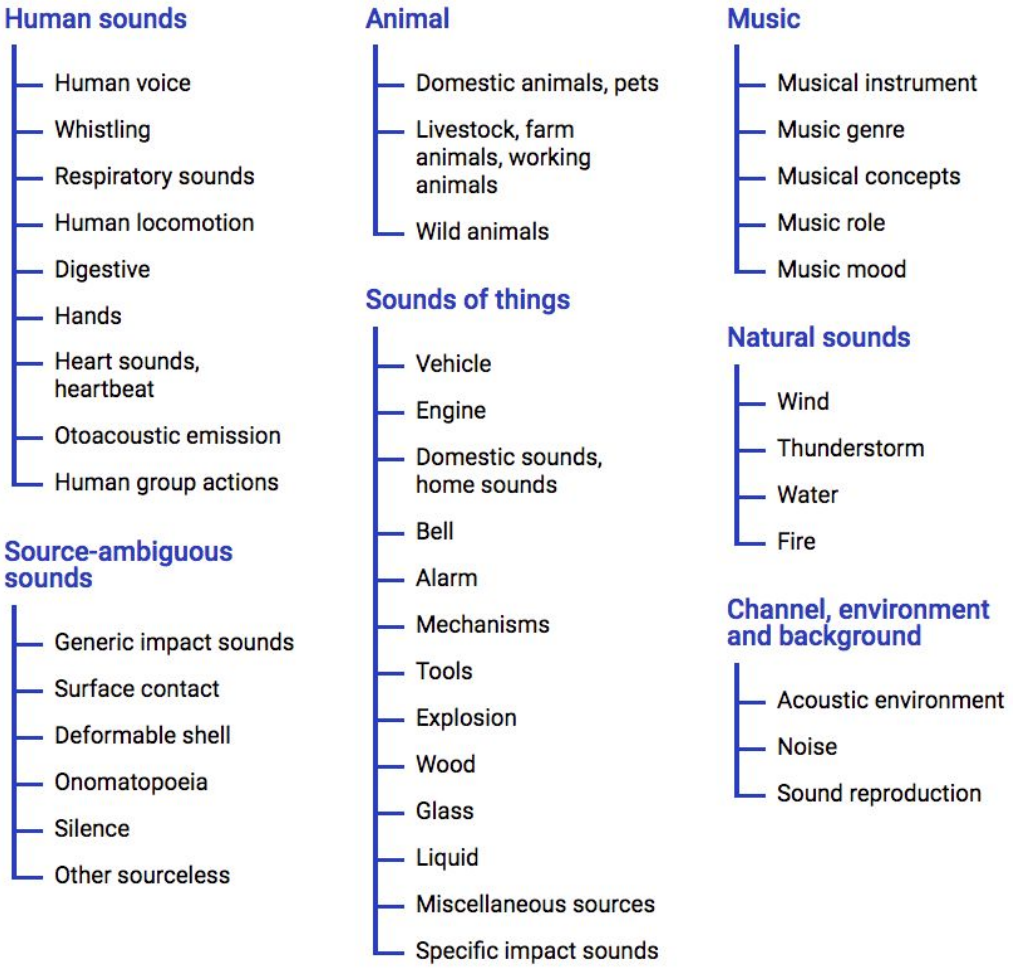

Figure 14. The first two levels (from the top) of the AudioSet Ontology

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

This hierarchy can be represented as a SKOS taxonomy of ac:AudioCategory instances in order to be used with Audio Commons ontology.

An excerpt of a possible SKOS representation is shown in Listing 2, expressed in RDF using the Turtle syntax.

```
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix ac: <http://audiocommons.org/ns/audiocommons#> .
@prefix as: <http://audioset.audiocommons.org/m/> .

as:schema a skos:ConceptScheme;
          dcterms:title "AudioSet Ontology by Google";
          dcterms:description "The AudioSet ontology is ...";
          skos:hasTopConcept as:m\/0dgw9r, as:t\/dd00098, as:m\/0jbk, as:t\/dd00041, ... .

as:m\/0dgw9r a ac:AudioCategory;
             skos:inScheme as:schema;
             skos:prefLabel "Human sounds"@en;
             dcterms:description "Sounds produced by the human body through the ..."@en;
             skos:narrower as:m\/09l8g, as:m\/01w250, as:m\/09hlz4, as:m\/0bpl036, ... .
```

Listing 2. A representation of the AudioSet taxonomy using RDF (Turtle syntax)

The JSON-LD context in Listing 3 maps the classes in the JSON file to instances to a set of related instances of ac:AudioCategory that represent the AudioSet taxonomy as SKOS.

```
{
    "@base": "http://audioset.audiocommons.org/",
    "id": "@id",
    "foaf": "http://xmlns.com/foaf/0.1/",
    "dcterms": "http://purl.org/dc/terms/",
    "skos": "http://www.w3.org/2004/02/skos/core#",
    "name": "skos:prefLabel",
    "description": "dcterms:description",
    "citation_uri": "skos:definition",
    "positive_examples": "skos:example",
    "child_ids": {"@id": "skos:narrower", "@type": "@id"},
    "restrictions": "skos:scopeNote"
}
```

Listing 3. JSON-LD context that permits to interpret the AudioSet Ontology JSON as RDF

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

Any JSON-LD processor, using this context, can directly interpret the AudioSet JSON as a SKOS taxonomy.

The Music Instruments Taxonomy is already defined as SKOS. In Listing 4 there is an excerpt (again expressed in RDF using the Turtle syntax).

```
@prefix dc: <http://purl.org/dc/elements/1.1/>.
@prefix foaf: <http://xmlns.com/foaf/0.1/>.
@prefix skos: <http://www.w3.org/2004/02/skos/core#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix fo: <http://www.ivan-herman.net/foaf.rdf#>.
@prefix wik: <http://wiki.musicbrainz.org/>.
@prefix mit: <http://purl.org/ontology/mo/mit#>.
@prefix ins: <http://purl.org/ontology/mo/instruments#>.

<> dc:creator fo:me;
   dc:date "2007-04-04";
   a skos:ConceptScheme;
   rdfs:comment
       "This is a SKOS version of the Instrument Taxonomy developed by MusicBrainz",
       "Version 0.1";
   rdfs:seeAlso wik:AdvancedInstrumentTree;
   skos:hasTopConcept mit:Electronic_instruments, mit:Other_instruments,
                      mit:Percussion_instruments, mit:String_instruments,
                      mit:Wind_instruments.

fo:me a foaf:Person;
      foaf:name "Ivan Herman" .

mit:Wind_instruments a skos:Concept;
                     skos:inScheme ins:Musical_instruments;
                     skos:narrower mit:Brass, mit:Organ, mit:Woodwind;
                     skos:prefLabel "Wind instruments".
```

Listing 4. An excerpt of the Music Instruments Taxonomy

The taxonomy as it is, can be directly used to classify sounds with mechanism provided by the Audio Commons ontology. The added information that the concepts in this taxonomy are of type ac:AudioCategory (and specifically of type mo:Instrument) can be inferred by adding an external set of OWL assertions.

Different applications/users may use different taxonomies to classify audio content. In some cases multiple taxonomies can to be integrated for an added value. For example, a user or a developer may be interested in using multiple taxonomies at once for wider coverage. Additionally, a user may prefer to use a taxonomy different to the one that was used to classify the sounds. In both cases, a useful consideration is that multiple coexisting taxonomical organisations are supported by the AudioCommons Ontology and these may be used to support a broader range of domain specific

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

queries than what would have been possible with a strong commitment to a selected or designated taxonomy.

As an example, the AudioSet taxonomy has a component classifying musical sounds by instruments. This component can be integrated with the Music Instrument Taxonomy.

Some concepts have a one-to-one equivalent while others, as expected in different schemes, are related in more complex ways.

For example, the mit:Percussion_instruments concept in the Music Instrument Taxonomy includes the instruments that in the AudioSet taxonomy are classified under three separate concepts, Percussion (as:m\/0l14md), Bell (as:m\/0395lw), and Singing bowl (as:m\/0l14t7). Furthermore the AudioSet Percussion concept is broader than mit:Drums, mit:Other_percussion, and all the subconcepts of mit:Tuned_percussion but mit:Bells and mit:Singing_bowl. Conversely, the concept of Singing bowl in the two taxonomies (mit:Singing_bowl and as:m\/0l14t7) has direct correspondence and an equivalence may be defined.

Listing 5 describes formally the above mentioned example relationships among concepts in different taxonomies (RDF in Turtle syntax, again).

```
@prefix skos: <http://www.w3.org/2004/02/skos/core#>.
@prefix mit: <http://purl.org/ontology/mo/mit#>.
@prefix as: <http://audioset.audiocommons.org/m/>.

mit:Percussion_instrument skos:broadMatch as:m\/0l14md, as:m\/0395lw, as:m\/0l14t7.

as:m\/0l14md skos:broadMatch mit:Drums, mit:Other_percussion,
                             mit:Chimes, mit:Gongs, mit:Mbira, mit:Steelpan,
                             mit:Timbales, mit:Triangle, mit:Whistle, mit:Xylophone.

mit:Singing_bowl skos:exactMatch as:m\/0l14t7.
```

Listing 5. An example of SKOS-based coordination among multiple taxonomies

This kind of information can benefit a user interface which aims to integrate multiple taxonomies seamlessly.

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

5 Ontology Use in the Audio Commons API

The main application of the Audio Commons ontology in our ecosystem is to provide a common way to represent multiple data models and APIs and to support the Audio Commons API in terms of formal definition of functions and response formats compared to what API specifications normally provide This pushes the boundaries of communication and interoperability between stakeholders in the audio domain.

This chapter discuss how the ontology may be used in the Audio Commons API through an example call to the Audio Commons mediator.

The Audio Commons API is described in D 2.4. Here we briefly recap some elements in order to better understand how the ontology presented here integrates with the API.

The Audio Commons API defines different endpoints that may be accessed. The main endpoint is the *search* endpoint (at `https://m.audiocommons.org/api/v1/search/text/`) that offers search functionality on audio content that may be in any of the integrated services (currently Europeana, Jamendo, Freesound).

The service implements and asynchronous messaging protocol, i.e., it does not directly return the results as part of the HTTP response. It instead reply with a URL that can be used by the client to poll a resource, called *aggregated response*, this response will contain the results as soon as they arrive from the different services. The use of the Audio Commons API may be exemplified by showing how an aggregated response may be represented.

Listing 6 is an example of part of an aggregated response using the basic, plain JSON, representation. This represent the output of the search on one of the integrated services, in this case Freesound.


```
{
      "num_results": 2,
      "results": [
        {
            "id": "Freesound:199261",
            "url": "https://freesound.org/people/felix.blume/sounds/199261/",
            "name": "A dog barking (close recording)",
            "author": "felix.blume",
            "license": "CC0",
            "preview_url": "https://freesound.org/data/previews/199/199261_1661766-hq.ogg"
        },
        {
            "id": "Freesound:137981",
            "url": "https://freesound.org/people/felix.blume/sounds/137981/",
            "name": "A dog groaning, moaning, sniffing, panting and grunting.",
```


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

```
            "author": "felix.blume",
            "license": "CC0",
            "preview_url": "https://freesound.org/data/previews/137/137981_1661766-hq.ogg"
          }
        ]
}
```

Listing 6. An example of part of an aggregated response from the API mediator.

Listing 7 shows an RDF representation (JSON-LD syntax) of this response using the Audio Commons ontology.

```
{
  "@context": {
    "ac": "http://audiocommons.org/ns/audiocommons#",
    "dc": "http://purl.org/dc/terms/",
    "cc": "https://creativecommons.org/ns#",
    "ccLicenses": "https://creativecommons.org/licenses/by-nc/",
    "fs": "http://audiocommons.org/ns/freesound/",
    "fsSound": "http://audiocommons.org/ns/freesound/sounds/",
    "fsArtist": "http://audiocommons.org/ns/freesound/artists/"
  },
  "@type": "ac:AudioCollection",
  "ac:membersCount": 2,
  "ac:collectionAsList": { "@list": [
    {
      "@id": "fsSound:199261",
      "@type": "ac:AudioClip",
      "ac:homepage": {"@id": "https://freesound.org/people/felix.blume/sounds/199261/"},
      "dc:title": "A dog barking (close recording)",
      "ac:compiled": {"@id": "fsArtist:felix.blume"},
      "cc:license": {"@id": "ccLicenses:CC0"},
      "ac:preview": {"@id": "https://freesound.org/data/previews/199/199261_1661766-hq.ogg"}
    },
    {
      "@id": "fsSound:137981",
      "@type": "ac:AudioClip",
      "ac:homepage": {"@id": "https://freesound.org/people/felix.blume/sounds/137981/"},
      "dc:title": "A dog groaning, moaning, sniffing, panting and grunting.",
      "ac:compiled": {"@id": "fsArtist:felix.blume"},
      "cc:license": {"@id": "ccLicenses:CC0"},
      "ac:preview": {"@id": "https://freesound.org/data/previews/137/137981_1661766-hq.ogg"}
    }
  ] }
}
```

Listing 7. JSON-LD version of the API response, using the presented ontology

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

The list of results is represented as an ac:AudioCollection, serialized as an rdf:List (utilising the @list construct of JSON-LD). In this way the serialization is still concise.

In the previous snippet, the @context is lightweight, and it is used just to assign the prefixes. Having a richer @context would simplify the data part, bringing it close to the original one, as shown in Listing 8. Please note that practical implementation of this response is future work commencing as the ontology is finalised with this deliverable.

```
{
  "@context": {
    "ac": "http://audiocommons.org/ns/audiocommons#",
    "dc": "http://purl.org/dc/terms/",
    "cc": "https://creativecommons.org/ns#",
    "ccLicenses": "https://creativecommons.org/licenses/by-nc/",
    "fs": "http://audiocommons.org/ns/freesound/",
    "fsSound": "http://audiocommons.org/ns/freesound/sounds/",
    "fsArtist": "http://audiocommons.org/ns/freesound/artists/",
    "membersCount": {"@id": "ac:membersCount", "@type": "xsd:int"},
    "collectionAsList": {"@id": "ac:collectionAsList", "@container": "@list"},
    "homepage": {"@id": "ac:homepage", "@type": "@id"},
    "title": "dc:title",
    "compiler": { "@reverse": "ac:compiled" },
    "license": {"@id": "cc:license", "@type": "@id"},
    "preview": {"@id": "ac:preview", "@type": "@id"}
 },
  "@type": "ac:AudioCollection",
  "membersCount": 2,
  "collectionAsList": [
    {
      "@id": "fsSound:199261",
      "@type": "ac:AudioClip",
      "homepage": "https://freesound.org/people/felix.blume/sounds/199261/",
      "title": "A dog barking (close recording)",
      "compiled": "fsArtist:felix.blume",
      "license": "ccLicenses:CC0",
      "preview": "https://freesound.org/data/previews/199/199261_1661766-hq.ogg"
    },
    {
      "@id": "fsSound:137981",
      "@type": "ac:AudioClip",
      "homepage": "https://freesound.org/people/felix.blume/sounds/137981/",
      "title": "A dog groaning, moaning, sniffing, panting and grunting.",
      "compiled": "fsArtist:felix.blume",
      "license": "ccLicenses:CC0",
      "preview": "https://freesound.org/data/previews/137/137981_1661766-hq.ogg"
    }
  ]
}
```

Listing 8. JSON-LD more concise version of the API response, obtained by using a richer @context

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

The example given here deals with the representation of the data. Guiding the implementation is the main goal of this deliverable. The full semantic description of the web API (semantic description of Web services, requests, etc.) will need some added concepts that will considered in the design and development of the future versions of the Audio Commons API.

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

6 Conclusion

In this deliverable, we presented the first complete version of the Audio Commons ontology. The Audio Commons ontology gathers different audio and music related ontologies and defines relationships and classes that tie them together in a way that maximises their utility in the domains of interest within Audio Commons. This covers many categories of sounds and their potentially disjoint taxonomical organisations that may exist in the ecosystem. One example is the Music ontology, which provides a comprehensive yet easy to use and easily extended domain specific knowledge representation for describing music related information.

Ontologies are shared vocabularies and they should grow and evolve through time, depending of the artefacts and services that will be involved in Audio Commons ecosystem. This deliverable presents the Audio Commons ontology and how it can be used in practice. The actual use of the ontology will provide useful feedback that will possibly lead to the evolution of the Audio Commons ontology depending on future application requirements and practical use cases scenarios emerging in the applications deployed in Audio Commons.

The Audio Commons ontology generalises concepts from existing music and sound data models. Furthermore, it provides a common and flexible model to provide multiple classification of sounds, represented as SKOS taxonomies. It finally describes concepts needed to express and formalise the interaction with the Audio Commons API. This pushes the boundaries in communication and interoperability between stakeholders in the audio domain. More generally, the publication of this ontology and its integration with our API has the potential to yield a step change in how REST-ful APIs are defined and linked to ontologies, bringing together the convenience and ease of use of REST APIs and JSON serialisation, with the versatile and knowledge rich representation formats of the Semantic Web such as RDF, OWL ontologies, Graph-based data structures and JSON-LD.

The focus of this document was to explain prominent modelling constructs in the Audio Commons ontology, provide justifications for the most important modelling decisions and guide subsequent implementations relying on the ontology. This deliverable complements the formal specification published online[3].

[3] https://github.com/AudioCommons/ac-ontology

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

7 References

[AudioSet] Audio Set - https://research.google.com/audioset/

[CC-REL] Creative Commons Rights Expression Language - https://creativecommons.org/ns

[CreativeCommons] Creative Commons - https://creativecommons.org/

[Del] Audio Commons deliverables: http://www.audiocommons.org/materials/

[EBUCore] EBU Core ontology - https://www.ebu.ch/metadata/ontologies/ebucore/

[Europeana] Europeana Data Model - http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation

[Fazekas] G. Fazekas. "Knowledge representation issues in audio related metadata model design", 133rd Audio Engineering Society Convention 2012, AES 2012

[FRBR] B. Tillett. "What is FRBR? A conceptual model for the bibliographic universe", The Australian Library Journal Vol. 54 , Iss. 1,2005

[MIT] Music Instrument Taxonomy - http://wiki.musicbrainz.org/AdvancedInstrumentTree

[MUTO] Modular Unified Tagging Ontology - http://muto.socialtagging.org/core/v1.html

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

[SKOS] SKOS Simple Knowledge Organization System Reference - W3C Recommendation 18 August 2009 - http://www.w3.org/TR/2009/REC-skos-reference-20090818/

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