AudioCommons

H2020-ICT-2015, GA 688382

Table of Contents

Section 1: Excellence

1.1 Objectives

The aim of this project is to bring Creative Commons licensed audio content to the creative industries, enabling its creation, access, retrieval and reuse in innovative ways. The digital value chain that connects content creators and content users follows a traditional model that has not yet been adapted to the modern communication society. The democratisation of multimedia content creation and the social web revolution have changed the way in which multimedia content is created, shared and (re)used all over the world. Online multimedia sharing has experienced a huge growth in the last decade, yielding significant amounts of user-generated multimedia resources, a big part of them shared under open licenses. Furthermore, significant amounts of professionally created multimedia content released with closed licenses, are now becoming available as their licenses expire, but nevertheless remain unused. At the same time, media industries (content users) need to reduce production costs in order to remain competitive. There is, therefore, an opportunity for media industries to incorporate such content in their productions, but there is a lack of technologies for easily accessing and incorporating that type of content in their creative workflows. In the particular case of sound and music, a huge amount of audio material like sound samples, soundscapes and music pieces, is available and released under Creative Commons licenses, both coming from amateur and professional content creators. We refer to this content as the *Audio Commons*. However, there exist no practical ways in which Audio Commons can be embedded in production workflows of the content users in the creative industries, and licensing issues are not easily handled across the production chain. As a result, most of this content remains unused in professional environments. Closing the gap between content users' needs and the available Audio Commons generated by content creators would hence be beneficial to both parties.

In this project (AudioCommons), we will bring the Audio Commons content to the creative industries. We will enable access and retrieval of audio content in innovative ways that fit the requirements of different use cases of media creation, benefiting content users by reducing production costs and benefiting content creators by exposing their works to professional environments and allowing them to licence their content. The Audio Commons technology layer that will be developed in this project fills in all the necessary gaps that are needed to make the **Audio Commons Ecosystem** (ACE) possible, from metadata specification to content annotation, content analysis, gathering of user feedback and licensing procedures. The Audio Commons Ecosystem will allow *content creators* to publish their content through *content providers*, that will host it and make it available to *content users* using embeddable tools developed by *tool developers*. The Audio Commons Ecosystem proposes a shift in current business models in which content creators will easily interact with content users, proposing an open and *circular model* through which the communication between both parties will be simplified and everyone will be able to contribute to that ecosystem with valuable audio content and further information about that content.

Besides making Audio Commons content available, the project will focus on all the legal and business-related challenges posed by the core ideas of AudioCommons. We will provide standard procedures for content creators to publish their content as Audio Commons through content providers. Audio Commons content will be able to accurately incorporate information about licenses so that it can be displayed to content creators and content users in understandable and meaningful ways, putting a special emphasis on simplifying the (re)licensing process for those pieces of content that require it (e.g., from CC-BY-NC to CC-BY). Finally, the project will investigate about the long-term sustainability of the proposed model, and propose innovative business models that can arise from the concept of Audio Commons.

The AudioCommons consortium is formed by leading research institutes in sound and music computing and key players in the creative industries. Academic partners include the Music Technology Group from

Universitat Pompeu Fabra (MTG-UPF, project coordinator), the Centre for Digital Music from Queen Mary University of London (QMUL), and the Centre for Vision, Speech and Signal Processing (CVSSP), the Institute of Sound Recording (IoSR) and the Business School from University of Surrey (denoted as 'Surrey' in this proposal). On the industry side, the Audio Commons consortium is formed by Jamendo (one of the biggest platforms for sharing independent Creative Commons music, with more than 470,000 music pieces from 35,000 different artists), AudioGaming (an innovative company developing next generation audio tools for sound designers and video game developers), and Waves (world's leading developer of audio DSP solutions for the professional, broadcast, and consumer electronics audio markets). The key objectives of AudioCommons can be summarised as follows:

Objective 1	**Enable the retrieval of audio content in innovative ways by bringing current state-of-the-art methods of semantic sound and music description to higher standards, and make these methods available as software packages.** To achieve this, we will ensure that the content is represented in meaningful ways and annotated using advanced sound and music annotation systems.
Objective 2	**Develop and deploy the technological layer that will allow the interconnection of all stakeholders that participate in the ACE.** This will enable the communication of content creators, content providers, content users, tool developers and further potential actors, for publishing, licensing and consuming Audio Commons content through unified and standardized procedures.
Objective 3	**Create and set up the Audio Commons Ecosystem, publish Creative Commons audio content through it and build tools that can consume the content and be embedded in existing creative workflows.** The ACE will be initially populated with 400,000+ music pieces and 230,000+ sound samples coming from two content providers (members of the AudioCommons consortium). We will also build tools for consuming that content that will be directly embeddable in the existing production workflows of creative industries (e.g., plug-ins or add-ons for existing software).
Objective 4	**Define standard procedures for new stakeholders to join and participate in the ACE, fostering its future growth and sustainability.** These procedures will define how to produce and consume content within the ACE, and will ensure its existence after the lifetime of the current project. The ACE will provide an example model for content reuse in the audio domain that could be extrapolated to other fields and content types.

1.2 Relation to the work programme

AudioCommons is highly aligned with the specific challenges of the H2020 ICT-19 call. What follows is a series of quotations from the call with brief explanations on how AudioCommons fits in the particular challenges:

"developments related to content creation, access, retrieval and interaction offer a number of opportunities and challenges, also for the creative and media industries"
The Audio Commons vision is particularly focused on changing the way in which creative industries create, access and retrieve audio content. The Audio Commons Ecosystem will establish a new channel through

which content users can take advantage of Creative Commons audio content, and content creators can make their content easily accessible in the professional market and for the public in general.

"those industries need to explore new ways of creating and accessing content"

Besides the new communication channels between content users and content creators that the Audio Commons technology will provide, content published in the Audio Commons Ecosystem will be browsable and accessible in innovative ways. The use of semantic sound and music annotations will enable unprecedented access to content repositories. Moreover, these repositories will be continuously updated and enhanced with further contributions from content creators.

"The opportunity to establish new forms of content and user engagement could be transformative to many businesses in creative and media industries."

In the core of the Audio Commons vision there is the idea of the *circular model*, in which content users can easily become content creators, feeding back into the Audio Commons Ecosystem (as allowed by the use of Creative Commons licenses). The tools used to access Audio Commons content will also enable the reintroduction of derivatives and new content to the system, establishing in this way new forms of user engagement and participation in the ecosystem.

"The focus is on research, development and exploitation of new or emerging technologies (e.g. 3D and augmented reality technologies) for digital content creation to support the creative and media industries and for unlocking complex information and media and interacting with them"

An important part of the research challenges of AudioCommons are focused on the semantic description of sound and music content. Semantic technologies allow expressive and rich representation of content, but have not yet been intensively applied into the sound and music domain, particularly for the description of content such as sound effects, soundscapes or musical samples. We will use semantic technologies in the real-world scenario of creative and media industries, impacting in the way these industries interact with the content.

"Research in new technologies and tools to support creative industries in the creative process from idea conception to production"

The Audio Commons Ecosystem will provide access to huge amounts of sound and music content, browsable in innovative ways, and easily embeddable in existing creative workflows. This will have a clear impact on the whole creative process, allowing content users to to access information from evolving repositories using query definitions based on high-level semantic terms and other facets specific to the audio domain.

"The proposed tools should explore the potential of technology to enhance the human creative process from the expression of ideas to experiment solutions"

The Audio Commons Ecosystem will enable fast experimentation of new ideas by allowing immediate retrieval of audio content and later licensing if required (as opposed to the current paradigm in which users first license the content and then use it). Because the content will be released under Creative Commons licenses, content users can potentially reuse, transform and even re-publish the content, all within their existing production workflows.

"Where possible, collaboration and user-community interaction should be improved based on research leading to a deeper understanding of the dynamics of co-creative processes"

The circular model proposed by the Audio Commons Ecosystem, which enables seamless communication between content users and content creators and easy publishing of new content to be feed back to the ecosystem, will establish an Audio Commons community, formed by all individual users, industries and other entities that participate in the ecosystem, and that can reuse their material. This will effectively enable the emergence of co-creation processes, and generate metadata describing the "lifetime" of Audio Commons content that the scientific community will be able to analyse and learn from.

"The tools should be cost effective, intuitive, and be demonstrated in real-life environments relevant for the creative industries (such as advertising, architecture, arts, design, fashion, films, music, publishing, video games, TV and radio)"

The Audio Commons technology will be built on top of standard Internet technologies, and will consist in a number of software packages to represent, annotate and retrieve Audio Commons content. Hence, the tools will be cost effective, a significant number of them released under open source licenses. The tools will be demonstrated in different real-world use cases including video games production, music production, sound design and audio-visual production.

1.3 Concept and approach

Useful terminology and definitions

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

1.3.1 Concept

AudioCommons aims to **build an ecosystem of content, technologies and tools that change the way in which content users and content creators interact with audio resources**. The scope of AudioCommons poses both technological and non-technological challenges, including sound and music computing, rights management and business models topics and their application in different (yet related) use cases. The excellence and transdisciplinary nature of the members of the consortium guarantees however the ability of approaching every challenge from a privileged perspective. The key ideas underneath the Audio Commons vision are described in the following bullet points. Figure 1.1 exemplifies the aforementioned ideas and concepts of the Audio Commons Ecosystem, and shows graphically how the different actors of the ecosystem will be interconnected.

- **Creative Commons audio content:** The audio content available in the ACE will be released under Creative Commons licenses. Creative Commons licenses allow understanding audio sharing in a much broader way, in which the possibilities for content reuse are not arbitrarily restricted according to traditional business models. In the Internet era, content sharing should be easy and should be done under clear and understandable terms. The Creative Commons modular licenses perfectly fit our vision because they **provide a clear framework in which content creators can specify the possibilities for reusing their published content**. Furthermore, there already exist large amounts of sound and music content released under Creative Commons licenses, as exemplified by sites like Freesound and Jamendo[1] (both run by members of the AudioCommons consortium), and other sites like the Internet Archive, Soundcloud, ccMixter and significant amounts of the content aggregated by Europeana[2]. Within this project, we want to bring the sound and music Creative Commons content to the industry plane, providing standard procedures for putting in contact content creators and content users and, when required, (re)license the content for commercial use.

- **Circular interaction model and co-creation:** The same "open" sharing concepts represented by the Creative Commons licenses are applied to the whole design of the ACE. The ecosystem we propose allows content users to consume reusable content published by content creators and hosted by content providers. However, it **also allows content users to easily become content creators and feed the ACE back with new material or derivatives of existing material** (see feedback arrows in Figure 1.1). In this way, the Audio Commons Ecosystem also fosters co-creation of audio resources, through a model in which everyone can seamlessly contribute and benefit. The circular model will be supported by the tools and technologies of the ACE, allowing to both consume and contribute within the same environment.

- **Audio content enhanced with annotations and metadata:** In the core of the Audio Commons vision, it lays the idea that audio content is much more that just audio files. The Audio Commons content will potentially come from very different sources and, particularly in the case of user-generated content, will typically lack a uniform structure to allow proper access and retrieval. AudioCommons targets different kinds of audio content such as **music pieces** (i.e., audio recordings corresponding to complete songs), **music samples** (i.e., individual music elements such as single

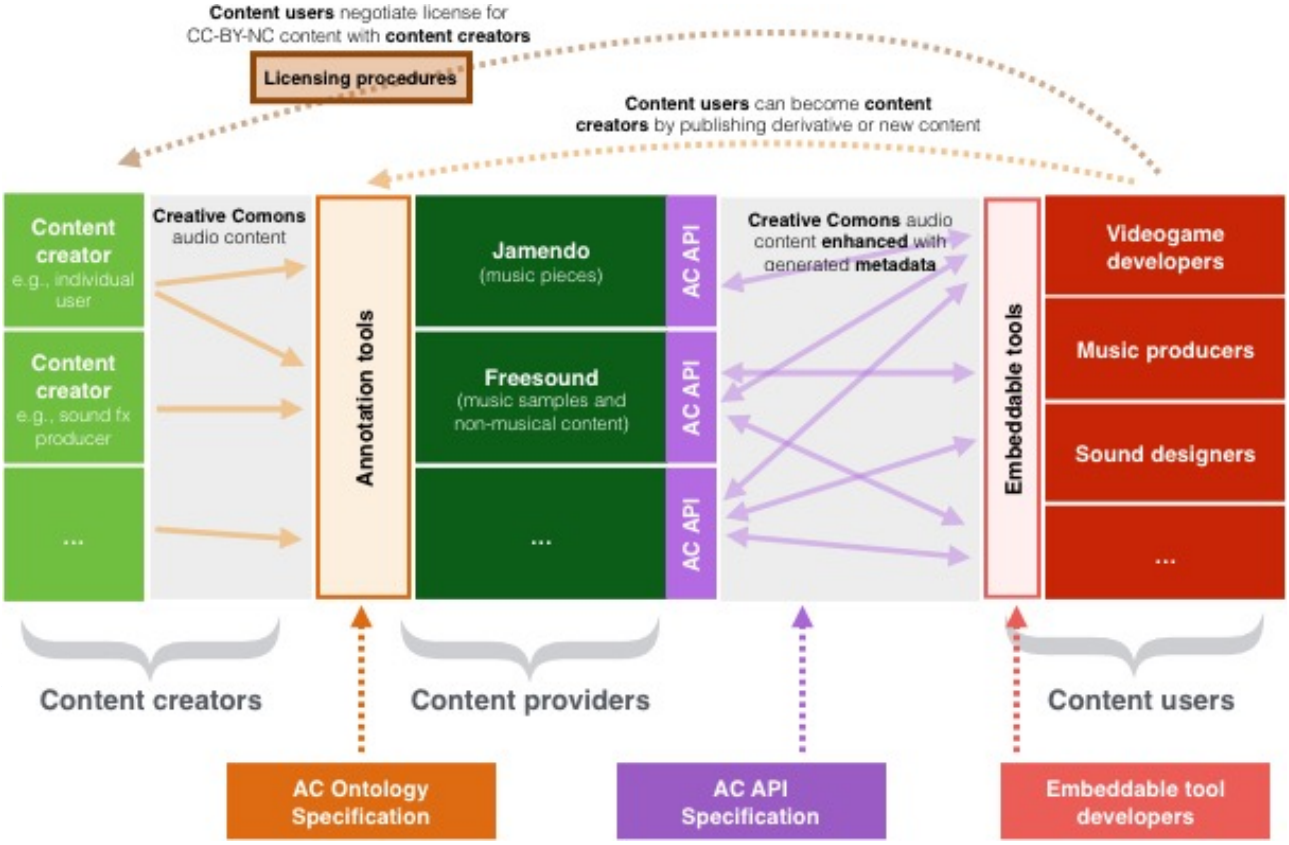

Figure 1.1: Audio Commons Ecosystem Conceptual Diagram

notes, percussive hits, chords, melodies and loops), **sound effects** (i.e., non-musical sound events such as foley, footsteps, opening and closing doors, alarm sounds, cars passing by, animals and all kinds of noises or artificially created glitches) and **soundscapes** (i.e., complex non-musical recordings such as environmental recordings, street ambiances or artificially constructed sonic environments). We propose a series of research developments that will set a new standard in state-of-the-art **sound and music description and semantic representation technologies**. The combination of semantic representation of audio content with the use of an ontology (i.e., the **Audio Commons Ontology** or AC Ontology) with the annotation tools (both manual and automatic) that will be developed within the project, will enable searching and browsing Creative Commons audio content in unprecedented and innovative ways by, for example, browsing sound effects on the basis of their timbral characteristics. The tools that we will develop for annotating content will, in essence, *enhance* the content by providing extra metadata with a uniform representation compatible with the ACE.

- **Distributed framework:** The Audio Commons Ecosystem is designed for incorporating many different agents that will be able to connect to the ecosystem for publishing, consuming, licensing or working content. For this reason, we propose a distributed architecture in which an indeterminate number content providers process and host resources created by the content creators, and offer API endpoints which follow a common API specification (i.e., the **Audio Commons API specification** or AC API). The tools embedded in production workflows will then access the different content providers by implementing a single API specification. During the lifetime of this project, we will evaluate the ecosystem with **Freesound and Jamendo as the initial content providers** and "early adopters" of the technology, and release embeddable tools that will allow the access to that content within production environments. The first embeddable tools for the ACE will be developed by the industry partners of the consortium, and will include **two audio plugins** (developed by Waves and AudioGaming), an **add-on for a well-known open source Digital Audio Workstation** (also developed by Waves), and a **web interface** for accessing music pieces of the Audio Commons Ecosystem (developed by Jamendo). Nevertheless, an important part of the work carried out in AudioCommons will be the publication of **clear and standard procedures** describing to potential new actors of the ACE how to create new embeddable tools for accessing (and publishing) content, how to become a content provider of Audio Commons, or how to interact with the ACE in any of the other ways enabled by our model (e.g., licensing content, providing metadata, etc.).

1.3.2 Technology Readiness Levels of the Audio Commons Ecosystem

AudioCommons developments, as a Research and Innovation action, are positioned in technology readiness levels from TRL1 to TRL4 (i.e. from "basic principles observed" to "technology validated in lab"). As mentioned, during this project we will deploy an Audio Commons Ecosystem populated with audio content from two content providers (Freesound and Jamendo), initially including three embeddable tools developed by the industrial partners of the consortium, and with clear guidelines for other potential actors to join the ecosystem (e.g., for potential new tool developers, content providers, etc.). The Audio Commons technology and tools will be validated in the lab, informed by real-world needs from industrial partners and AudioCommons associated experts.

The Audio Commons Ecosystem will take advantage of some existing technologies that are already well established in the market, but will also need to develop some others that require further advancements to be usable in operational environments. On the one hand, the Audio Commons Ontology will rely on **Semantic Web** technologies that have already been proven in the real-world (TRL6-9). Moreover, the Audio Commons API will be defined as a web-based **RESTful API** with which any regular HTTP client will be able to connect. RESTful APIs are widely used in the Internet and have been proven to work well even with huge amounts of requests (TRL9). Similarly, the concept of content providers of the Audio Commons

Ecosystem is already well represented in the real-world not only by Freesound and Jamendo but also by many other sites such as the ones above mentioned Internet archive, Soundcloud, ccMixter and Europeana. These sites have been proven in operational environments (TRL9), and also offer access to their content via RESTful APIs. The embeddable tools developed by the non-academic partners of the consortium will follow **well-defined industry standards** for their implementation (e.g., audio units technology), which have also been proven in operational environments (TRL9).

On the other hand, the core of the Audio Commons Ecosystem includes the use of computational methods for **the semantic annotation of audio content** that will need to be developed and advanced during this project. The technology readiness level of existing methods for the semantic annotation of audio content (typically referred to as high-level audio feature extraction [Serra13]) varies a lot depending on the methods. In general, methods for the automatic characterisation of music pieces are the most advanced (e.g., music fingerprinting methods are proven and used in operational environments, TRL9), whereas methods for the characterisations of other content such as **sound effects or music samples are less developed** and generally only proven in labs (TRL1-4). Similarly, there are other ACE aspects such as the **service discovery using semantic web technologies** whose technology readiness level is below TRL4. Our research and innovation action will particularly contribute in advancing the technology readiness level of these components (see Section 1.4.1).

1.3.3 Related research and innovation activities

MTG-UPF, QMUL and Surrey leverage research technologies and knowledge from previous European projects in research, innovation and FET projects, as well as national research projects. The following table gives an overview of projects or initiatives in which the academic partners of the consortium have been (or are) involved and which have a close relation with AudioCommons. Extended summaries of these projects can be found in Section 4 of this proposal.

MTG-UPF
Freesound Freesound is an online collaborative sound database where people from different disciplines share recorded sound clips under Creative Commons licenses. It currently hosts more than 230,000 sounds and has more than 4 million registered users (as of February 2015). Freesound was started in 2005, and is still developed and maintained at MTG-UPF. Freesound offers a RESTful API through which researchers and developers can access its content programmatically. See section 1.4.2 for more details about Freesound. **CompMusic** (ERC Advanced Grant 267583) CompMusic is a research project funded by the European Research Council and coordinated by Xavier Serra from MTG-UPF. It aims to advance in the automatic description of music by emphasizing cultural specificity. It carries research within the field of music information processing with a domain knowledge approach, taking advantage of semantic technologies and advanced signal processing techniques. **MIReS (**FP7 ICT-2011.1.5 Project No. 287711) MIReS project aims to create a research roadmap of Music Information Retrieval (MIR) field to provide a meta-analysis of the MIR discipline, address emergent contexts and major challenges, formulate research evaluation standards for the discipline, contribute to the establishment of music production and digital library management standards, engage a variety of stakeholders from different disciplines of academia and industry and deliver innovative platforms for co-creative workshops focusing on horizon-scanning and technology foresight. **Metaverse1** (ITEA2 07016)

The Metaverse1 Project, "Setting Global Standards among Real and Virtual Worlds", started in 2008 and finished in 2011. It was an European ITEA2 project with the participation of industrial partners (Philips, Alcatel-Lucent, OrangeLabs, Technicolor, other) and research institutions (MTG-UPF, INRIA, Utrecht University…). MTG-UPF's participation focused in the area of analysis and synthesis of environmental sounds, developing a real-time engine for the autonomous generation of soundscapes and an automatic classification for environmental sounds.

SIMAC (FP6-507142 IST-2.3.1.7)
SIMAC was a project coordinated by the MTG-UPF between 2004 and 2006. SIMAC's main task was the development of prototypes for the automatic generation of semantic audio descriptors and the development of prototypes for exploration, recommendation, and retrieval of music. One special feature was the development and use of semantic descriptors. That is, ways to tag music that are close to the user's way of describing its contents. MTG-UPF developed tools to be used by music consumers, music distributors, and music creators.

QMUL

FAST-IMPACt (EP/L019981/1)
The Fusing Semantic and Audio Technologies for Intelligent Music Production and Consumption project brings the very latest technologies to bear on the complete music industry, end-to-end, producer to consumer. It aims at making the production process more fruitful, the consumption process more engaging, and the delivery and intermediation more automated and robust. The project addresses 3 premises: (i) that Semantic Web technologies should be deployed throughout the content value chain from producer to consumer; (ii) that advanced signal processing should be employed in the content production phases to extract "pure" features of perceptual significance and represent these in standard vocabularies; (iii) that this combination of semantic technologies and content-derived metadata leads to advantages and new products and services at many points in the value chain, from recording studio to end-user (listener) devices and applications.

OMRAS2 (EP/E017614/1)
In "OMRAS2: A Distributed Research Environment for Music Informatics and Computational Musicology" , several computational algorithms were developed for automatic annotation of musical audio, including novel methods for audio transcription, chord recognition, key recognition, tempo and beat detection, structural segmentation and music similarity. Several Semantic Web ontologies were created for describing and publishing music related metadata on the Semantic Web. The Music Ontology, a core framework connecting the OMRAS2 ontologies became a de-facto standard in music-related data publishing.

EASAIER (IST-FP6-033902)
The "Enabling Access to Sound Archives through Integration, Enrichment and Retrieval" was a two and a half year European research project addressing access to and preservation of cultural and scientific resources. An innovative approach to knowledge representation led to development of the Music Ontology and the Audio Features Ontology, which are now widely used outside the consortium. In the field of automatic feature extraction, QMUL developed new high performance methods to identify and characterise sound objects (emotion detection, laughter detection, key extraction, tempo identifier...). In the area of presentation of multimedia material, novelty was presented in sound source separation, equalisation and noise reduction algorithms.

M4 (TS/J002283/1)
Making Musical Mood Metadata (M4) was a collaborative technology transfer project with BBC R&D and I Like Music. Over eighteen months, the team developed new and innovative methods of extracting high-

level metadata from music contact, including information about the mood and emotional content of tracks. Large-scale trials involving feature extraction from over 1M tracks were performed with BBC users using an online platform providing recommendations of sounds-like tracks for audio-visual productions.

The following two projects were collaborations with other partners of the consortium:
MIReS (FP7 ICT-2011.1.5 Project No. 287711) - MTG-UPF
SIMAC (FP6-507142 IST-2.3.1.7) - MTG-UPF
Audio Data Exploration: New Insights and Value (EP/M507088/1) - Surrey (description below)

Surrey

S3A: Future Spatial Audio for an Immersive Listener Experience at Home (EP/L000539/1)
S3A is a major five-year UK research collaboration between internationally leading experts in 3D audio and visual processing, the BBC and UK industry. Its goal is to enable listeners to experience the sense of "being there" at a live event, such as a concert or football match, from the comfort of their living room through delivery of immersive 3D sound to the home. University of Surrey's IoSR role in S3A is to: (i) ascertain which attributes of reproduced spatial audio are most important to listeners; (ii) identify any important attributes missing from previous studies; (iii) determine the relationships between the important attributes and listener preference; (iv) model overall spatial quality in terms of the important perceptual attributes; and (v) model these perceptual attributes in terms of their physical correlates.

Audio Data Exploration: New Insights and Value (EP/M507088/1)
This project, funded by Innovate UK and EPSRC, is a collaboration between Audio Analytic Ltd., QMUL and University of Surrey (CVSSP). The project is undertaking advanced audio data analysis and modelling techniques in Automatic Environmental Sound Recognition, to create value across a variety of applicative domains. As well as current application areas such as Professional Security and Home Security, a range of novel markets can be developed in relation to Multimedia Database Indexing, Environmental and Industrial Monitoring, the Internet of Things and more.

Musical Audio Repurposing using Source Separation (EP/L027119/2)
This EPSRC-funded project aims at developing a new approach to high quality audio repurposing (upmixing and remixing), based on high quality musical audio source separation. To achieve this, University of Surrey's CVSSP will combine new high resolution separation techniques with information such as musical scores, instrument recognition, onset detection, and pitch tracking. In parallel, perceptual evaluation measures for source separation, remixing and upmixing will be investigated, and new diagnostic evaluation techniques tailored to measure different aspects of the repurposed outcome will be developed. The outcomes of this project will allow music consumers to enjoy their favourite songs in interactive remixing apps and games, even where the original separate "stems" are not available.

ENDuRE: European Network of Design for (u)Resilient Entrepreneurship (554337-EPP-1-2014-1-IT-EPPKA2-KA)
This project, in partnership with iVeridis, Blue Ocean Robotics, IDEA, CEDIT and the Universities of Pisa, Surrey (Business School) and Southern Denmark is looking at emergent business models and methodologies for inter-european entrepreneurship. Part of the remit of ENDuRE is to address how new economic models in the digital economy can survive the pressures of market change, and take advantage of cross-subsidisation from other industries, such as music resold/used within games for health.

QESTRAL (EP/D041244/1)
This project, in partnership with Bang & Olufsen, DK, and BBC R&D, UK, developed a set of test signals

and related physical measures that can be used to evaluate the performance of spatial audio processing systems in relation to a range of perceptually important spatial quality attributes of a reference system, and modelled the reproduced sound field generated by selected transducer arrays at a number of monitoring positions within a listening space. This process required the identification of key perceptual attributes and the refinement and expansion of existing metrics for these attributes, to build into a comprehensive sound characterisation system.

Furthermore, there exist other highly related initiatives with which we will establish contact and collaborate when possible. Of particular relevance to our project is the initiative carried out by Europeana and the Europeana Sounds project[3] for the definition of a Sound profiling model (as part of the **Europeana Data Model**) and the aggregation of audio content from different providers. Europeana provides a web API through which HTTP clients can search and browse cultural heritage multimedia contents (including sound and music) aggregated from several public libraries across Europe. The Europeana initiative and AudioCommons are highly compatible in their vision and goals. Whereas Europeana is focused on allowing open access to audio content (and part of the content is released under Creative Commons licenses), Audio Commons is focused on fostering the reuse of audio content in creative industries. Communication with Europeana and Europeana Sounds will be fluent as Surrey's PI Mark Plumbley is on the Advisory Board of the Europeana Sounds project. Therefore, a natural future step for extending the Audio Commons Ecosystem after its development and evaluation during this project, would be the inclusion of Europeana aggregated content in the Audio Commons Ecosystem, as well as the inclusion of Audio Commons content in Europeana API.

1.3.4 Overall approach and methodology

The approach with which we plan to achieve the goals of AudioCommons has several broad steps that we roughly divide into two working areas that will run in parallel. On the one hand, there will be a number of steps aimed at fulfilling Objectives 1-3, and will be focused on the research about sound and music **semantic annotation and representation**, the technological **service orchestration**, and the development and evaluation of **tools for accessing Audio Commons** content (WP2, WP4-6). On the other hand there are a number of steps aimed at fulfilling Objectives 3-4 by working on the **rights management** issues related with the Audio Commons Ecosystem, working on establishing **procedures for new actors to join the ecosystem**, and by exploring **emerging business models** and the future self-sustainability of the ecosystem (WP3,7). What follows is an overview of the different broad steps that will be carried out for accomplishing the objectives of both working areas. The actual work plan and work package descriptions are given in Section 3.1:

Steps related to the technological challenges of AudioCommons (WP2, WP4-6):

- *Gathering of requirements from creative industries*
  Research and technological developments of AudioCommons will be driven by the **requirements of different use cases** of creative industries represented by the industry partners. Industry partners know their users and their needs, and will therefore contribute in the definition of innovative ways of retrieving audio content and selecting some of them as targets for AudioCommons technological developments.

- *Definition of the Audio Commons Ontology and API specification*
  We will define the Audio Commons Ontology capable of representing audio content in ways that allow further retrieval in innovative and meaningful ways. The AC Ontology will be a domain-

[3] http://www.europeanasounds.eu

specific ontology capable of representing **information about audio content in general** (not only about music pieces), and **also about the intellectual property and licensing aspects** of the content. Besides the AC Ontology, we will define an Audio Commons API that will specify the way in which the **different actors and technologies of the Audio Commons Ecosystem should interact** among them. The API will standardise the ways in which embeddable tools will communicate with content providers in order to consume audio content, and will specify how content users will communicate with content creators to perform licensing and attribution when required. Best practices and guideline documents will be released to facilitate potential new actors to implement the API specification and join the ACE.

- *Research and validation*
  Intensive research and validation will be carried out to develop the necessary technological advancements that will make the different tools of the Audio Commons Ecosystem possible. On the one hand, tools for the semantic annotation of audio content will be researched, and the audio content from **Jamendo and Freesound will be enhanced with the generated metadata** and considered to be *Audio Commons Ready*. Content providers will expose their data offering endpoints conforming to the Audio Commons API specification (more than 470,000 music pieces from Jamendo and more than 230,000 sound samples from Freesound). On the other hand, the embeddable tools that will consume Audio Commons content (and by extension the whole ACE) will be developed and evaluated within **different use cases of the creative industries** including video games production, music production, sound design and audio-visual production, represented by the industrial partners of the consortium.

Steps related to rights management challenges, business models and dissemination of AudioCommons (WP3, WP7):

- *Clarification of rights management aspects related with the ACE*
  The Audio Commons Ecosystem will potentially lead to **complex audio content reuse patterns** that will challenge the basic usage of Creative Commons licenses for rights management. For this we will write documentation with different use cases that exemplify possible situations with intellectual property conflicts and provide solutions for them based on the use of these licenses. Besides rights management aspects related to Audio Commons content, we will also describe how these should be extended to the different components and tools in the ecosystem and how are these affected by content licenses.

- *Exploration of emerging business models for the ACE*
  Because the Audio Commons vision challenges some of the established business models in the creative industries, an important part of the project will be focused on evaluating how our proposed ecosystem fits and impacts the current industries, and how the interactions of the different actors in the ACE can **give birth to new business models and ultimately support the existence of the ecosystem itself**.

- *Demonstration and dissemination*
  The developments of AudioCommons will be demonstrated and disseminated to the **scientific community, the creative industries and the public in general**. We will define procedures and write guidelines that will explain to potential new content creators, content providers, tool developers and content users how to connect to the Audio Commons Ecosystem to ensure its subsistence after the lifetime of the current project. We will actively talk with potential future actors for the ACE, explaining the vision and how to get involved.

All different steps to achieve the goals of AudioCommons will be evaluated. The evaluation activities of AudioCommons will assess the project's hypotheses, RTD outputs, ontologies, APIs, business models as well as the system as a whole. We will perform: (i) **analytical evaluation** of tools targeting specific creative domains, feature extraction algorithms, data representations, ontologies and business models; and (ii) **holistic evaluation** of the ACE from the perspective of facilitating creative output and the success of integrating Audio Commons content in existing workflows. The evaluation of scientific and business hypotheses and RTD outcomes will be performed in respective work packages (WP2-5). The evaluation of the ACE as a whole and of the developed tools within their creative domains will be performed in a separate work package (WP6).

1.4 Ambition

1.4.1 Progress beyond state-of-the-art

The ambition of the AudioCommons project is to create an Audio Commons Ecosystem through **advancing the state of the art in the different fields and facets covered by the project**. In the following table we summarise the current state-of-the-art, the challenges that AudioCommons will pose over current state-of-the-art, and the ambitions and approaches that our project will follow to overcome these challenges. We divide the progress beyond state-of-the-art in four facets, which will be tackled in a number of particular tasks of different work packages (see Section 3.1).

A) Access to Creative Commons audio resources within existing creative workflows (WP2, WP6)	
State-of-the-art	• Existing APIs for programmatically accessing content from sites like Jamendo, Freesound, Soundcloud, ccMixter, Europeana, etc. • Limited access due to lack of high quality and unified metadata [Aslam12]. • No unified access mechanism for APIs (APIs have different specifications). • Inadequate retrieval tools • Creative Commons audio content is not frequently used in professional environments as a result of the above points.
Challenges	• Finding and defining metadata requirement in creative applications. • Designing appropriate ontologies for data representation. • Providing reliable metadata to facilitate access to Audio Commons content. • Building access to Audio Commons content into existing creative workflows.
Our ambitions and approaches	• Seamless integration of industry specific tools and Audio Commons content. • Develop an Audio Commons ontology and API specification to store and access Audio Commons content (see C). • Improve content-based tools for the semantic annotation of audio content (see D). • Improve manual tools for the semantic annotation of audio content (see D). • Develop embeddable tools that can search and retrieve Audio Commons content in the advanced ways allowed by the ontology, enabling integration of Audio Commons content into existing creative workflows in use by industry.
B) Licensing procedures for commercial reuse of Creative Commons resources (WP3)	

State-of-the-art	• Existing platforms for licensing Creative Commons music pieces for commercial use (e.g., Jamendo Licensing[4], Audiosocket[5], Magnatune[6]). • Uncertainties about licensing and other legal requirements for using Creative Commons content in commercial productions. • Lack of standardised procedures for (re)licensing content with Creative Commons licenses that allow content reuse in commercial productions. • Uncertainties about business models that can emerge from the publication and consumption of audio data licensed under open Creative Commons licenses.
Challenges	• Clarification of the usage of Creative Commons licenses for audio content in scenarios with high content reuse and presence of derivative works. • Designing standard and easy procedures that allow content creators to (re)license their content to content users with suitable licenses. • Identifying emerging business models that can make the Audio Commons Ecosystem self-sustainable.
Our ambitions and approaches	• Enabling a sustainable ecosystem (ACE) based on open knowledge • Provide guidelines about best-practices for the usage of Creative Commons audio content within the ACE, with examples and use-cases from the point of view of the different actors involved in the ecosystem (content creators/providers/users, tool developers). • Definition of standard flows for the (re)licensing of Creative Commons audio content and specification of these flows into the Audio Commons API, which will allow easy licensing procedures within the embeddable tools. • Research and evaluation of business models that can emerge within the Audio Commons Ecosystem, describing business opportunities that can be carried out by new or existing actors joining the ecosystem.
C) Semantic representation of sound and music (WP2)	
State-of-the-art	• Generic metadata for media ontologies such as the Dublin Core ontology[7] from the Dublin Core Metadata Initiative (DCMI). • Audio-specific ontologies developed at QMUL such as The Music Ontology[8] [Raimond07], the Audio Features Ontology[9], the Studio Ontology [Fazekas11b] and the ongoing work on Europeana Data Model for sounds[10]. • Existing audio-specific ontologies only suited for representing music pieces. • Rights management ontologies in the media domain such as the Media Value Chain Ontology (MVCO), standardised by MPEG [Kudumakis14]. • Established Semantic Web technologies (proposed by W3C) such as Resource Description Framework (RDF) and the OWL Web Ontology Language.
Challenges	• Define an ontology capable of representing a broader scope of audio content than that represented by the already existing Music Ontology (i.e., including non-musical

[4] https://licensing.jamendo.com/en
[5] https://www.audiosocket.com
[6] https://magnatune.com
[7] http://dublincore.org
[8] http://musicontology.com
[9] http://motools.sourceforge.net/doc/audio_features.html
[10] http://pro.europeana.eu/get-involved/europeana-tech/europeanatech-task-forces/edm-profile-for-sound

	sounds and music samples). ● Seamlessly integrate rights management concepts in the aforementioned ontology, with enough expressive power to represent complex licensing patterns that will potentially arise within the ACE.
Our ambitions and approaches	● Definition of the Audio Commons Ontology, which will enable the semantic representation music pieces (based on the Music Ontology) and also of music samples and non-musical content such as sound effects and soundscapes. The Audio Commons Ontology will also be capable of representing complex licensing information for the annotated content. ● Turn the Audio Commons Ontology into a de-facto standard, adopted at first by the initial content providers of the Audio Commons Ecosystem (Freesound and Jamendo) and then by other sites hosting audio content. ● As a long-term ambition, create the Audio Commons Foundation which will maintain the Audio Commons Ontology and API specification.
D) Semantic annotation of sound and music (WP4, WP5)	
State-of-the-art	● Common audio feature extraction algorithms [Peeters03]. ● Methods for mid/high-level characterisation/annotation of music signals, relevant to AudioCommons: ○ Chord and key recognition [Gomez06, Mauch10] ○ Tempo/rhythm/meter characterisation [Gouyon05, Klapuri06] ○ Melody extraction [Salamon12] ○ Pitch estimation and multi-pitch estimation [Gerhard03, Klapuri08, Yeh10] ○ Genre/mood recognition [Tao10, Bischoff09] ○ Music structure analysis [Levy08, Paulus10] ○ Instrument recognition [Herrera03] ○ Sound effects classification [Zhang99, Casey02, Roma10] ● Methods for low/mid-level timbre characterisation of audio signals [Lakatos00, Terasawa05, Caclin05, Zwicker07]. ● Satisfactory accuracies for the above methods only achieved with models and algorithms trained for particular use-cases, typically for particular kinds of music pieces.
Challenges	● Consolidate high-level music annotations into actual semantic annotations with concepts defined in reusable ontologies. ● Develop new audio features optimised for the characterisation of music samples and non-musical content instead of music pieces. ● Improve high-level music characterisation tasks such as genre and mood classification by using broader training datasets including not only professionally-produced music but also user-generated content. ● Optimise music signal characterisation methods for the particular case of music samples. ● Develop higher-level timbre characterisation methods to semantically annotate non-musical audio content.
Our ambitions and	● Leverage existing approaches for high-level characterisation/annotation of music pieces to develop semantic annotation tools and improve their performance using big datasets comprising professionally and non-professionally generated material. ● Create new audio features tailored to the annotation tasks of music samples and non-

approaches	musical content (e.g., features for the description of the envelope of short events or features for the identification of loopable periodicities). ● Develop semantic annotation tools for music samples with useful musical properties for the reuse of the content (using new audio features). ● Develop semantic annotation tools for timbre properties, specially useful for describing non-musical audio content and providing innovative ways of retrieving that content. ● Use the aforementioned Audio Commons Ontology (see C) for representing the output of the semantic annotation systems.

1.4.2 Innovation potential

Our consortium has substantial **collective experience developing innovative research concepts into mature industry deployments** and operational environments. Because of the expertise that the members of the consortium have in the field of sound and music computing, semantic technologies, signal processing and sound and music repositories and business and IPR aspects; we have a **precise idea of the improvements and developments that should be carried out for the technologies** listed below in order to make the Audio Commons Ecosystem possible. We are therefore in a privileged position for approaching the research and innovation challenges of AudioCommons. The academic partners of the consortium have developed and matured innovative technologies and have access to content repositories with a close relation with AudioCommons. This, combined with the tools and technologies of the industrial partners, represents a huge innovation potential for the project. We now outline AudioCommons related technologies and tools created by different institutions of the consortium, which illustrates our innovation potential:

Technologies

- **Essentia (MTG-UPF)** [Bogdanov13]
  Essentia is an open-source C++ library for audio analysis and audio-based music information retrieval released under the Affero GPLv3 license. It contains an extensive collection of reusable algorithms which implement audio input/output functionality, standard digital signal processing blocks, statistical characterization of data, and a large set of spectral, temporal, tonal and high-level music descriptors. In addition, Essentia can be complemented with Gaia, a C++ library with python bindings which implement similarity measures and classifications on the results of audio analysis, and generate classification models that Essentia can use to compute high-level description of music. Essentia has been exploited for industrial applications: it has been used in industrial products by BMAT and Stromatolite (music selection), Yamaha's BODiBEAT (automatic playlist generation for runners), Steinberg's LoopMash (audio-based sample matching for music production), Korg's Cortosia/Artistry (real-time assessment of sound quality for instrumental sounds), and Brace Yourself Games' Crypt of the Necrodancer (beat tracking for video games).

- **Sonic Visualiser (QMUL)** [Cannam10]
  Sonic Visualiser[11] is an application for viewing and analysing the contents of music audio files. Since its initial release, it has been downloaded over 300,000 times. The aim of Sonic Visualiser is to be the first software to reach for when studying a musical recording rather than simply listening to it. Sonic Visualiser is of particular interest to musicologists, archivists, signal-processing researchers and anyone else looking for a friendly way to take a look at what lies inside the audio file.

- **Music Ontology (QMUL)** [Raimond07]

[11] www.sonicvisualiser.org

The Music Ontology provides a vocabulary for publishing and linking a wide range of music-related data on the Web. Music Ontology data can be published by anyone as part of a web site or an API and linked with existing data, therefore creating a music-related web of data. The Music Ontology is specified using the Ontology Web Language (OWL), which provides a set of constructs to describe domain models in Resource Description Framework (RDF). These standards and technologies are supported by the W3C and were chosen to tackle the many competing requirements of the music domain with a standalone format. By using RDF, the Music Ontology gains a powerful extensibility mechanism, allowing Music-Ontology-based data to be mixed with claims made in any other RDF vocabulary.

- **Vamp Audio Analysis API (QMUL)**
  Vamp[12] is an audio processing plugin system for plugins that extract descriptive information from audio data — typically referred to as audio analysis plugins or audio feature extraction plugins. Similarly to audio effects plugin (such as a VST), a Vamp plugin is a binary module that can be loaded by a host application and fed audio data. However, unlike an effects plugin, a Vamp plugin generates structured symbolic information. Typical features that a Vamp plugin might calculate include the locations of moments such as note onset times, visualisable representations of the audio such as spectrograms, or curve data such as power or fundamental frequency.

- **Sonic Annotator Web Application (QMUL)**
  SAWA[13] demonstrates audio feature extraction technology developed at the Centre for Digital Music, the standard plugin API for audio analysis and a Semantic Web ontology based representation for returning the results of the analysis. It is primarily aimed at researchers who wish to become familiar with these technologies. SAWA may be used as a batch feature extractor, for example, for automatically generating a small reference data set in RDF format. The resulting files can be loaded in a suitable RDF database and queried over, they may be browsed using a Semantic Web browser (e.g. the OpenLink RDF browser) or loaded and visualised with Sonic Visualiser a tool for studying the content of audio files.

Content repositories and datasets

- **Freesound (MTG-UPF)** [Font13]
  Freesound is a collaborative repository of Creative Commons licensed audio samples with more than 230,000 sounds (as of February 2015). Sounds are uploaded to the website by its users, and cover a wide range of subjects, from field recordings to music samples and synthesized sound effects. Audio content in the repository can be tagged and browsed by folksonomic means as well as standard text-based search. Audio content in the repository is also analysed using the open-source audio analysis tool Essentia, which powers the similarity search functionality of the site. Freesound has a RESTful API through which third-party applications can access and retrieve audio content and its metadata. Freesound's original goal was to give support to sound researchers, who often have trouble finding large royalty-free sound databases to test their algorithms, and to sound artists, who use pre-recorded sounds in their pieces. Nine years since its inception, Freesound has become one of the most popular sites for sharing sound samples, with 46,000 unique visits per day and more than four million registered users.

[12] http://www.vamp-plugins.org
[13] http://isophonics.net/sawa

- **Jamendo and Jamendo Licensing (Jamendo)**
  Jamendo is one of the biggest platforms for independent music in the world. The goal of the website is to bring an entirely free service for all music lovers to discover music, to provide the most simple ways for an artist to share its music online, and to offer a variety of music for professionals to license for commercial at a fair price. Started in 2005, it currently hosts 470,000 tracks from 35,000 unique artists, featuring 2.3 million members and 50,000 unique visitors/day. Jamendo Licensing (20,000 clients in 2014) is an extension of Jamendo, offering all the Jamendo music that is available for commercial use (use in film, TV, advertising, in-store background music) thanks to an opt-in option for all artists. Jamendo Licensing has almost 200,000 tracks (40% of the whole Jamendo catalog), the ones that the Jamendo artists have chosen to add to the program. All the content in Jamendo will be analysed and used to train AudioCommons algorithms and models (along with other datasets).

- **AcousticBrainz (MTG-UPF) and CALMA (QMUL)** [Wilmering15]
  The AcousticBrainz[14] project aims to crowd-source acoustic information for all music in the world and to make it available to the public. This information describes the acoustic characteristics of music and includes low-level spectral information and information for genres, moods, keys, scales and much more. The goal of AcousticBrainz is to provide music technology researchers and open source hackers with a massive database of information about music, which is accessible through a RESTful API. As of April 2015, AcousticBrainz has already computed audio features for 1.9 million music tracks which are linked to MusicBrainz[15] IDs providing standardised metadata. This database will spur the development of new music technology research and allow music hackers to create new and interesting applications. In relation to AcousticBrainz, CALMA[16] is a repository of audio features extracted from the Live Music Archive (LMA) holding copyright free live concert recordings uploaded to the Internet Archive. CALMA already includes audio features for more than 300,000 music tracks from 17,000 distinct live performances. Both AcousticBrainz and CALMA datasets will be used to train AudioCommons algorithms and models.

Creativity support tools

- **Waves Audio Plug-ins (Waves)**
  Waves have developed over 200 audio plug-ins which are used to improve sound quality in virtually every sector of the audio market, from recording, mixing, mastering and post-production to broadcast, live sound, and consumer electronics. As Waves products are being used by the majority of audio professionals around the world, Waves plugins are an unbeatable choice for incorporating AudioCommons in the workflow of sound professionals.

- **AudioWind, AudioRain, AudioMotors and AudioSteps (AudioGaming)**
  AudioWind, AudioRain, AudioMotors, AudioSteps are audio plug-ins based on mathematical models that generate real-time sounds for audio post-productions. Because of their quality and innovative nature, they have no equivalent on the market at this day. These products have been used by companies like LucasFilm, Ubisoft or Sounddelux on projects like Assassin's creed licence or Django Unchained for the most famous ones.

[14] http://acousticbrainz.org
[15] http://musicbrainz.org
[16] http://etree.linkedmusic.org/about/calma.html

Section 2: Impact

2.1 Expected impacts

AudioCommons is not only aligned with the target outcomes of ICT-19-2015 very closely, but it will deliver the expected impact in the specific topic as well as the broader work theme of *Content technologies and information management* of the current ICT work programme. In addition, **AudioCommons will make significant steps towards delivering transformational and universal impact on science, technology and society**. This will be achieved through an example of transforming the traditional linear value chain of audio media; in which content is delivered from producer to consumer, to a fully interconnected model that supports communication between all stakeholders. Innovative new technologies for annotation, access, transformation and licensing will play key role in supporting content reuse in unprecedented and unexpected ways, benefiting content creators and content users and effectively eliminating the boundaries between them. Open data reuse is a key aspect of the Digital Agenda for Europe (see IP/10/581, MEMO/10/199 and MEMO/10/200) and it is recognised as a key driver to develop content markets in Europe, which not only generate new business opportunities and jobs but also provide consumers with more choice and more value for money.

This work shall be carried out in the context of European level collaboration supported by researchers whose relations and professional networks **span a broad range of experts in multidisciplinary research** that is required to fulfil the goals and deliver the impacts of the project. The consortium is formed by top-level academic and industrial partners, including two of the world's biggest online repositories for user contributed Creative Commons sound and music content (Freesound and Jamendo), world's leading and award winner audio plugin developers (Waves), and innovative sound design and game development audio tools developers (AudioGaming). Furthermore, AudioCommons counts with the support of key persons from European companies and research institutions that constitute the AudioCommons associated experts and that will regularly be consulted and invited to AudioCommons discussions (see "Support letters" section at the end of the second part of this technical annex). This provides a strong basis for shaping research in *creativity support technologies* in ways that will prove economically and socially relevant in the long term.

2.1.1 Expected impact specific to objective ICT19a: Technologies for creative industries, Research and Innovation Action

The focus of AudioCommons is the development of tools that support content creation in a Web-based ecosystem, targeting two areas of the creative industries with the highest rate of dematerialisation (a process signifying a shift towards service economy) worldwide in 2013: *music production and video games development*. These areas are also within sectors that produced the highest growth rates in Europe in terms of employment in the last decade [Nem14]. The project will demonstrate cross-cutting impact in these industries by (i) **lowering the cost of content creation** through facilitating the reuse of existing Creative Commons content in commercial applications, (ii) **streamlining the creative process through novel technologies for annotation, search, retrieval and transformation of existing content**, and (iii) **increasing the quality of output** by exposing a large community of content creators to industries typically relying on a small number of experts, thus increasing competition and competitiveness. AudioCommons will achieve the impact expected for our specific topic (quoted in box) using the mechanism outlined below.

"Validated novel ICT technologies and tools supporting the creation process and delivering measurable benefits for the creative industries as regards time and resource investment, and quality of output."

The Audio Commons Ecosystem that will be developed during this project will have a clear impact on the

creation process of all kinds of multimedia content by interceding in the way in which audio resources are integrated in that content. Research on the definition of the Audio Commons ontology and on the annotation of sound and music content that meets the requirements of the ontology will allow unprecedented ways of accessing audio content, facilitating that access and reducing the time needed to get to the right content. Furthermore, the Creative Commons nature of the audio content shared in the Audio Commons Ecosystem will reduce production costs as an important part of the content is released with non restrictive licenses that allow their commercial usage, enable the reuse of the content in different scenarios, and foster co-creativity in audio solutions across a wide range of multimedia applications. For instance, 88% of the audio content released in Freesound and 40% of music released in Jamendo is licensed either CC0 or CC-BY terms.

Industrial partners of the AudioCommons consortium will develop **tools that can be directly integrated in existing professional workflows of media production**. Tools for integrating Audio Commons content in music production workflows will be developed by Waves. Similarly, AudioGaming will develop tools for the integration of Audio Commons content in sound design tasks for video game production or audio-visual production. Furthermore, Jamendo will develop a web interface for browsing music pieces released through the Audio Commons Ecosystem for being reused in creative industries such as advertising, films, music, publishing, video games, TV and radio. The industrial partners of the AudioCommons consortium expect that the use of such tools might considerably reduce the time required for audio content production in video games or audio-visual works by at least a 15%.

2.1.2 Enhancing innovation capacity, integration of new knowledge, and delivery to market

Sustained increase in productivity and innovation capacity must necessarily come from changes in technology [Basu01]. The AudioCommons project aims to **deliver long lasting and transformational changes in the content industries by creating an ecosystem that demonstrates novel ways of accessing, licensing and utilising audio media in the creative processes**. Although our focus is on audio, the Audio Commons Ecosystem will provide an example of content reuse that has the potential to transform entire sectors within the creative economy and open new pathways in the exploitation of user-generated video, narrative, literature, social media or open source software.

The project will introduce a wide range of audio resources in the creative industry, both professional and amateur. Being able to access and modify the already available Creative Commons content will make it grow over time, and also with the inclusion of new content providers. This means that content **users will have access to a living repository, continuously updated and enhanced with further contributions from content creators**, and opening an abundance of new possibilities. Finally, ACE will deliver cutting-edge new sound and music computing technologies which facilitate the use of audio commons in production environments, thus closing the gap between content creators and content users. The industry partners of the project will release tools that will enable in-place consumption of audio commons resources in the creative industries. These tools will provide easy access to vast resources of Creative Commons content that may be utilised directly or transformed for specific applications. The process will thus lower the cost of content creation and increase efficiency and productivity, as the traditional chain of commissioning and producing specialised content will be bypassed. Our platform will also enable the reuse of content by autonomous tools, such as sound retrieval agents in video games relying on procedural audio.

The impact of AudioCommons will necessarily depend on successful demonstration of acquired new knowledge. For example, this may come from the *evolution* of creativity support tools developed in WP6 and the *evaluation* thereof. Such **demonstrations should attract interest from stakeholders in creative industries as well as broader technology fields** and illustrate how current technology for supporting creativity may be developed beyond the state-of-the-art. The project will also provide an exemplar of delivering innovations to the market by **building tools that embed directly into existing products used by creative professionals** and evaluate how these tools facilitate human creativity (in WP6) and increase

competitiveness (in WP3). This evaluation will include analysis of user behaviour in the creative process, and research findings will be exposed to public, stakeholders and researchers through the activities in WP7.

2.1.3 Other environmental and socially important impacts

AudioCommons will provide a mechanism for creative industries to reach out and incorporate large bodies of creative individuals, who normally are limited to play the role of the consumer, into mainstream creative workflows. The system will ensure Creative Commons content is delivered with the quality of experience required by professionals, e.g., ease of access through high-quality annotation and search facilities. The project will also develop **new methods to support creativity and measure its success**. These elements of basic ICT research within AudioCommons have far-reaching social implications. We recognise that the impact will depend on uptake and willingness to relinquish some degree of autonomy in the creative process and integrate new smart technologies into everyday workflows. However, there are already numerous examples of the uptake of pervasive technologies and their impact on society, e.g. smart tools in word processors eliminate the need for copy editing for most types of written publications. Undoubtedly, improvements in creativity support tools will change the way people interact with technological artefacts whose role will be elevated from mere tools to communicating peers, perhaps in the distant future.

As mentioned, AudioCommons is also well aligned with a broader set of expected impacts in the *Content technologies and information management* work programme. It is expected to enable intelligent use of data from different sources and facilitate the creation, access, exploitation and reuse of a wide range of digital content. Through the deployment of smart support tools for creative industries, including licensing tools, it will facilitate communication and the exchange of cultural content fostering the deployment of the *Digital Single Market*[17].

The objectives of the project map to those of several related ICT topics and therefore some of their expected impacts will be realised through AudioCommons as well. Our ontologies (WP2) will enable turning large data volumes into **semantically interoperable data assets** and knowledge, facilitate **open data integration** and reuse and foster the development of open data supply chains (ICT15). AudioCommons as a whole may provide an example infrastructure for *Big Data* objectives (ICT16) in the audio and music domains. We will address the development of novel data structures (WP2), algorithms (WP4, WP5) and software architectures (WP2) for "*Big Music Data*" (in which the consortium has experience [Weyde14]), we will investigate and address scalability of data mining in WP4 and WP5, and focus generally on industry validated (WP6) and user-defined challenges (WP3). AudioCommons will *support the growth of ICT innovative creative industries SMEs* (ICT18) by stimulating the adoption and deployment of innovative ICT solutions and supporting them in leveraging emerging ICT technologies (WP3). AudioCommons will support composing and reusing audio content and services. The project will thus address interoperability requirements in a digital ecosystem (WP2, WP3), then develop and integrate tools (WP6) that **provide an example** which may be applicable in the context of interoperable learning and teaching tools, addressing objectives in *Technologies for better human learning and teaching* (ICT20). On of the industrial partners of AudioCommons is specialised in audio for video games (AudioGaming) and will create and evaluate tools for this sector. It will thus stimulate technology transfer via small scale experiments (WP3, WP6). These tools may become applicable and reduce the cost of developing games in a non-leisure context, addressing objectives of the ICT21 topic. Finally, AudioCommons will develop annotation tools (WP4, WP5) and novel Human-Computer interfaces (WP6) that facilitate search and retrieval from audio repositories.

[17] http://ec.europa.eu/digital-agenda/our-goals/pillar-i-digital-single-market

2.1.4 Barriers and obstacles to achieve the impact

The success of the project in delivering its expected impacts will largely depend on overcoming a number of barriers that are beyond our direct control and lay outside of the scope of technical risk factors detailed in Section 3.2.4. This section addresses how we are planning to overcome these.

Perhaps the most significant problem is presented by the barriers to the **uptake of certain AudioCommons technologies by the creative industries at large**. Barriers to uptake in the context of this project may manifest in two distinct forms. First, insufficient utilisation of the Audio Commons API and lack of industry-scale agreement on the Audio Commons Ontology. This may stem from conflicting views of the world by stakeholders with respect to the domain models prescribing how data is structured and accessed via our APIs. Second, barriers may be presented by limited uptake of the creativity support tools developed in the project. As hinted before, this may originate from the unwillingness to relinquish autonomy in the creative process, as well as mistrust in intelligent tools that connect the production environment to the "cloud".

We plan to overcome the first barrier by (i) carefully **assessing the requirements of content industries** (T2.1) in our domain, (ii) implement a **cyclic process** to enable the revision of ontologies and API specifications (T2.5), (iii) lay our work on the foundation of already established ontologies and APIs such as the Music Ontology, the Freesound API or the Vamp audio feature extraction API, and (iv) consulting **AudioCommons associated experts** in the field and from other relevant fields. The barrier to uptake in creativity support tools will be investigated during the first phase of the project. Prototyping will be carried out in collaboration with our industry partners in two phases. This will allow for incorporating tools for accessing and integrating AC content in existing workflows in a way that delivers the most benefit and enable unexpected, novel ways of content creation. The two phase process will also allow for any usability issues to be addressed on time. By designing tools that are engaging and increase productivity we are hoping that the benefits will prevail over potential issues of losing creative autonomy and mistrust.

Low commercial readiness of certain technologies may also present a barrier. For instance, certain semantic technologies for automatic content annotation (see WP4 and WP5) are only tested in a lab environment, typically using small datasets or close to ideal conditions. It is expected that professional applications require more precision or general applicability of these technologies than what is scientifically validated. The project offers an opportunity to overcome this barrier by **deploying, testing, validating and gradually improving technologies with large industry relevant audio content and the user base available to our partners**. Using an iterative user-centred design, we will focus our core technologies to offer the features that users naturally feel ready in practice and desire to use.

2.2 Measures to maximise impact

In order to maximise the impact value of AudioCommons, we plan to take advantage of strategies for developing and deploying our innovations and technology in open ways. Consortium members have already strong track records in the public deployment, dissemination and exploitation of research technology. Our initial dissemination plan includes the activities outlined in the following sections.

2.2.1 Scientific dissemination

The AudioCommons consortium is committed to knowledge production and transfer through academic communities. To maximise AudioCommons impact on the scientific community we will disseminate our work in top conferences and journals. A non-exhaustive list of relevant conferences which the academic partners of the consortium usually attend includes ACM International Conference on Multimedia Retrieval (ICMR), IEEE International Conference on Multimedia & Expo (ICME), Sound and Music Computing Conference (SMC), International Computer Music Conference (ICMC), International Society for Music

Information Retrieval Conference (ISMIR), Annual Hawaii International Conference on System Sciences, IEEE International Conference on Multimedia and Expo (ICME), International Semantic Web Conference (ISWC), and Audio Engineering Society Convention. A non-exhaustive list of relevant journals includes ACM Transactions on Multimedia Computing, Communications and Applications (TOMCCAP), ACM Transactions on Information Systems and Technology (TIST), Journal of Web Semantics, International Journal on Semantic Web and Information Systems (IJSWIS), IEEE Transactions on Multimedia (TMM), Computer Music Journal (CMJ), IEEE Multimedia, International Journal of Multimedia Information Retrieval (IJMIR), Advances in Multimedia Applications of Multimedia Mining, Multimedia Tools and Applications (MTAP), ACM Transactions on Information Systems (TOIS), Journal on Knowledge-based Systems (KNOSYS), IEEE Transactions on Audio, Speech and Language Processing, Journal of the Audio Engineering Society, Journal of the Acoustical Society of America, and IEEE Journal of Selected Topics in Signal Processing.

2.2.2 Research data management

Throughout the activities carried out in AudioCommons, multiple types of data will be generated and acquired, both for sharing/dissemination activities as well as for internal use. The core ideas of AudioCommons share open data and open metadata principles, which we will apply to the management of generated research data. This subsection outlines what the data types will be, what standards/formats will be used and how they will be shared and curated. It should be noted that AudioCommons has applied to participate in the **H2020 Open Research Data Pilot**, and as a result we have included the definition of a comprehensive Data Management Plan as a task in WP1, to be carried out within the first six months of the project and then iteratively updated. Therefore, the currently provided research data management plan will be subject to refinement during the project's initial steps.

Types of data
Data collected from the listening tests and evaluation sessions:

- User activity tracking data
- Questionnaire data

Data to be collected from the content providers of the consortium, MTG-UPF (Freesound) and Jamendo:

- Audio recordings
- Existing audio metadata

Data to be generated:

- Numerical features extracted from raw audio signals
- Datasets of annotated audio content
- Semantic metadata for audio content
- Documents (deliverables, reports)

Standards, formats, data sharing and preservation
Audio recordings will be collected in the formats that their authors originally provided, with a preference for lossless standard audio formats (WAVE, AIFF, FLAC). Before processing, audio content will be transcoded to the lossless WAVE format. The datasets that we publish will consist of annotations for particular audio resources. Annotations will be stored using standard formats such as JSON and YAML, and Semantic Web formats such as RDF/XML and N3. These datasets will include references to the original audio content from their respective content provider sites (i.e., we will not redistribute audio data not generated by ourselves) and respecting their original format. Complete datasets will be easily downloaded using scripts provided by the AudioCommons consortium and making use of the Audio Commons API. Questionnaire data will be stored using interchangeable spreadsheet formats such as the Comma-Separated Values (CSV).

Documents generated for the needs of the project such as deliverables and reports will be stored on the AudioCommons website with username/password protection as necessary for access control. It is in the

interests of our consortium to have an open data policy to the maximum degree that is afforded, both in the service of contributing to standardisation and good practices as well as fostering research in the field. Therefore, most of the deliverables (see Table 3.1c) and tasks of the work plan will produce public documents (not requiring username/password for their access). In all cases where industrial partners' data or user data are sensitive or required to be anonymised, they will be treated as research data in line with the ethics policy available in Section 5.

2.2.3 Knowledge management, protection and IPR

AudioCommons, as a project within Information and Communication Technologies EU scheme will comply with **open access policies** where publications and similar outputs are concerned. The consortium will comply with the following objectives:

- To deposit articles resulting from AudioCommons into an institutional or subject-based repository (such as MTG-UPF's OpenAire compliant repository e-Repositori[18]).
- To ensure open access to these articles within six months (including gold open access when required).

Wherever possible, we aim to provide universal online access, free of charge to all content and metadata. The Audio Commons Ontology **specification will be open sourced**, as well as the Audio Commons API specification and the other service integration and automatic annotation tools developed by the academic partners of the consortium. Generated **metadata will be released with a Creative Commons 0** (Public Domain) license. The Embeddable tools developed by the industrial partners of the consortium will be **accessible to the public** according to their developers' commercial exploitation model. Nevertheless, one of the industrial partners (Waves) will develop an **add-on for an open source digital audio workstation** which will effectively be open source as well (see description of WP6).

MTG-UPF, QMUL and Surrey consider the principles set out in the Principles for an internal intellectual property policy and the Principles for a knowledge transfer policy of the Code of Practice included in the Commission Recommendation on the management of intellectual property in knowledge transfer activities[19]. Researchers working on AudioCommons will deposit final articles or manuscripts into the institutional repositories of the research institution with which they are affiliated, compatible with the Open Access initiative:

- **Spain**: Spain currently has around 59 institutional repositories and a fully operable national repository network. Most Spanish repositories (78 %) are institutional, but there are also research institutions, like the Spanish National Research Council[20] (CSIC) and even private organisations involved in the development of different kinds of Open Access repositories. Library Consortia[21] play a key role in the creation of repositories, especially in Catalonia (CBUC) and Madrid (Consorcio Madroño) whose respective territories account for most of the existing institutional repositories in Spain.
- **UK**: The vast majority of open access projects and initiatives have been funded by the JISC. The JISC is responsible for the disbursement of funding for the development of IT in higher education. There are currently 192 repositories containing full-text content within the UK[22]. This figure includes institutional and disciplinary repositories.

[18] http://repositori.upf.edu
[19] http://ec.europa.eu/invest-in-research/pdf/ip_recommendation_en.pdf
[20] http://www.csic.es/
[21] http://icolc.net/
[22] http://www.opendoar.org/

Before the project starts, a Consortium Agreement will be designed based on the DESCA model for H2020 and signed, preventing all and any unauthorised exploitation of any existent background IP by any consortium member and protecting each partner's background IP in perpetuity, during and beyond the project. The Consortium Agreement will be co-developed and overseen by each entity's legal representatives and will enable a clear exchange of innovation technologies/models that we can build on together to form all tools interconnected in the Audio Commons Ecosystem.

In terms of Background, the AudioCommons partners may ask for access rights to any of the other partner's Background technologies used in the project as far as they need this access in order to carry out their tasks. In addition, to ensure minimum possible conflict between innovation partners (e.g. between academic institutions working the ontology/API definition and the tools for the semantic annotation of audio content, and industry partners in the development of embeddable tools), the Consortium Agreement will establish the basis also in terms of Foreground IP which will be owned by the developing party, in line with EU IP guidelines. In the case where the IP of background is not fully owned by a specific consortium member, this member will be responsible of reaching required agreements that allow providing other partners with access rights to the background technologies if needed. The Foreground produced during the project execution will be identified, captured, assessed, managed and nurtured. The general policy will be to avoid joint ownership of the Foreground in order to prevent potential blocking factors for the exploitation of project technologies beyond the project scope and duration. In this way, each partner's IPR will be fully protected, being brought to the project as existent and complete, available for use by all other consortium partners during the project and also beyond through dual license scheme models (including xGPL, MIT & proprietary licenses).

2.2.4 Dissemination to the professional sector and to wider audiences

AudioCommons dissemination activities include the dissemination of the project to wider audiences including the **professional sector through different communication channels**. Firstly, general dissemination of the project will be carried out through **online communications**. AudioCommons online dissemination strategy aims to be not just a record of project outcomes, but a dedicated dissemination portal. The project website will contain a constantly updated stream of audio, video and blogging activities detailing our progress. It will create awareness and deliver access to the Audio Commons API and provide information for joining the Audio Commons Ecosystem. Our web presence will serve to reinforce our existing contacts with music technology actors. All industrial partners websites will syndicate their main project outputs and marketing force through this portal, driving traffic from their customer base. Finally, the website will contain major project outputs including the AudioCommons API and all guidance frameworks and best practice documentation.

Secondly, we will target the dissemination of project outcomes to the creative industries in the professional sector. The partners in the consortium, particularly the industrial partners, will disseminate AudioCommons technology and tools by attending **specialised trade shows and industry events** such as MIDEM (Cannes, France), MaMA (Paris, France), Music Tech Fest (San Francisco, USA / London, UK), South By SouthWest (Austin, USA), Game Connection (Europe and USA), Game Developers Conference (GDC, San Francisco), AES Convention and AES Audio for Games conference. In these trade shows and conferences, AudioCommons outcomes will be disseminated in stands for the project showing outcomes and products from the different partners of the consortium, and with the intention to spread the word of AudioCommons, promote developed tools and technologies, and gather new content providers and tool developers to connect to the Audio Commons Ecosystem. In events only attended by one partner of the consortium, the project outcomes will be disseminated in their stand along with their own products. Members of the consortium will also be present in more targeted events like the Audio Meetup organised by AudioGaming for south of France in Toulouse, in which AudioCommons outcomes will also be disseminated. Furthermore, we will disseminate the AudioCommons ideas, technology and developed tools in specialised magazines and online blogs for the professional sector such as Sound On Sound and Create Digital Music.

Thirdly, AudioCommons aims to **disseminate project outcomes in public events**. All three research partners have excellent experience in presenting novel interface technology to generalist audiences in public events in high profile venues. QMUL has presented at Music Tech Fest, Music Hack Day events[23] and Digital Shoredich, a two week tech community event, inviting the entire world of tech, creative and converging industries[24]. MTG-UPF has presented, among others, at Barcelona's CCCB Centre for Contemporary Culture and at Sonar festival[25], and Surrey has presented at London's Science Museum and to members of the UK Houses of Parliament. MTG-UPF organises the Barcelona chapter of the Music Hack Day (MHD), a 24 hour coding session in which participants conceptualise, program and present their projects. The MHD has gathered over 2000 participants worldwide, building hundreds of hacks and with over 125 music and tech companies supporting the events. MTG-UPF has hosted MHD Barcelona since 2010 in partnership with the Sonar festival. The Audio Commons API and tools will be disseminated and open for feedback at this event.

We will promote the outputs of this research through **events and other activities aimed towards the general public**, including exhibitions, newspaper and magazine articles such as New Scientist and the BBC World Service who have been involved in previous publicity of work with QMUL. We will disseminate the work in specialist forums such as Assistive Technology magazines and bulletin boards. We will also develop targeted information for educational institutions (e.g. through cs4fn and Audio! Magazine; cs4fn.org), and engage a broad public via YouTube videos and public outreach events such as QMUL sponsored TeenTech[26], which holds 10 events a year to over 300 students each time. Wider dissemination activities are also exemplified by the Human Harp project[27], turning bridges into instruments and art installations using open sound repositories[28].

Finally, AudioCommons outcomes will also be **disseminated in workshops and educational activities**. Workshops in schools provide public engagement opportunities and provide an opportunity to gather user experience during the innovation design process. MTG-UPF carries out educational workshops at schools through the initiative Sons de Barcelona[29], regularly organises concerts in the frame of the PHONOS project[30] and is very active in the organisation of International Conferences (SMC'10, TEI'13), including its Summer Schools and Graduate Students' activities. In this way, community relations, user testing, grassroots promotional activities and knowledge transfer are combined. AudioCommons sessions at such workshops will be important in informing the development of core pedagogical materials based on the consortium's technology as ways to maximize take-up and adoption.

2.2.5 Commercial exploitation

The Audio Commons Ecosystem will provide **several opportunities for commercial exploitation ranging from direct commercial conversion to indirect knowledge transformation**. Immediate commercialisation of the different components and technologies will be available to early adopters willing to provide value added services within ACE. Examples of these include the provision of new types of content, specialised licensing services, or on demand content analysis that provide commercially valuable metadata not yet available in the system. The creativity support tools developed by the consortium will become part of the products currently offered by Waves, Jamendo and AudioGaming. This is not only a pathway to direct

[23] http://musichackday.org
[24] http://digitalshoreditch.com
[25] http://sonar.es/en/pg/sonarpro
[26] http://www.teentech.com
[27] http://humanharp.org
[28] http://www.openmusicarchive.org/projects
[29] http://barcelona.freesound.org
[30] http://phonos.upf.edu/?language=en

commercialisation, but it also creates **new markets for content and metadata services through technology push**. To proactively support this process, the project will **research market opportunities**, create business plans available for partners, early adopters and third parties (WP3) and assess monetisation pathways for creative content to be sold and consumed as well as reused within the ACE.

Knowledge transformation opportunities will primarily arise from the scientific work carried out in WP2, WP4 and WP5, focussing on information management solutions and the semantic analysis of audio content. The knowledge created in our **core RTD activities will be transferrable to the wider content industries** after additional work. This includes the ontologies, which may be extended to describe video content or support rights management in audio-visual media. Semantic audio analysis technologies will bear relevance on information management technologies that rely on advanced AI and machine learning methods.

Section 3: Implementation

3.1 Work plan — Work packages, deliverables and milestones

In Section 1.3.4 of this proposal we outline the approach and methodology that we will follow to achieve the goals of AudioCommons. In this section we explain how the different steps of our work plan are represented as work packages and the interrelations between them. The overall structure of the project is divided in seven work packages as shown in the Pert chart below (Figure 3.1). Project coordination and dissemination activities will be carried out in WP1 and WP7, which run in parallel throughout the duration of the whole project. Intellectual property and business model aspects will be addressed in WP3 and will also start at the beginning of the project, running in parallel with the rest of the work packages for most of the project duration. The remaining work packages (WP2, WP4, WP5 and WP6), include the work to be done in relation to the semantic representation (i.e., the ontology) and annotation of sound and music content, the development and evaluation of the tools that will be embedded in creative workflows and access Audio Commons content, and the orchestration of the different services that conform the Audio Commons Ecosystem. These work packages will start progressively, and will be interconnected via a number of deliverables that will inform tasks to be carried out in other work packages. WP2 and WP3 will also be interconnected as the definition of the semantic representation of Audio Commons content (WP2) will include licensing information whose requirements will be determined in WP3.

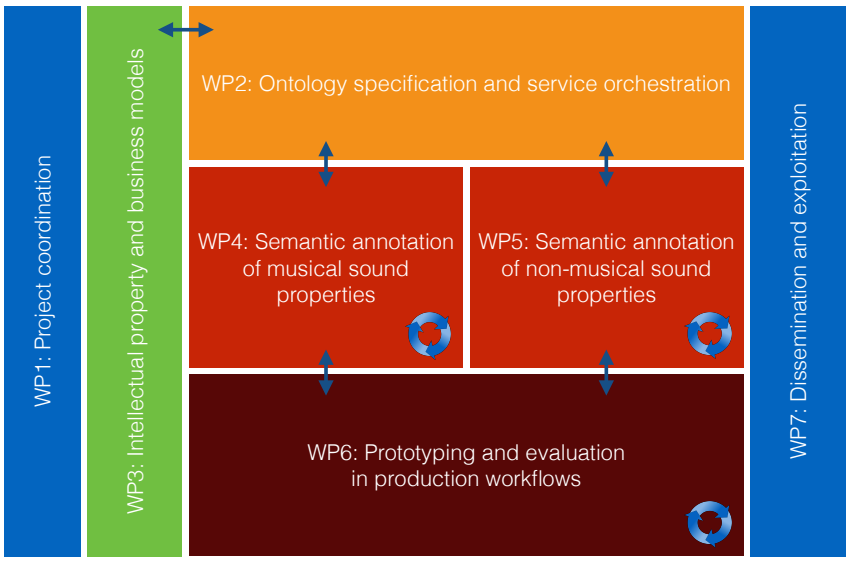

Figure 3.1: Work Packages Pert chart

The first milestone of AudioCommons (MS1) is set at month 9 (M9). In MS1, WP3 will have carried out work regarding the intellectual property aspects related with Audio Commons. Similarly, WP2 will have defined that part of the ontology that is concerned with licensing information and in providing the process through which content creators will be able to license content (when required) to content users.

After 15 months from the start of the project, there will be a second milestone (MS2) in which the whole Audio Commons Ecosystem will have been set up with prototype versions of all of its components. By that time, WP4, WP5 and WP6 will have already entered a research/development and evaluation cycle in which the prototypes will be refined to generate the second and final versions of the ACE, to be delivered in MS3 and MS4 respectively. Figure 3.2 shows a Gantt Chart illustrating the duration of the different tasks and

deliverables of each work package, and the milestones of AudioCommons. What follows is a brief description of the objectives of each work package. Detailed descriptions of milestones, tasks and deliverables of each work package are given below in tables 3.1 and 3.2.

WP1 activities address the overall **coordination and management** of the project. A suitable structure (see Section 3.2) will be in place in order to guarantee fulfilment of the project objectives. It covers high-level issues and daily management, as well as the integration of regular daily activities and the setup and implementation of efficient coordination and management of the overall project.

WP2 will be focused on the definition of the **Audio Commons Ontology** that specifies how to represent Audio Commons data in terms of licensing information and audio properties that should be annotated. The definition of these audio properties will be done according to requirements provided by the industry partners of the consortium and interviews with content users. Closely following the Audio Commons Ontology, this work package will also 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). Finally, WP2 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. The content providers of the AudioCommons consortium, Jamendo and MTG-UPF (Freesound), will integrate these technologies.

We have explained in Sections 1 and 2 that commonly understood frameworks for publishing and master rights to particular audio and music recordings will be challenged within the ACE's framework. The objective of **WP3** is to make those challenges understandable, and ultimately, useful for the industry. In the first instance, this will involve **understanding the rights management requirements** from the perspective of how audio is described and tagged at the point of creation and ingestion into the ACE. Then, use case analyses will be performed on the different **Creative Commons licensing possibilities**, and recommendations made as necessary. Finally, guidelines will be published in order for the creative industries to understand how new actors in the Audio Commons Environment can participate, and research into **emerging business models possibly created by ACE** interaction with publishers/creators/consumers carried out.

In **WP4** we will carry out the research and evaluation of **methods for semantically annotating musical content** to be exposed in the Audio Commons Ecosystem (from single notes, percussive hits, chords, melodies and loops to complete music pieces). The content will be annotated in terms of musically meaningful audio properties as defined in the Audio Commons Ontology. With this work package we will enhance the usefulness of existing content and facilitate its creative uses by: (i) developing a system to automatically add musically-meaningful metadata to existing content; and (ii) developing better tools for manually annotating music samples and music pieces. In this work package, the content providers included in the Audio Commons consortium that host musical content (Jamendo and Freesound) will integrate the developed tools in their audio description systems.

WP5 will be devoted to the **semantic annotation of non-musical content** to be exposed in the Audio Commons Ecosystem (e.g., sound effects and soundscapes). In contrast with WP4, the annotations in this case will be centered around timbral attributes of audio content as defined in the Audio Commons Ontology. This Work Package will enhance the usefulness of existing non-musical content and facilitate more creative uses by: (i) developing better tools for manually annotating sound effects and soundscapes; and (ii) developing a system to automatically add timbral metadata, such that content can (additionally, or instead) be searched by perceptual sound quality (piercing, crunchy, rich, etc.). Similarly to WP4, the content providers included in the Audio Commons consortium that host non-musical content (Freesound) will integrate the developed tools in their audio description systems.

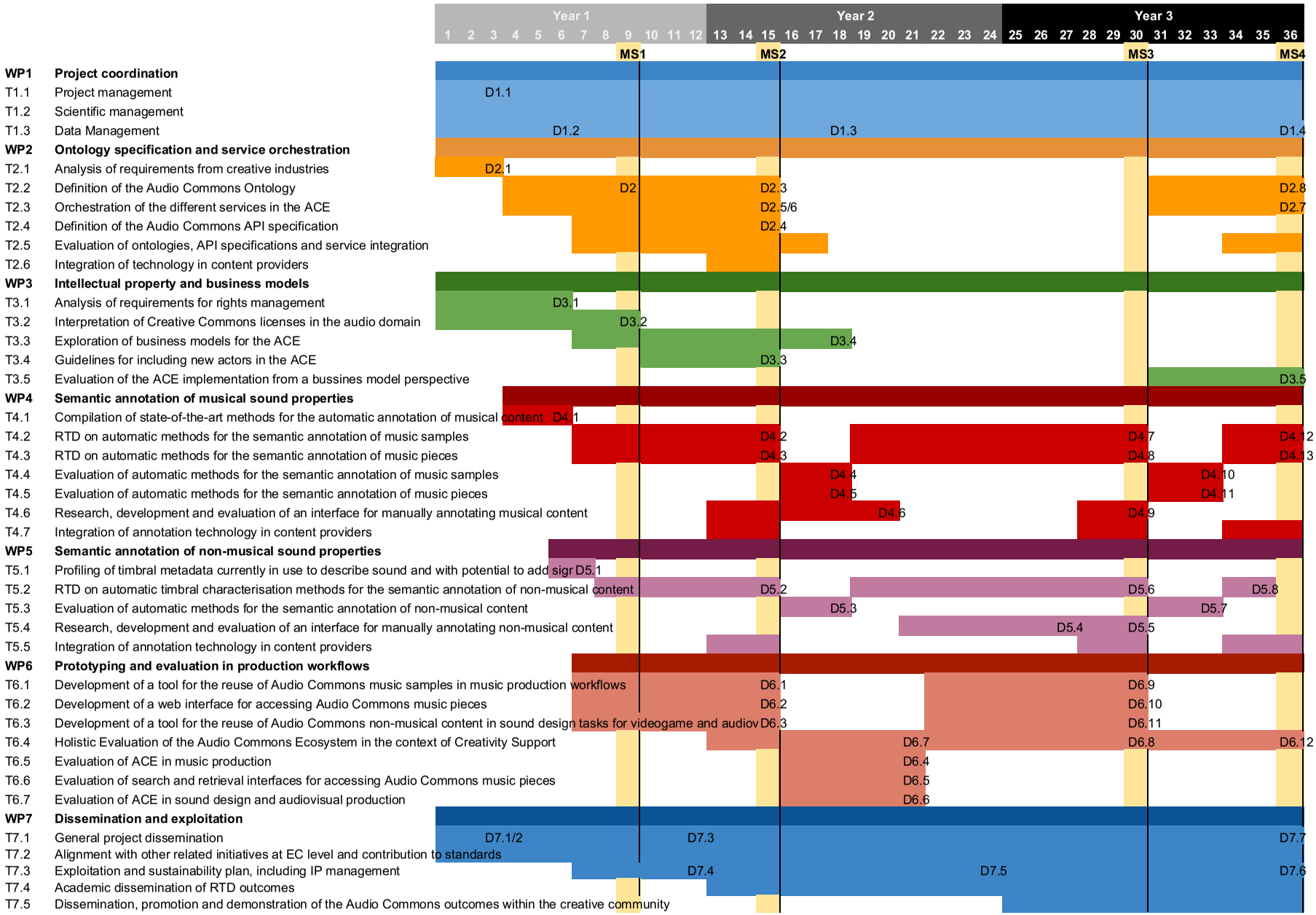

Figure 3.2: Tasks, Deliverables and Milestones Gantt chart

WP6 activities will be focused on the **development and evaluation of tools to access Audio Commons content** and to be embedded in production workflows of the creative industries, and in the evaluation of the ACE as a whole. Even though the industrial partners of the consortium will be responsible of the development of these tools, they will all be coordinated by one of the research partners (QMUL) which will provide support for the service orchestration and lead the evaluation of the prototypes with content users. The different tools that will be developed will be focused on accessing different kinds of audio content in the Audio Commons Ecosystem (i.e., music samples, music pieces and non-musical sounds), and will be evaluated in use-cases of the creative industries including **video games production, music production, sound design and audio-visual production**.

Finally, **WP7** will address the **dissemination of the project's results on various levels**, ensuring the proper outreach of Audio Commons activities to all relevant stakeholders in the creative industries (content users and content creators, potential new content providers and tool developers) and the scientific community. The work package will also include an exploitation and sustainability plan of the Audio Commons Ecosystem beyond the duration of the project, including the different business models that can arise in the ecosystem as analysed in WP3.

Table 3.1: Work package description

WP1 - Project coordination						
Work package No.	WP1			**Starting date**		M1
Work package title	Project coordination					
Participant number	1	2	3			
Short name of participant	**MTG-UPF**	QMUL	Surrey			
Person/months per participant	18	6	6			

Objectives
● Efficient communication with EC, including timely and complete reporting on project status and compliance with programme procedures. ● Ensuring efficient internal communication and providing timely information to project partners. ● Planning and overall coordination of all WPs. ● Ensuring effective coordination and collaboration between project partners. ● Constant monitoring of the project progress in accordance with the project objectives. ● Ensuring that the overall goals of the project are met within the time and budget constraints. ● Guaranteeing coherence in project development and in the achievement of the expected results. ● Ensuring that any planned alterations are approved by the Steering Committee. ● Ensuring effective coordination between the project and external organisations. ● Organizing effective project meetings and reviews. ● Coordinating the preparation, quality assurance and distribution of all reports. ● Risks monitoring, prevention and management.

- Arbitrating of conflict resolution procedures if required.

Description of work and role of partners
WP1 is aimed at ensuring that correct procedures are carried out during the project and that all deadlines and obligations met. Consensus on project steering will be sought by preparing clear instructions on actions to be performed. In order to design, implement, facilitate and monitor the execution of the actions described in the work plan, the WP1 responsible partners will be in charge of the preparation of project meetings, where project goals, plans, procedures and relevant information on progress will be discussed. WP1 will be executed mainly by the Project Coordinator (MTG-UPF), with the support of QMUL and Surrey, and will be divided in the following tasks: **Task 1.1: Project management [M1-M36]** This task addresses the administration and financial reporting to the Commission: monitoring progress of the project in its various administrative aspects (finance, reporting to the Commission, notification and monitoring of deadlines, coordinating management meetings, etc.); establishing communication procedures; ensuring submission of deliverables, analysing potential risks, remaining aware of their palliative solutions, and driving the process of implementing corrective actions when required via the project's predefined decision mechanisms; management of the consortium and mediating in the event of conflicts and disputes. The Project Coordinator is also responsible for ensuring that quality deliverables are submitted on budget, and that milestones are reached according to the set timelines in close cooperation with the coordinators of other Work Packages Leaders. Internal progress reports will be delivered every 6 months in order to ensure a proper project monitoring. The list of activities aimed at achieving gender balance within the workforce is as follows: set targets to achieve gender balance in decision-making positions, promote women's participation in Consortium research activities, committees and working groups and family friendly working conditions, promoting the work from home when needed. This task will be carried out by MTG-UPF. Interdependencies: (none) Expected results: Efficient and effective project coordination, deliverable D1.1. Success criteria: Possible deviations from DoA timely detected and addressed, timely deliverable of management reports. **Task 1.2: Scientific management [M1-M36]** This task is responsible of coordination and constant monitoring of the technical activities of the project. It will be in charge of preparing, monitoring and updating a detailed project work-plan, of organizing technical meetings, of boosting collaboration between partners (e.g., through exchanges) for handling technical issues, of proposing strategies for solving possible technical problems. Other activities included in this WP are developing Guidelines for Quality Assurance in Research carried out in the project, monitoring the quality of scientific output, detecting risks and taking corrective action as necessary. This task will be led by MTG-UPF, in collaboration with other WP leaders. Interdependencies: (none) Expected results: Efficient and effective scientific project coordination. Success criteria: Possible deviations from DoA timely detected and addressed, timely deliverable of S/T reports. **Task 1.3: Data management [M1-M36]** This task is devoted to handle the multiple types of data that will be generated throughout the project activities, both for sharing/dissemination tasks and also for internal use. For disseminating the project results, open access policies will be followed for the publication of research findings and best practices. This task will be led by Surrey, in collaboration with MTG-UPF and QMUL. Interdependencies: Tasks T7.1 and T7.3. Expected results: Deliverables D1.2, D1.3 and D1.4

Success criteria: All data types being produced properly identified and managed, data and metadata produced accessible by 3rd parties to mine, reproduce and disseminate during and beyond the AudioCommons project lifespan.

Deliverables

D1.1 Project handbook and quality plan [M3]
Project handbook and quality plan including all procedures, communication channels and operational framework.

D1.2 Draft data management Plan [M6]
First draft of the data management plan (DMP) following EC guidelines, describing the data to be generated during the project and how will this be made public and accessible.

D1.3 Second draft data management Plan [M18]
Second draft of the data management plan (DMP) following EC guidelines, describing the data to be generated during the project and how will this be made public and accessible.

D1.4 Data management Plan [M36]
Data management plan (DMP) following EC guidelines, describing the data to be generated during the project and how will this be made public and accessible.

WP2 - Ontology specification and service orchestration						
Work package No.	WP2			**Starting date**		M1
Work package title	Ontology specification and Service Orchestration					
Participant number	1	2	4	5	6	
Short name of participant	MTG-UPF	**QMUL**	Jamendo	AudioGaming	Waves	
Person/months per participant	18	30	8	8	8	

Objectives

- Analysis of requirements from the creative industries.
- Definition of the Audio Commons Ontology.
- Specification of the Audio Commons API.
- Release of tools that allow the interconnection of different actors in the ACE.
- Integration of technologies in the content providers of the consortium (Freesound and Jamendo).

Description of work and role of partners

WP2 is mainly aimed at the specification of the Audio Commons Ontology and the orchestration of the

different services that will enable the Audio Commons Ecosystem. The Audio Commons ontology will specify how to represent Audio Commons data not only in terms of which audio properties need to be annotated in order to provide meaningful retrieval, but also in terms of the license information that needs to be embedded to allow licensing procedures and safe reuse of content. The service orchestration includes the definition of the Audio Commons API and the release of the necessary tools that enables the interconnection of different actors in the distributed nature of the Audio Commons Ecosystem. This work package will be led by QMUL, and will include the following tasks:

Task 2.1: Analysis of the requirements from creative industries [M1-M3]
Current practices in creative workflows within our target domains, such as game audio or sound design, involve the use of existing tools and services. This task will rely on (i) knowledge within the consortium to analyse requirements for successfully integrating Audio Commons content in these workflows and (ii) structured interviews with a broader set of content users to enable formalising 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 focussing the work on sound and music analysis algorithms and end user prototypes. This task will be led by QMUL, and carried out in collaboration with MTG-UPF and the industrial partners in the consortium.
Interdependencies: (none)
Expected results: Deliverable D2.1.
Success criteria: Identification of factors that affect the use of Audio Commons content in creative workflows within our target communities.

Task 2.2: Definition of the Audio Commons Ontology [M4-M15, M31-M36]
The heterogeneous nature of information related to audio and music makes it difficult to design metadata schema to fulfil diverse and often conflicting requirements within this domain and our perspective applications [Kolozali13]. The use of Semantic Web technologies, such as the Resource Description Framework (RDF) and the OWL Web Ontology Language, presents a solution that has been proposed by the W3C to provide interoperability, flexibility and extensibility for structuring information on the Semantic Web, augmenting the common Web using machine interpretable data formats. This task aims to define a formal ontology for the Audio Commons Ecosystem, the Audio Commons Ontology, that provides the anchor for the design of Web APIs, audio feature extraction and processing APIs and rights management. The task will extend successful ontologies such as the Music Ontology [Fazekas11] with terms relevant in the broader context of ACE, e.g., to support the description of non-musical sounds and music samples. AC Ontology will also examine or extend rights management ontologies in the media domain such as the Media Value Chain Ontology (MVCO) standardised by MPEG [Kudumakis14]. At the end of this task, we will also evaluate the defined ontology in terms of its adaptation within ACE and future maintenance strategies. This task will be carried out by QMUL in collaboration with MTG-UPF.
Interdependencies: Tasks T2.1 T3.1 and T5.1
Expected results: Deliverables D2.2, D2.3 and D2.8.
Success criteria: Formal verification of AC Ontology for consistency and other logical properties. Adaptation of the ontology throughout and possibly outside of ACE.

Task 2.3: Orchestration of the different services in ACE [M4-M15, M31-M36]
The Audio Commons Ecosystem will feature a set of distributed services that include content management (e.g. search and retrieval, recommendation), license management and negotiation, service discovery as well as support services for lower-level tools such as automatic content annotation. The primary challenge

we anticipate is in accessing heterogeneous information from different content repositories and the successful integration of services within ACE. Our ecosystem may interlink for instance a content repository such as Jamendo, a licensing server that allows the commercial exploitation of Jamendo content (Jamendo Licensing) and a feature extraction server, which exposes or calculates content-based metadata on demand. In essence, ACE will allow delegating responsibilities to different entities. This will provide scalability and sustainability for our system, as well as create business opportunities for our commercial partners and third parties. However, this comes at an expense of additional technical challenges. The system will require sophisticated service orchestration and choreography (a method of service composition) that allows for instance redirecting a query to a specific service end-point that is able to answer it, or composing service end-points for complex queries on demand based on a process description that prescribes how the query should be answered. An example of this is a complex query that asks for a recording with a specific tempo and mood and one that can be licensed for commercial use. We may assume that mood annotation is available from the content holder, tempo annotation needs to be obtained from an on-demand feature extraction service while license information is obtained from a third party service. The composition of a federated system that is able to answer such queries and execute underlying processes defined within this task may rely on the calculus of communicating systems, also known as pi-calculus [Milner, 1999].

This task will develop the necessary technological solutions for the interconnection of the different actors in the Audio Commons ecosystem. The task will also perform evaluation of these solutions and provide guidelines for their implementation. Within this task we will push the boundaries of the theory and implementation of services built on the Semantic Web stack. We will assess and potentially advance technologies such as the SPARQL 1.1 Federated Query specification, Business Process Model Notation (BPMN), the Web Services Choreography Description Language (WS-CDL), and OWL-S, a Semantic Web framework for describing services. Software libraries will be built to support creating and accessing ACE. This task will be carried out by QMUL, with the collaboration of MTG-UPF and with inputs from Jamendo and Waves for their experience in licensing procedures.

Interdependencies: Tasks T2.1, T2.2, T2.3, T3.1, T4.2, T4.3 and T5.2

Expected results: Deliverables D2.5, D2.6 and D2.7.

Success criteria: Release of software tools and guidelines for integration.

Task 2.4: Definition of the Audio Commons API specification [M7-M15]

Common Application Programming Interfaces (APIs) are key to interoperability between tools and services that will comprise the Audio Commons Ecosystem. This development will follow two correlated streams to define APIs for end user tools and APIs for tools that are used in ACE's backend processes. End-user tools will connect to ACE compliant services using web APIs that follow the REST architectural style [Fielding00]. This choice is partly motivated by the ability to directly relate the AC Ontology to resources offered by ACE through state of the art technologies recommended by the W3C. These include the Linked Data Platform specification (www.w3.org/TR/ldp/) and JSON-LD (www.w3.org/TR/json-ld/), a flexible messaging format for building interoperable web services. The end-user APIs will serve two main purposes. Firstly, they will provide access to metadata as well as search and retrieval facilities and other content management services offered by ACE, second, they will facilitate content licensing following the requirements outlined by D3.1.

An additional set of APIs will be designed for backend processes within ACE, for instance, audio content annotation. These will support automatic and semi-automatic content annotation protocols for utilising the algorithms provided in T4.2, T4.3 and T5.3. We consider building a feature extraction API by extending

the VAMP audio analysis plugin API (vamp-plugins.org) defined in C++ with appropriate higher-level language bindings [Cannam11]. A number of audio analysis plugins have already been developed that support this API incorporating the Essentia framework (http://essentia.upf.edu/) by the MTG as well as the Queen Mary Vamp Plugin set and many others (www.vamp-plugins.org/download.html). We aim to push the boundaries by incorporating algorithms for the broad range of content types within ACE and modify our existing APIs as well as create new APIs that support their requirements. This task will be carried out by MTG-UPF, in collaboration with QMUL and the industrial partners of the consortium.
Interdependencies: Tasks T2.1, T2.2, T2.3, T3.1, T4.2, T4.3 and T5.3.
Expected results: Deliverable D2.4.
Success criteria: Release of basic tools to follow the API specification.

Task 2.5: Evaluation of ontologies, API specifications and service integration [M9-M17, M34-M36]
Similarly to ontologies in other domains, the ontologies supporting ACE are expected to be continuously evolving artefacts. Therefore, ontology development and evaluation as well as the evaluation of dependent systems is a continuous and cyclic activity. This task will support the continuous assessment of ontologies, API specification and service orchestration through the lifetime of the project. This task will be carried out by QMUL, in collaboration with MTG-UPF and the industrial partners of the consortium.
Interdependencies: Tasks T2.2, T2.3, T2.4, T2.6, T*6.4
Expected results: Successful refinements of ontologies, APIs and services in ACE
Success criteria: Successful ontologies, APIs and services developed in ACE

Task 2.6: Integration of technology in content providers [M13-M15]
In this task, the initial content providers of the Audio Commons Ecosystem (Jamendo and Freesound), will integrate the services and technologies developed in this WP in order to connect to the Audio Commons Ecosystem and expose their content. This task will be carried out by the content providers of the Audio Commons consortium, MTG-UPF (Freesound) and Jamendo.
Interdependencies: Tasks T2.2, T2.3 and T2.4.
Expected results: Implementation of Audio Commons service orchestration technologies in Freesound and Jamendo.
Success criteria: At least 70% of Freesound and Jamendo content exposed in the ACE.

Deliverables

D2.1 Requirements report and use cases [M3]
A report outlining requirements in creative workflows of the target user communities to facilitate the use of audio commons content. A set of use cases against which our prototypes will be developed and evaluated. The report will also inform the development of the Audio Commons Ontology and API.

D2.2 Draft ontology specification [M9]
A draft first version of the Audio Commons Ontology implemented in OWL. This will facilitate the development of tools, APIs and service interoperability.

D2.3 Final ontology specification [M15]
The second version of the AC Ontology formalised in OWL. This version incorporates feedback from other technology development tasks in WP4 and WP5.

D2.8 Ontology evaluation report [M36]

The final version of the ontology will be evaluated at the end of the project with respect to deployment and success of adaptation within ACE as well as with respect to its logical properties. A report will inform future efforts to adopt and maintain the ontology.

D2.4 API specification [M15]
Specification of the end-user Web APIs and internal audio analysis and processing APIs.

D2.5 Service integration technologies [M15]
Software systems for enabling access to distributed repositories and services within ACE. This includes client side and server side software modules that can be integrated into existing or new products and services.

D2.6 Service integration draft guidelines [M15]
Draft guidelines for integrating ACE technologies. This will involve tutorial documents to help the adaptation of ACE technologies by content providers and users.

D2.7 Service integration guidelines [M36]
Final service integration guidelines that enable third party adaptation of ACE technologies, including the improvements that will have been introduced in the technologies.

WP3 - Intellectual property and business models						
Work package No.	WP3			**Starting date**		M1
Work package title	Intellectual property and business models					
Participant number	1	3	4			
Short name of participant	MTG-UPF	**Surrey**	Jamendo			
Person/months per participant	4	25	6			

Objectives

- Analysis of requirements for the rights management procedures involved in the ACE.
- Clarification of intellectual property aspects of the ACE. Documentation with different use cases that exemplify all possible rights and IP situations and provide solutions. Explain and analyse how the ACE fits and disrupts current creative industries IP models.
- Relate the previous objectives with the different Creative Commons licensing strategies available in the ACE.
- Define procedures that explain to content creators, content providers, tool developers and content users how to interact with the ACE (e.g., how to publish their content, how to consume it, how to license, i.e., how to become “Audio Commons Ready”).
- Research on emerging business models and long-term sustainability models for the ACE, including documentation on ways established audio/music industry entities can get involved in the ACE from a contributor/consumer perspective.

Description of work and role of partners
The ACE will present several opportunities and challenges for established gaming, music, and film/advertising industry entities. Commonly understood frameworks for publishing and master rights to particular audio and music recordings will be challenged within the ACE's framework, and this work package's remit will be to make those challenges understandable, and ultimately, useful for the industry. In the first instance, this will involve understanding the rights management requirements from the perspective of how audio is described and tagged at the point of creation and ingestion into the ACE. Then, use case analyses will be performed on the different Creative Commons licensing possibilities, and recommendations made as necessary. Finally, guidelines will be published in order for the creative industries to understand how new actors in the ACE can participate, and research into emerging business models possibly created by ACE interaction with publishers/creators/consumers carried out. **Task 3.1: Analysis of requirements for rights managements [M1-M6]** This task will investigate rights management requirement from an industry perspective, taking input from the project's commercial partner in the catalogue music-provision industry (Jamendo), and outlining use cases and rights management provisions that should be taken into account in the Audio Commons Ontology and API specification, in order to enable future developers and service consumers to be able to granularly specify rights uses cases and have taken care of from within ACE. This task will be carried out by Surrey in collaboration with Jamendo. Interdependencies: Task T2.1. Expected results: Deliverable D3.1. Success criteria: Satisfactory evaluation of all rights management cases within possible API use. **Task 3.2: Interpretation of Creative Commons licences in the audio domain [M1-M9]** In this task we will analyse different use cases for reusing Creative Commons content in the audio domain. Creative Commons licensing that allows for reuse, such as 3.0 Attribution Unported (CC-BY-3.0), or Attribution 4.0 International (CC-BY-4.0), might not specifically set out, expressively enough, how commercial use might be made and following which models in the complex reuse scenarios potentially arising from ACE. If necessary, we will make recommendations of modifications in the licenses to enable cross-subsidisation of business models from other industries to use and pay for audio and music from the Audio Commons. Initiatives like ccMixter have previously dealt with this by establishing their own extensions to CC-BY, such as CC-Plus, that allow them to set terms to potential commercial agreements. Based on these use cases, we will write documentation with recommendations to Audio Commons actors on how to use the licences, clarifying what content users need to do to "safely" reuse content from Audio Commons, and the process of licensing. This task will inform tools development on how to present licensing information to users (in terms of what data is needed and has to be displayed). This task will be carried out by Surrey, with input from Jamendo and MTG-UPF (content providers). Interdependencies: Task T2.1. Expected results: Deliverable D3.2. Success criteria: All possible use cases successfully handled with Creative Commons licenses. **Task 3.3: Exploration of business models for the ACE [M7-M18]** The Audio Commons Ecosystem and use cases explored within previous tasks, will create opportunity for research into emerging business models, in the interaction between creators, consumers, services, and creative platforms such as music production tools or game engines. This task will explore some of these directly, in order to create documentation that illustrates the potential business models and ways to implement them within industry. The investigation will particularly look for models that cross-subsidise the value of audio/music to emerging industry needs, such as health applications whereby the audio/music acts as agents for behaviour change (music for fitness apps or games for wellness), and where that behaviour affects health and provision of care or life insurance. This task will be carried out by Surrey.

Interdependencies: Task T3.4.
Expected results: Deliverable D3.4.
Success criteria: Identification of valid business models for the different perspectives from which actors can interact with the ACE.

Task 3.4: Guidelines for including new actors in the ACE [M10-M15]
In this task, use case comparisons to all current scenarios whereby a user or creator interacts with an available service such as Jamendo will be explored. Guidelines will be created for how new actors should be able to interact with the ACE, from the contributor and content provider level (creating and hosting music and/or audio to bring to the ACE) to the consumer level (buying and using ACE audio), to the tool creator level (writing plugins that are used within production workflows and that consume particular kinds of audio content), also including other levels of interaction in which only a particular service is offered such as audio content annotation or licensing management. This task will be carried out by Surrey, with input from Jamendo and MTG-UPF (content providers).
Interdependencies: Task T3.3.
Expected results: Deliverable D3.3.
Success criteria: Identification of flexible and modular interaction patterns with the ACE.

Task 3.5: Evaluation of ACE Implementation from a business model perspective [M31-M36]
In this task we will carry out an evaluation, from the business model perspective, of how ACE will be implemented across industry and through creative industry participants such as freelance composers, producers and audio designers, as well as game engine makers and their stores. This evaluation will report on possible commercialisation and monetisation pathways for creative content to be sold and consumed as well as reused within the ACE. This task will be carried out by Surrey.
Interdependencies: Tasks T3.3, T3.4 and T7.3.
Expected results: Deliverable D3.5.
Success criteria: Favorable evaluation of the ACE implementation.

Deliverables

D3.1 Report on Rights Management requirements [M6]
Release of a document outlining rights management requirements for the ACE and the Audio Commons Ontology in particular.

D3.2 Report on usage of Creative Commons licences [M9]
Release of a document describing use cases of the Creative Commons licenses in the Audio Commons Ecosystem and identifying potential shortcomings (with proposals for improvements if required). Guidelines on how to present licensing information to users.

D3.3 Guidelines for including new actors in the ACE [M15]
Document describing the guidelines for interacting with the Audio Commons Ecosystem, from the contributor and content provider level to the consumer level and to the tool creator level.

D3.4 Report on business models emerging from the ACE [M18]
Report on potential business models that can emerge from the Audio Commons Ecosystem.

D3.5 Report on the evaluation of the ACE from a business model perspective [M36]
Evaluation report of the implementation of the Audio Commons Ecosystem carried out during the project, from a business model perspective.

WP4 - Semantic annotation of musical sound properties						
Work package No.	WP4			**Starting date**		M4
Work package title	Semantic annotation of musical sound properties					
Participant number	1	2	4			
Short name of participant	**MTG-UPF**	QMUL	Jamendo			
Person/months per participant	35	15	6			

Objectives

- Research, development and evaluation of automatic methods for semantically annotating musical properties of audio content (both at the piece level and at the sample level).
- Research, development and evaluation of an interface for manually annotating musical properties of audio content (both at the piece level and at the sample level).
- Release of the aforementioned tools and integration in Audio Commons content providers (Freesound and Jamendo).

Description of work and role of partners

In WP4 we will carry out research and evaluation of semantic annotation of musical properties of music samples and music pieces. Music samples can be used in music production contexts to, for example, build a sampler application using recordings of isolated instrument notes, or to create a music composition by overlaying different loops. Music pieces can be used in a creative context as, for example, the soundtrack of a video or as remixable material for a music producer. This Work Package aims to enhance the usefulness of existing content and to facilitate more creative uses by: (i) developing a system to automatically add musically-meaningful metadata to existing content; and (ii) developing better tools for manually annotating music samples and music pieces. The work in WP4 will be divided in the following tasks:

Task 4.1: Compilation of state-of-the-art methods for the automatic annotation of musical content [M4-M6]
In this task consists we will review current state of the art techniques for the automatic annotation of music pieces and music samples. Traditionally, the Music Information Retrieval field has been focused on the analysis of music pieces based on audio features [Peeters03], hence the main focus of this task will be the compilation and comparison of these existing methods [e.g., Klapuri06, Paulus10]. The selection of the methods to compare (i.e., the different annotation tasks) will be carried out according to an estimation of what audio properties could be relevant for the audio representation of the Audio Commons Ontology. Nevertheless, part of the focus will be assigned to collecting existing methods for audio annotation that could be suitable for the annotation of the different kinds of music samples [e.g., Gomez06]. It is possible that, as part of this task, some relevant methods need to be implemented from scratch to be usable in later stages of the project. This task will be carried out by MTG-UPF.
Interdependencies: (none)
Expected results: Deliverable D4.1.

Success criteria: Availability of the aforementioned executables to facilitate research in subsequent tasks.

Task 4.2: RTD on automatic methods for the semantic annotation of music samples [M7-M15, M19-M30, M34-M36]

As mentioned before, current MIR techniques are focused on the analysis and characterisation of music pieces. The highest innovation focus of this work package will therefore be on the semantic annotation of music samples. Music samples feature significantly different characteristics than music pieces, and hence their characterisation requires a different approach than that of music pieces. For simpler music samples we will focus on semantically annotating musical properties such as the envelope, the particular note being played in a recording, or the instrument that plays that note. The particular audio properties that we will analyse will depend on the actual definition of the Audio Commons ontology and will require the development of new audio features. For more complex music samples such as loops, we will focus on semantic annotations that will be useful on a creative context such as the Bpm or the exact start and end points for a perfect looping, as well as other annotations such as genre or mood which are also applicable to music pieces (see below). To evaluate the semantic annotation of music samples, we will use data from Freesound and from other open music sample libraries such as the Modular Samples library[31]. This includes the construction of datasets to train the models required to characterise the targeted audio properties. The output of the annotation methods will consist of semantic descriptions represented in standard semantic web formats and following the Audio Commons Ontology definition.

The methods developed in this task will be evaluated in Task 4.4, and a second research and development cycle will take place to improve the technology taking into account the input from the evaluations (D4.4, D4.10) and, if relevant, from the first evaluation of the embeddable tools developed in WP6 (#D6.4). This task will be carried out by MTG-UPF, with the collaboration of QMUL.

Interdependencies: Tasks T2.1, T2.2, T2.4, T4.1, T4.3 and T6.4.

Expected results: Deliverables D4.2, D4.7 and D4.12.

Success criteria: Tools being released in their prototype and final states.

Task 4.3: RTD on automatic methods for the semantic annotation of music pieces [M7-M15, M19-M30, M34-M36]

In this task, and building on top of the methods reported in D4.1, we focus on researching about the automatic semantic annotation of musical pieces. We will advance current state-of-the-art in this field by training models for predicting musical aspects such as genre, mood or instrument recognition using big datasets (typically bigger than those used in MIR [Serra13]), consisting of data coming from Jamendo, AcousticBrainz, and the Computational Analysis of the Live Music Archive of the Internet Archive (CALMA dataset [Wilmering15]). In order to train the models required to characterise the targeted music properties, parts of these datasets will need to be annotated. We will also focus on other relevant music piece characterisations such as Bpm, tonality or structure. The specific selection of audio properties to include in the semantic annotation will depend on the requirements of the Audio Commons Ontology. The output of the annotation methods will consist of semantic descriptions represented in standard semantic web formats and following the Audio Commons Ontology definition. This task is, in fact, well aligned with the current MIR challenges listed in the MIReS roadmap [Serra13], a document describing guidelines for future MIR research.

The methods developed in this task will be evaluated in Task 4.5, and a second research and development cycle will take place to improve the technology taking into account the input from the evaluations (D4.5, D4.11) and, if relevant, from the first evaluation of the embeddable tools developed in WP6 (#D6.5). This task will be carried out by QMUL with the collaboration of MTG-UPF, and with the contribution of Jamendo for the construction of music datasets.

Interdependencies: Tasks T2.1, T2.2, T2.4, T4.1, T4.4 and T6.5.

Expected results: Deliverables D4.3, D4.8 and D4.13.

[31] http://modularsamples.com

Success criteria: Tools being released in their prototype and final states.

Task 4.4: Evaluation of automatic methods for the semantic annotation of music samples [M16-M18, M31-M33]
This task will consist in the evaluation of the automatic annotation tools developed in Task 4.2. Evaluation will be carried out following strict scientific standards, with thorough statistical analysis of methods' accuracy under different circumstances and including the participation of human evaluations (if required) via the design of listening tests. This task will be carried out by MTG-UPF in collaboration with QMUL.
Interdependencies: Task T4.2.
Expected results: Deliverables D4.4 and D4.10.
Success criteria: Satisfactory accuracy of the evaluated methods.

Task 4.5: Evaluation of automatic methods for the semantic annotation of music pieces [M16-M18, M31-M33
This task will consist in the evaluation of the automatic annotation tools developed in Task 4.3. Evaluation will be carried out following strict scientific standards, with thorough statistical analysis of methods' accuracy under different circumstances and including the participation of human evaluations (if required) via the design of listening tests. This task will be carried out by QMUL in collaboration with MTG-UPF.
Interdependencies: Task T4.3.
Expected results: Deliverables D4.5 and D4.11.
Success criteria: Satisfactory accuracy of the evaluated methods.

Task 4.6: Research, development and evaluation of an interface for manually annotating musical content [M13-M20, M31-M33]
Besides the automatic annotation of music samples and music pieces that is addressed in previous tasks, in this task of WP4 we will focus on the manual annotation of music samples and music pieces. In particular, we will develop a web-based interface that intelligently guides users on the annotation process of music samples and music pieces, and maximises the quality and usefulness of the annotations according to the guidelines of the Audio Commons ontology, making it also suitable as ground truth data [Font14]. That interface will also take advantage of the aforementioned automatic annotation tools to pre-fill some of the aspects that need to be annotated. For example, for note recordings the automatic tools can be used to estimate the note name. The interface developed in this task is mainly thought for being used during the process of publishing audio content in the Audio Commons Ecosystem (e.g., when content creators upload content to Freesound or Jamendo), but will also be usable as a post-annotation step in which users of a platform can collaboratively contribute to the annotation of content and ground truth generation. The definition and implementation of the tools developed in this task will be opened so that future content providers (besides Freesound and Jamendo) and AudioCommons publishing tools can easily adopt them.
This task will include an evaluation of the interface in terms of its usability and its expressive power for annotating music samples and music pieces. That evaluation will be carried out with real users and in combination with the evaluation of Task 5.4. This task will be carried out by MTG-UPF.
Interdependencies: Tasks T4.2, T4.3.
Expected results: Deliverables D4.6 and D4.9.
Success criteria: Satisfactory evaluation of the annotation interface.

Task 4.7: Integration of annotation technology in content providers [M13-M15, M29-30, M34-M36]
In this task, the initial content providers of the Audio Commons Ecosystem (Jamendo and Freesound), will integrate the annotation methods developed in this WP. This task will be carried out by the content providers of AudioCommons consortium, MTG-UPF (Freesound) and Jamendo.
Interdependencies: T4.2, T4.3 and T4.6.
Expected results: Implementation of Audio Commons annotation technologies in Freesound and Jamendo.

Success criteria: At least 70% of Freesound (musical content) and Jamendo content annotated with Audio Commons metadata as defined in the Audio Commons Ontology.

Deliverables

D4.1 Report on the analysis and compilation of state-of-the-art methods for the automatic annotation of music pieces and music samples [M6]
This deliverable will include a document listing existing methods for several audio characterisation tasks (related to music pieces and music samples) along with pointers into already existing implementations or specifically developed implementations that can be used in AudioCommons.

D4.2 First prototype tool for the automatic semantic description of music samples [M15]
Release of the first prototype of the software tool for the automatic annotation of music samples.

D4.4 Evaluation report on the first prototype tool for the automatic semantic description of music samples [M18]
Report on analysis results of the accuracy and effectiveness of the first prototype for the automatic annotation of music samples. Identification of key areas for improvement/extension.

D4.7 Second prototype tool for the automatic semantic description of music samples [M30]
Release of the second prototype of the software tool for the automatic annotation of music samples. This version of the tools contains improvements included after the first evaluation cycle.

D4.10 Evaluation report on the second prototype tool for the automatic semantic description of music samples [M33]
Report on analysis results of the accuracy and effectiveness of the second prototype for the automatic annotation of music samples. Identification of key areas for improvement/extension.

D4.12 Release of tool for the automatic semantic description of music samples [M36]
Final release of the software tool for automatically annotating music samples using semantic concepts from the Audio Commons Ontology. This version of the tool contains the improvements included after the second evaluation cycle.

D4.3 First prototype tool for the automatic semantic description of music pieces [M15]
Release of the first prototype of the software tool for the automatic annotation of music pieces.

D4.5 Evaluation report on the first prototype tool for the automatic semantic description of music pieces [M18]
Report on analysis results of the accuracy and effectiveness of the first prototype for the automatic annotation of music piece. Identification of key areas for improvement/extension.

D4.8 Second prototype tool for the automatic semantic description of music pieces [M30]
Release of the second prototype of the software tool for the automatic annotation of music pieces. This version of the tools contains improvements included after the first evaluation cycle.

D4.11 Evaluation report on the second prototype tool for the automatic semantic description of music pieces [M33]
Report on analysis results of the accuracy and effectiveness of the second prototype for the automatic annotation of music pieces. Identification of key areas for timbral analysis improvement/extension.

D4.13 Release of tool for the automatic semantic description of music pieces [M36]
Final release of the software tool for automatically annotating music pieces using semantic concepts from

the Audio Commons Ontology. This version of the tool contains the improvements included after the second evaluation cycle.

D4.6 Release of tool for the manual annotation of musical content [M20]
Release of the tool for manually annotating music samples and music pieces following the guidelines of the Audio Commons Ontology.

D4.9 Evaluation report on the tool for manual annotation of musical content [M30]
Evaluation report on the tool for manually annotating music samples and music pieces following the guidelines of the Audio Commons Ontology.

WP5 - Semantic annotation of non-musical sound properties						
Work package No.	WP5			**Starting date**		M6
Work package title	Semantic annotation of non-musical sound properties					
Participant number	1	3				
Short name of participant	MTG-UPF	**Surrey**				
Person/months per participant	15	35				

Objectives

- Determine the categories of timbral attributes most useful to automatic characterisation of sound library elements.
- Research, development and evaluation of automatic methods for semantically annotating non musical sounds
- Research, development and evaluation of an interface for manually annotating non musical sounds (i.e., sound effects and soundscapes, and possibly musical instruments in a non-musical context)
- Release of the aforementioned tools and integration in Audio Commons content providers that feature non-musical content (Freesound).

Description of work and role of partners

In WP5 we will carry out the research and evaluation of semantic annotation of non-musical sounds. Non-musical sounds can be used, for example, in creative contexts such as audiovisual production or video games development to create the non-musical soundtrack of a game or movie. Typically, much sound library content is annotated with tags according to their sound source/object (scream, car-crash, orchestra, etc.). This Work Package aims to enhance the usefulness of existing content and to facilitate more creative uses by: (i) developing better tools for manually annotating sound effects and soundscapes; and (ii) developing a system to automatically add timbral metadata, such that content can (additionally, or instead) be searched by perceptual sound quality (piercing, crunchy, rich, etc.). The work in WP5 will be divided in the following tasks:

Task 5.1 Profiling of timbral metadata currently in use to describe sound and with potential to add significant value [M6-M7]
In this task, timbral metadata in existing content from Freesound and potentially other sources will be gathered, and supplemented with descriptors from verbal elicitation experiments using common content types. Verbal profiling analysis will then be employed to sort and prioritise terms according to meaning and use frequencies. This will provide a hierarchical ontology of timbral descriptors with potential to add significant value to audio content, which will feed the definition of the Audio Commons Ontology. This task will be carried out by Surrey, with the collaboration of MTG-UPF for gathering existing data from Freesound.
Interdependencies: (none)
Expected results: Deliverable D5.1.
Success criteria: Meaningful ontology with terms corresponding to timbral properties useful to annotate audio content.

Task 5.2 RTD on automatic timbral characterisation methods for the semantic annotation of non-musical content [M8-M15, M19-M30, M34-M35]
In this task, and following the hierarchical ontology of timbral descriptors (incorporated in the Audio Commons Ontology), we will build a tools for automatically annotating non-musical audio content. Timbral metrics and perceptual models exist in many different forms for a variety of timbral attributes [e.g., Lakatos00, Terasawa05, Caclin05, Zwicker07]. In a first phase of this task, these existing metrics and models will be surveyed to determine the best existing approaches and to formulate a system structure for comprehensive timbral characterisation. Existing metrics and models will then be built into the formulated system, and optimised for consistency and efficiency, to provide a prototype capable of rating a wide range of timbral attributes and generating metadata accordingly.
On a second phase of the task, timbral perception literature and further listening experiments, combined with objective signal analysis, will inform the specification of required enhancements to existing metrics, and of modelling approaches for significant timbral attributes not covered by the prototype system. These enhancements and new models will be implemented for cases where the associated feature extraction can be accomplished using existing signal processing techniques, to provide an improved timbral characterisation system. In some situations, existing techniques may not be effective to extract the necessary acoustic features required by the enhancements and new models specified in this task. For these cases, novel signal processing techniques will be developed, potentially involving elements of source separation for content comprising multiple elements, and sparse representation for efficient feature extraction. These novel techniques will feed into the further development of timbral perception models.
The methods developed in this task will be evaluated in Task 5.3, and a second research and development cycle will take place to improve the technology taking into account the input from the evaluations (D5.3, D5.7) and, if relevant, from the first evaluation of the embeddable tools developed in WP6 (D6.6). This task will be carried out by Surrey.
Interdependencies: Tasks T2.1, T2.2, T2.4, T5.1, T5.3 and T6.6.
Expected results: Deliverables D5.2, D5.6, D5.8.
Success criteria: Tools being released in their prototype and final states.

Task 5.3 Evaluation of automatic methods for the semantic annotation of non-musical content [M16-M18, M31-M33]
In this task, the prototypes developed in Task 5.2 (D5.2, D5.6) will be evaluated against the timbral descriptor hierarchy defined in Task 5.1, and validated using listening tests where appropriate, in order to identify key areas for improvement (where existing approaches do not adequately predict listener perception) and the most significant gaps in capabilities (where descriptors higher in the hierarchy are not yet represented). This task will be carried out by Surrey.
Interdependencies: Task T5.1 and T5.2.

Expected results: Deliverables D5.3 and D5.7.
Success criteria: Satisfactory accuracy of the evaluated methods.

Task 5.4 Research, development and evaluation of an interface for manually annotating non-musical content [M21-M30]
In close relation with the Task 4.6 of WP4, in this task we will work on an intelligent interface for manually annotating non-musical audio content.. The interface will focus on standard important properties of non-musical audio content (e.g., sound effects and soundscapes) such as the sound sources present in a recording, but will also focus on helping users annotate the timbral terms defined in Task 5.1. The interface developed in this task will be mainly thought for being used during the process of publishing audio content in the Audio Commons Ecosystem, but will also be usable as a post-annotation step in which users of a platform could collaboratively contribute to the annotation of content and ground truth generation. The definition and implementation of these tools will be opened so that future content providers (besides Freesound) and AudioCommons publishing tools can easily adopt them. Also as a part of this task, the manual annotation tools developed in Task 4.6 will be integrated with these developed in this task, and will be evaluated in combination. This task will be carried out by MTG-UPF.
Interdependencies: Tasks T4.6, T5.1 and T5.2.
Expected results: Deliverables D5.4 and D5.5.
Success criteria: Satisfactory evaluation of the annotation interface.

Task 5.5: Integration of annotation technology in content providers [M13-M15, M29-30, M34-M36]
In this task, the initial content providers of the Audio Commons Ecosystem that have non-musical content (Freesound), will integrate the annotation methods developed in this WP. This task will hence be carried out by MTG-UPF (Freesound).
Interdependencies: Tasks T5.2 and T5.4.
Expected results: Implementation of Audio Commons annotation technologies in Freesound.
Success criteria: At least 70% of Freesound (non-musical) content annotated with Audio Commons metadata as defined in the Audio Commons Ontology.

Deliverables

D5.1 Hierarchical ontology of timbral semantic descriptors [M7]
Hierarchical ontology of timbral semantic descriptors with categories of timbral attributes most useful to automatic characterisation of non-musical content in sound library elements.

D5.2 First prototype of timbral characterisation tool for semantically annotating non-musical content [M15]
Release of the first prototype of the software tool for the automatic annotation of non-musical audio content.

D5.3 Evaluation report on the first prototype of the timbral characterisation tool [M18]
Report on analysis results of the accuracy and effectiveness of the first prototype for the automatic annotation of non-musical audio content. Identification of key areas for improvement/extension.

D5.6 Second prototype of timbral characterisation tool for semantically annotating non-musical content [M30]
Release of the second prototype of the software tool for the automatic annotation of non-musical audio content. This version of the tools contains improvements included after the first evaluation cycle.

D5.7 Evaluation report on the second prototype of the timbral characterisation tool [M33]

Report on analysis results of the accuracy and effectiveness of the second prototype for the automatic annotation of non-musical audio content. Identification of key areas for improvement/extension.

D5.8 Release of timbral characterisation tool for semantically annotating non-musical content [M35]
Final release of the software tool for automatically annotating non-musical audio content using semantic concepts from the Audio Commons Ontology. This version of the tool contains the improvements included after the second evaluation cycle.

D5.4 Release of tool for the manual annotation of non-musical content [M27]
Release of the tool for manually annotating non-musical content following to the guidelines of the Audio Commons Ontology.

D5.5 Evaluation report on the tool for manual annotation of non-musical content [M30]
Evaluation report on the tool for manually annotating non-musical content following the guidelines of the Audio Commons Ontology.

WP6 - Prototyping and evaluation in production workflows						
Work package No.	WP6			**Starting date**		M7
Work package title	Prototyping and evaluation in production workflows					
Participant number	2	4	5	6		
Short name of participant	**QMUL**	Jamendo	AudioGaming	Waves		
Person/months per participant	20	12	24	24		

Objectives

- Development of embeddable tools to access different kinds of Audio Commons content (music pieces, music samples and non-musical content) within different production workflows (music production, sound design, video games and audiovisual production…).
- Evaluation of the developed tools with content users, to provide feedback and further improve the tools themselves and the technology behind them.

Description of work and role of partners

This work package will be focused on the development and evaluation of tools to access Audio Commons content and to be embedded in production workflows of the creative industries, and in the evaluation of the ACE as a whole. Hence, this work package enables (i) testing the system as a whole by building a series of prototypes that rely on ACE and measure their impact on specific creative dimensions, (ii)

evaluate the prototypes from the perspective of users in specific creative domains. The industrial partners of AudioCommons will be responsible for the development of the tools, and QMUL will lead the evaluation activities and coordinate the industrial partners.

Task 6.1: Development of a tool for the reuse of Audio Commons content in music production workflows [M7-15, M22-30]

This task will enable the integration of ACE into music production workflows. We will research and develop a sampler plugin for Digital Audio Workstations (DAW) that uses Audio Commons content, and add-ons for Ardour[32], a well-known open-source DAW, and TracksLive, a multitrack recording software for live performance capture. These will connect the end user and music professional to the Audio Commons Ecosystem. This connection will allow the user to locate, access and obtain the license for the content he needs, such as music samples available on Freesound and delivered through the Audio Commons API**.** This task will be carried out by Waves.

Interdependencies: T2.4, T6.4, T6.5.

Expected results: Deliverables D6.1 and D6.9.

Success criteria: Tool effectively accessing Audio Commons content and successfully being evaluated in Task 6.5.

Task 6.2: Development of a web interface for accessing Audio Commons music pieces [M7-15, M22-30]

In this task, we will develop a prototype for navigating and exploring music tracks found in the Audio Commons Ecosystem. The prototype will consist of a web interface that will allow users to navigate music tracks in advanced ways according to the semantic music description of music pieces generated by the automatic annotation tools from WP4. This task will be carried out by Jamendo.

Interdependencies: T2.4, T6.4, T6.6.

Expected results: Deliverables D6.2 and D6.10.

Success criteria: Tool effectively accessing Audio Commons content and successfully being evaluated in Task 6.6.

Task 6.3: Development of a tool for the reuse of Audio Commons in sound design tasks for video game and audiovisual production workflows [M7-15, M22-30]

In this task, we will research and develop a sound design plug-in with a direct interfacing with Audio Commons non-musical content. This plugin will allow end users to efficiently get access to and upload sounds in the plug-in along with its licence. They will then be able to use the sound or part of the sound to create new sound effects by layering them with other sounds or transforming them in real-time using specific effects developed by AudioGaming. This plug-in will be compatible with most of today's industry standard software. AudioGaming will put a particular focus on interactivity. As such, the tools will emphasize as much as possible real-time interactions and will provide exports for most major audio and interactive audio formats and software. This task will be carried out by AudioGaming.

Interdependencies: T2.4, T6.4, T6.7.

Expected results: Deliverables D6.3 and D6.11.

Success criteria: Tool effectively accessing Audio Commons content and successfully being evaluated in Task 6.7.

[32] www.ardour.org

Task 6.4 Holistic Evaluation of the ACE in the context of Creativity Support [M13-M36]
We will develop novel methods to assess how our system and specific tools facilitate the *creative flow* [Csikszentmihalyi97], discovery, innovation [Schneidermann07] and other relevant dimensions of creative work. New methodologies will be devised to evaluate ACE as a whole from the perspective of the social implications of audio content reuse and its transformational character on content creation, complementing T3.5 focussing on the business perspectives. We will also measure the extent to which ACE enables new activities. The task will devise novel methods to evaluate specific tools relevant in three specific use cases addressed by T6.1-6.3. The work will rely on and extend factor evaluation [Carroll09] and cognitive theories of creativity support [Davis13], considering user behaviour metrics to learn about the creative process. This will inform the specific evaluation tasks T6.5-6.7 to ensure tool specific evaluations allow for drawing general conclusions. This task will be carried out by QMUL with the collaboration of the industrial partners of the consortium.
Interdependencies: T3.5, T6.1, T6.2, T6.3.
Expected results: Deliverables D6.7, D6.8, D6.12.
Success criteria: Successful evaluation of ACE and specific tools.

Task 6.5: Evaluation of ACE in music production [M16-M21]
This task will measure the utilities of ACE in typical music production workflows such as studio-based composition and music post-production where music samples are typically used to enhance recorded material. The evaluation will focus on how ACE enables new creative activities in music production and the utilisation of AC content in this scenario. The task will involve beta testers available from Waves and students of Queen Mary's Media and Arts Technology (MAT) programme studying music production. This task will be carried out by QMUL in collaboration with Waves.
Interdependencies: T6.1, T6.4.
Expected results: Deliverable D6.4.
Success criteria: Favorable evaluation of the prototype and relevant feedback obtained.

Task 6.6: Evaluation of search and retrieval interfaces for accessing Audio Commons music pieces [M16-M21]
This task will focus on the evaluation of ACE in different creative dimensions, primarily focussing on exploration, engagement and expressiveness [Carroll09]. We will assess how ACE supports information seeking activities in creative music production using the web-based interfaces created in T6.2. This task will be carried out by QMUL in collaboration with Jamendo.
Interdependencies: T6.2, T6.4.
Expected results: Deliverable D6.5.
Success criteria: Favorable evaluation of the prototype and relevant feedback obtained.

Task 6.7: Evaluation of ACE in sound design and audiovisual production [M16-M21]
In this task, ACE will be evaluated in the context of audio as a functional tool in software, audio-visual production, sound design and potential content reuse by autonomous systems such as video games using procedural audio. We will assess how ACE allows for new creative output by reuse and transformation of AC content, how collaboration is facilitated by ACE between content users and content creators. The task will involve beta testers available from AudioGaming and students from Surrey's Film and Video Production Engineering BA (Hons). This task will be carried out by QMUL in collaboration with

AudioGaming.
Interdependencies: T6.3, T6.4.
Expected results: Deliverable D6.6.
Success criteria: Favorable evaluation of the prototype and relevant feedback obtained.

Deliverables
D6.1 Prototype of an embeddable tool for integrating Audio Commons music samples [M15] Prototype of the DAW plugin and add-on for accessing Audio Commons music samples. D6.4 Evaluation report on the prototype of an embeddable tool for integrating Audio Commons music samples [M21] Report on the evaluation of the prototype of DAW plug-in and add-on for accessing Audio Commons music samples. Identification of key areas for improvement/extension. D6.9 Release of the embeddable tool for integrating Audio Commons music samples [M30] Release of the DAW plugin and add-on for accessing Audio Commons music samples. This version includes enhancements proposed by the evaluation carried out in Task 6.4. D6.2 Prototype of a web interface for accessing Audio Commons music pieces [M15] Prototype of the web interface accessing Audio Commons music pieces. D6.5 Evaluation report on the prototype of a web interface for accessing Audio Commons music pieces [M21] Report on the evaluation of the prototype of web interface for accessing Audio Commons music pieces. Identification of key areas for improvement/extension. D6.10 Release of the web interface for accessing Audio Commons music pieces [M30] Release of the web interface for accessing Audio Commons music samples. This version includes enhancements proposed by the evaluation carried out in Task 6.5. D6.3 Prototype of an embeddable tool for integrating non-musical Audio Commons content [M15] Prototype of the plugin for accessing non-musical Audio Commons content. D6.6 Evaluation report on the prototype of an embeddable tool for integrating non-musical Audio Commons content [M21] Report on the evaluation of the plugin for accessing non-musical Audio Commons content. Identification of key areas for improvement/extension. D6.11 Release of the embeddable tool for integrating non-musical Audio Commons content [M30] Release of the plugin for accessing non-musical Audio Commons content. This version includes enhancements proposed by the evaluation carried out in Task 6.6. D6.7 Guidelines for second phase implementation [M21] We will start by assessing state of the art in measuring the success of creativity support systems. Then, we will address shortcomings with novel designs and issue a set of guidelines helping the implementation of tools and enable the collection of relevant data. D6.8 Report on novel methods for measuring creativity support [M30] This report will detail our methods for the assessment of creativity support in ACE and summarise our findings on the performance of our tools.

D6.12 Report on the evaluation of the ACE from a holistic and technological perspective [M36]
Report on the evaluation of our system assessing how our tools support creative work in specific domains and how the system facilitates the reuse of audio content as a whole.

WP7 - Dissemination and exploitation						
Work package No.	WP7			**Starting date**		M1
Work package title	Dissemination and exploitation					
Participant number	1	2	3	4	5	6
Short name of participant	**MTG-UPF**	QMUL	Surrey	Jamendo	AudioGaming	Waves
Person/months per participant	12	6	6	4	4	4

Objectives

- Contribute to the achievement of all the project objectives and to ensure the long term impact of the project by disseminating and exploiting results to the industry, research community and general public.
- Coordination with other external initiatives related to the project
- Defining IPR strategies suitable for maximising the impact of the project.
- Spread the word about the Audio Commons Ecosystem.

Description of work and role of partners

This work package will address the dissemination of the project's results on various levels, ensuring the proper outreach of AudioCommons activities to all relevant stakeholders in the creative industries (content users and content creators, potential new content providers and tool developers) and the scientific community. The work package will also include an exploitation and sustainability plan of the Audio Commons Ecosystem beyond the duration of the project, including the different business models that can arise in the ecosystem as analysed in WP3.

Task 7.1: General project dissemination [M1-M36]
Visual design definition, implementation and maintenance of the web portal, including two years beyond the end of the project. This task will be carried out by the MTG-UPF, with contributions from other partners in the form of content (blog posts, etc…) for the web.
Interdependencies: (none)
Expected results: Deliverables D7.1 D7.2, D7.3 and D7.7.
Success criteria: 50 daily unique visitors to the AudioCommons web portal, (excluding bots), this value

being increased by at least 50% during time periods influenced by AudioCommons events.

Task 7.2: Alignment with other related initiatives at EC level and contribution to standards [M1-M36]
AudioCommons partners will be in contact with the most relevant initiatives related to the project interests (such as the Europeana Sounds project and, in particular, the Europeana Data Model profile for Sound), inviting representatives to join to some of the AudioCommons technical sessions, as well as the attendance to the events to be organised. This task will be carried out by MTG-UPF with the collaboration of other partners when required.
Interdependencies: (none)
Expected results: Contribution and collaboration with other initiatives or standards such as Europeana data model for sounds.
Success criteria: Contribution with at least three (3) related initiatives.

Task 7.3: Exploitation and sustainability plan, including IP management of AudioCommons technologies and tools [M6-M36]
The industrial partners will monitor the development of the market and activities in audience experience. We will prepare the Implementation and Use of Knowledge plans for the project, and plan the Exploitation taking into account the market reality and the inputs of the emerging business models researched in WP3. The partners will review the IP produced by the research teams to determine which items can be protected, initiating the protection process where appropriate, reviewing and clearing material for publication (to ensure there are no premature disclosures that academics can publish in a timely manner), and reaching agreements for the commercial exploitation of knowledge, since some of the technologies to be used are background of project partners. This task will be carried out by all partners in the consortium, leaded by Surrey.
Interdependencies: Task T3.3.
Expected results: Deliverables D7.4, D7.5 and D7.6.
Success criteria: Roadmap for proper future exploitation available, with stakeholders potential risks and bottlenecks identified and contingency plan.

Task 7.4: Academic dissemination of RTD outcomes [M13-M36]
Research partners will take part in dissemination to the research community through meetings, showcases, specially-organised conference sessions and publication of results in reviewed journals. This task will be carried out by the three research partners in the consortium, MTG-UPF, QMUL and Surrey.
Interdependencies: (none)
Expected results: Research publications in relevant conferences and journals.
Success criteria: Fifteen (15) publications.

Task 7.5: Dissemination, promotion and demonstration of the AudioCommons outcomes within the creative community [M25-M36]
This activity aims at widely disseminating and promoting the Audio Commons tools for publishing and retrieving Audio Commons content among the creative community through the active participation in key events and direct communication with key actors in the field. In addition, all software tools, APIs and datasets to be obtained will be presented to the industry and tested by the developer community through the Music Hack Day events. This task will be carried out by the industrial partners in the consortium, Jamendo, Waves, and AudioGaming, and with the collaboration of MTG-UPF.
Interdependencies: (none)
Expected results: Demonstration of the Audio Commons tools and products at relevant venues.
Success criteria: Five (5) media featuring the project, attendance to five (5) relevant events.

Deliverables

D7.1 Project Website [M3]
A multimedia website with public information on the project and its evolution, together with a private space for partners only as a discussion space and document repository.

D7.2 Visual identity of Audio Commons tools and ecosystem [M3]
Visual identity for the different parts and tools of the Audio Common ecosystem.

D7.4 Draft exploitation and sustainability plan [M12]
Draft of the exploitation and sustainability plan according to academic and industrial partner's perspectives.

D7.5 Second draft exploitation and sustainability plan [M24]
Second draft of the exploitation and sustainability plan according to academic and industrial partner's perspectives.

D7.6 Final exploitation and sustainability plan [M36]
Exploitation and sustainability plan according to academic and industrial partner's perspectives.

D7.3 Dissemination Plan [M12]
Dissemination plan covering the dissemination of AudioCommons RTD outcomes. This deliverable is transversal to several tasks of this work package.

D7.7 Report on dissemination and publication of results [M36]
Publication of results will be produced including the publications, conferences and trade shows attendances. This deliverable will summarise all dissemination activities carried out during the project, as well as AudioCommons contribution to other initiatives. This deliverable is transversal to several tasks of this work package.

Table 3.2: List of milestones

No	Name	Related WPs	Date	Leader	Means of verification
MS1	Intellectual property documents and early Ontology definition	WP2, WP3	M9	MTG-UPF	Intellectual property requirements defined, draft of the Audio Commons Ontology including intellectual property information and audio properties as estimated by the research and industrial partners of the consortium; early work on the research for semantic annotation of musical and non-musical audio content (first implementations using existing or simplified techniques).
MS2	1st ACE prototype	WP2-6	M15	MTG-UPF	Laboratory prototype of the Audio Commons Ecosystem, including integration with content providers and development of embeddable tools; Audio Commons Ontology and API specification delivered; first research and development cycle for semantic annotation technologies (not yet validated); exploration of potential business models for the Audio Commons Ecosystem.

MS3	2nd ACE prototype	WP2-6	M30	MTG-UPF	Laboratory prototype of the Audio Commons Ecosystem including the second iteration of the semantic annotation technologies and embeddable tools (validated in the laboratory by user groups and with standard MIR evaluation methodologies); academic dissemination of the work and early non-academic dissemination.
MS4	Release of ACE	WP2-7	M36	MTG-UPF	Final prototype if the Audio Commons Ecosystem including the third iteration of the semantic annotation technologies and embeddable tools (validated again in the laboratory by user groups and with standard MIR evaluation methodologies); academic and non-academic dissemination of the work; evaluation of the Audio commons Ecosystem from the business model perspective; sustainability plans and documents for facilitating adoption of Audio Commons technologies and inclusion to the ecosystem for potential new actors.

3.2 Management structure and procedures

Project management and coordination is active during the whole life-cycle of the project and is devoted to the effective coordination and management of the project activities. In particular, it ensures effective cooperation between partners, supervises management reports and organises reviews, meetings and public events.

The organisational structure of the AudioCommons project is composed of the Project Steering Committee (PSC), the Project Coordinator (PC) and the Technical Committee (TC). This structure is suitable for the project size, previous experience of partners in European Projects and has already been implemented with success in similar projects.

The **Project Steering Committee (PSC)** will ensure that the Consortium fulfills all its contractual obligations. The members of this board have enough seniority to take binding decisions on behalf of their entities (also in terms of budget or level of disclosure of the research data and findings). The PSC meets at least once a year, while extra meetings might be convened if needed. The PC chairs the PSC meetings. Decisions at PSC level are made on the basis of consensus; however, if this is not possible, they are made on the basis of a majority vote, with the PC having the casting vote. Each representative of the PSC is responsible for the internal project activities within his/her institution.

The role of the key persons inside the PSC is as follows:

Role	Partner nr	Partner short name	Key person
Chair	1	MTG-UPF	Xavier Serra
Secretary	1	MTG-UPF	Project Management Office
Member	2	QMUL	George Fazekas

Member	3	Surrey	Mark Plumbley
Member	4	Jamendo	Martin Guerber
Member	5	AudioGaming	Amaury La Burthe
Member	6	Waves	Meir Shaashua

The PC is led by Prof. Dr. Xavier Serra and supported by the management team at the Music Technology Group of UPF. The PC plays a crucial role in the design and monitoring of the project evolution and achievements in accordance with the objectives and risks identified throughout its complete duration, and in the implementation of contingency plans if required.

More specifically, the PC is responsible for the following tasks: establishing and maintaining efficient communication within the Consortium, monitoring project progress according to the work plan, time schedule and resources as established in DoA (identifying possible deviations) and reporting to the European Commission, resolving any potential conflicts, coordinating and leading the PSC meetings and monitoring risk elements.

MTG-UPF has extensive experience in participating in European projects since the 4th Framework Programme as partner but also coordinating them (SIMAC, RAPID-MIX, PHENICX, MIReS and GiantSteps at scientific level). In addition, MTG-UPF is an internationally recognised research group because of its track record in TechTransfer and innovation actions, having established 3 spin-off companies (BMAT, Reactable Systems and Voctro Labs) and participated in the development of commercially available products such as the Vocaloid singing voice synthesizer (exploited by Yamaha Corp.), the Good-sounds.org repository (in collaboration with Korg) or the Loopmash VST3 Instrument integrated into Cubase (exploited by Steinberg).

The **Technical Committee (TC)** involves the key persons within the WP leaders who will be in charge of the coherence of the tasks running into the WP, ensuring that they all target their specific task goals but also aligned to objectives at WP level and that all intra/inter-dependencies are well identified and coordinated, organising meetings, consolidating partner information, preparing the reports for the PC and overseeing the timely submission and quality of deliverables. The WP1 leader is not included in this grouping because its role is already covered by the liaison between the PC and the PSC. The list of key persons within the TC is as follows:

WP nr.	WP title	Partner nr.	Partner acr.	Key person
2	Ontology specification and service orchestration	2	QMUL	George Fazekas
3	Intellectual property and business models	3	Surrey	David Plans
4	Semantic annotation of musical sound properties	1	MTG-UPF	Frederic Font
5	Semantic annotation of non-musical sound properties	3	Surrey	Mark Plumbley
6	Prototyping and evaluation in production workflows	2	QMUL	Mathieu Barthet
7	Dissemination and exploitation	1	MTG-UPF	Xavier Serra

3.3 Consortium as a whole

The AudioCommons consortium is composed of:

- 3 world-leading academic partners (MTG-UPF, QMUL and Surrey) with expertise in carrying out applied research in music technology, signal processing and semantic web technologies from different and complementary perspectives, and with experience in technology transfer activities. Our academic consortium also has experience in rights management and business models for open data.
- 3 industry partners (Jamendo, AudioGaming and Waves) with broad experience in developing tools and providing content to the creative industries. Jamendo is one of the biggest platforms for sharing independent Creative Commons music in the world. AudioGaming is an innovative company developing next generation audio tools for sound designers and video game developers. Waves is an award winning and world's leading developer of audio DSP solutions.

Partners come from 5 European countries, including Spain (MTG-UPF, from Barcelona), United Kingdom (QMUL and Surrey, from London and Guildford, respectively), Luxembourg (Jamendo), France (AudioGaming, from Toulouse) and Israel (Waves, from Tel-Aviv). This consortium covers the full value-chain of AudioCommons: from research, innovation and product development to content provision and evaluation, including all main profiles of potential actors involved in the Audio Commons Ecosystem (content users/creators/providers and tool developers). Furthermore, the consortium has the support of a number of associated experts from academic institutions and potential future ACE actors which will be consulted and invited to project meetings (see "Support Letters" section on the second part of this technical annex). RTD and MGT activities foreseen in AudioCommons work plan are assigned to the consortium members making the most out of the research potential and valuable assets to be provided from each. In addition, due to the relation already existing among partners, the collaborative work between and fluent communication them is assured (both QMUL and MTG-UPF have been partners in SIMAC and MIReS projects, QMUL and Surrey collaborate in the "Audio Data Exploration: New Insights and Value" project). The following table summarizes the expertise and nature of each partner in AudioCommons. More details on each member of the consortium are given in Section 4.

Partner	Country	Expertise and role in the project
MTG-UPF	ES	MTG-UPF is the coordinator of the project and will provide expertise in a number of sound and music computing areas such as audio analysis, automatic sound and music description, and collaborative and creative sound communities. MTG-UPF will thus lead the coordination and dissemination work packages (WP1 and WP7), and the semantic annotation of musical content work package (WP4). Furthermore, MTG-UPF, as developers and maintainers of Freesound, will participate in the Audio Commons Ecosystem as a content provider of music samples and non-musical content (with more than 230,000 sounds), and will therefore contribute in the ontology and API specification (WP2), provide use cases for the intellectual property rights management (WP3), and collaborate in the semantic annotation of non-musical audio content (WP5). MTG-UPF will integrate AudioCommons technologies in Freesound, and thus make its content available in the ACE.
QMUL	UK	QMUL will provide expertise in user evaluation, knowledge engineering and audio semantics including the analysis of audio content to extract high-level descriptors and the design of Semantic Web ontologies for the structured representation of music related metadata. For these reasons, QMUL will lead the ontology/API specification and the service orchestration (WP2), and the

		prototyping and evaluation work package (WP6), coordinating the industrial partners in the development of their tools. Furthermore, QMUL will contribute in the coordination and dissemination activities (WP1 and WP7, respectively), and in the semantic annotation of music content (WP4).
Surrey	UK	Surrey participation in AudioCommons includes involvement from three research groups which bring complementary experience in the project. These are the Centre for Vision, Speech and Signal Processing (CVSSP), the Institute of Sound Recording (IoSR) and the Business School. On the one hand, CVSSP provide expertise in signal processing and machine listening, which combined with IoSR's extensive experience in listener-based experiments and psychoacoustics will provide perfect skills for carrying out the required tasks for the semantic annotation of non-musical content (WP5), including the construction of a timbral model based on listening experiments. On the other hand, the Business School will provide their expertise in open commercial software delivery models, startups creation, entrepreneurship and global strategy for carrying out the tasks related to intellectual property management and business models (WP3). Furthermore, Surrey will also collaborate in the coordination and dissemination activities of the project (WP1 and WP7, respectively).
Jamendo	LU	Jamendo will provide their experience as an industry partner with a business model based on the commercialisation of Creative Commons music pieces, and will be a content provider and early adopter of the technologies developed in AudioCommons. Jamendo's catalog includes more than 470,000 tracks from 35,000 artists, one of the biggest platforms for independent music in the world. For these reasons, Jamendo as a content provider will collaborate in the definition of the ontology and API (WP2), in providing datasets for the semantic annotation of music pieces (WP4) and also in the creation of a web interface for accessing Audio Commons music pieces (WP6). Furthermore, Jamendo will bring their business model experience for carrying out the tasks related to intellectual property management and business models (WP3), and contribute in the dissemination of the project (WP7).
AudioGaming	FR	AudioGaming's expertise is on the development of next generation audio tools for sound designers and video games developers, proposing innovative solutions based on cutting-edge audio research developed both internally and in partnership with internationally recognized research institutions. In AudioCommons, AudioGaming is acting both as a tool developer and a content user, and will participate in the development and evaluation of a tool for the reuse of Audio Commons non-musical content in sound design tasks (WP6), and in the definition of requirements for the ontology and API specification (WP2). AudioGaming will also contribute in the dissemination of the project in professional game development and sound design events and tradeshows (WP7).
Waves	IL	Waves is the world's leading developer of audio DSP solutions for the professional, broadcast, and consumer electronics audio markets, with over 20 years of expertise in the field, awarded with a Technical GRAMMY® Award in 2011. Because Waves products are being used by the majority of audio professionals around the world, Waves plugins are ideal for incorporating Audio Commons content in the existing workflows of the sound professionals. For

		these reasons, Waves will contribute in AudioCommons with the development and evaluation of a tool for the reuse of Audio Commons music samples in music production workflows (WP6), as well as in the dissemination of the project in industry events and tradeshows (WP7), and in the definition of user requirements in WP2. Furthermore, Waves will collaborate in WP2 by providing their experience in web-based content licensing technologies.

Section 4: Members of the consortium

4.1 Participants

4.1.1 Music Technology Group from Universitat Pompeu Fabra (MTG-UPF)

Description of the legal entity

The **Music Technology Group (MTG)** of the **Universitat Pompeu Fabra** (UPF) in Barcelona, part of its Department of Information and Communication Technologies, is specialised in sound and music computing. With more than 50 researchers coming from different and complementary disciplines, the MTG carries out research on topics such as audio signal processing, sound and music description, musical interfaces, sound and music communities, and performance modeling. The MTG combines expertise in engineering disciplines such as Signal Processing, Machine Learning, Semantic Technologies, and Human Computer Interaction, to apply them in sound and music related problems.

The Department of Information and Communication Technologies (DTIC) at Universitat Pompeu Fabra (ranked 17th in the Times Higher Education Under 50 World University Ranking 2012) was created in 1999 with the mission to become one of the leading European university departments in the broad range of fields created around the convergence of ICT with biomedical and cognitive sciences, computation and networks. The DTIC has since its beginnings emphasised scientific excellence and internationalisation as core aspects of its activities.

The international vocation is, among other indicators, illustrated by the over 60 FP7 projects participated and coordinated by DTIC staff across all programs (Capacities, Cooperation, Ideas, People), including 9 ERC grants, involvement in 2 of the 6 European FET Flagship candidates (Human Brain Project selected in 2013), 37 projects within Cooperation in ICT, Health and SSHH programmes and 13 Marie Curie actions. DTIC is also active in Research for SMEs, the non-FP7 CIP and Ambient-Assisted Living. A full list of FP7 projects can be found at http://www.upf.edu/dtic/en/recerca/dtic_projects_fp7.html

This is complemented by an active role at a national level; close ties and agreements with large companies (Yamaha, Korg, Philips, GE, IBM, Yahoo!, Microsoft, etc) and SMEs; Catalan TECNIO recognition for successful industrial transfer to 4 groups, and research projects awarded by reputed international entities such as the McDonnell Foundation and the Google Research Awards.

Main tasks in AudioCommons

The role of the MTG-UPF in the project includes the coordination of the overall project activities, leading WP1 (coordination), WP4 (semantic annotation of musical content) and WP7 (dissemination), and contributing to WP2 (ontology/API specification and service orchestration), WP3 (IP and business models) and WP5 (semantic annotation of non-musical content). Besides its contribution to research tasks related to audio analysis and automatic content description, it also acts as content provider through their Freesound repository which contains more than 230,000 sounds and four million registered users. MTG-UPF is a leading research group in the field of sound and music computing, and provides the AudioCommons project with expertise in a number of sound and music computing areas such as audio analysis and automatic sound and music description (WP4, WP5), and collaborative and creative sound communities (WP2, WP3). The track record of this group in technology transfer and innovation activities is also a plus for the project. The Music Technology Group has established 3 spin-off companies (BMAT, Reactable Systems and VoctroLabs), usually collaborates with industry in joint projects and is behind the development of successful products available in market like Vocaloid (Yamaha Corp.), Cortosia/Artistry (Korg) or loopmash for Cubase software (Steinberg).

Key persons

Prof. Xavier Serra (M) is Associate Professor of the Department of Information and Communication Technologies and Director of the Music Technology Group at the Universitat Pompeu Fabra in Barcelona. After a multidisciplinary academic education he obtained a PhD in Computer Music from Stanford University in 1989 with a dissertation on the spectral processing of musical sounds that is considered a key reference in the field. In 2011 he was awarded an Advanced Grant of the European Research Council and since then he has been leading an international and interdisciplinary research team that works on music information research from a multicultural perspective.

Frederic Font (M) received a three-year degree in telecommunications engineering from Universitat Politècnica de Catalunya, Barcelona, Spain, in 2007. He obtained a MSc. in Sound and Music Computing from Universitat Pompeu Fabra, Barcelona, Spain, in 2010. He is currently enrolled in a PhD program at the Music Technology Group, and will defend his PhD thesis on June 2015. His research interests include the description of sound and music content and the characterisation of communities of users in online sound and music sharing sites to support musical creativity.

Dr. Dmitry Bogdanov (M) received a degree in applied mathematics and informatics at the Lomonosov Moscow State University, Moscow, Russia, in 2006. Since 2007, he is a member of the Music Technology Group at UPF, where he received the PhD. degree in information, communication, and audio-visual media technologies in 2013. Currently he is a post-doc researcher and the main developer of Essentia, a library for audio and music analysis renown by researchers and practitioners in the field of music technology and used in a number of industrial products.

Alastair Porter (M) received his Masters in Music Technology from McGill University in 2013. Since then he has been a researcher in the Music Technology Group at the Universitat Pompeu Fabra. He is an active member of the MusicBrainz project, an open database of information about music and musicians, both as a data contributor and a developer. In 2014 he helped to launch the AcousticBrainz project, which encourages people to contribute their audio to a collection of audio features extracted using the Essentia feature extractor.

Resources provided and previous experience related to the call content

Related initiatives and technologies:

Freesound[33]

Freesound is a collaborative repository of Creative Commons licensed audio samples with more than 230,000 sounds and 4 million registered users (as of February 2015). Sounds are uploaded to the website by its users, and cover a wide range of subjects, from field recordings to synthesized sound effects. Audio content in the repository can be tagged and browsed by folksonomic means as well as standard text-based search. Audio content in the repository is also analysed using the open-source audio analysis tool Essentia, which powers the similarity search functionality of the site. Freesound has a RESTful API through which third-party applications can access and retrieve audio content and its metadata. Nine years since its inception, Freesound has become one of the most popular sites for sharing sound snippets (46,000 unique visits per day, more than four million registered users, and more than 230,000 uploaded sounds).

[33] http://www.freesound.org

Essentia[34]
Essentia is an open-source C++ library for audio analysis and audio-based music information retrieval released under the Affero GPLv3 license. It contains an extensive collection of reusable algorithms which implement audio input/output functionality, standard digital signal processing blocks, statistical characterization of data, and a large set of spectral, temporal, tonal and high-level music descriptors. In addition, Essentia can be complemented with Gaia, a C++ library with python bindings which implement similarity measures and classifications on the results of audio analysis, and generate classification models that Essentia can use to compute high-level description of music. Essentia has been exploited for industrial applications: it has been used in industrial products by BMAT and Stromatolite (music selection), Yamaha's BODiBEAT (automatic playlist generation for runners), Steinberg's LoopMash (audio-based sample matching for music production), and Brace Yourself Games' Crypt of the Necrodancer (beat tracking for videogames). It also powers the online collaborative sound database Freesound.

AcousticBrainz[35]
The AcousticBrainz project aims to crowd-source acoustic information for all music in the world and to make it available to the public. This information describes the acoustic characteristics of music and includes low-level spectral information and information for genres, moods, keys, scales and much more. The goal of AcousticBrainz is to provide music technology researchers and open source hackers with a massive database of information about music, which is accessible through a RESTful API. As of April 2015, AcousticBrainz has already computed audio features for 1.9 million music tracks which are linked to MusicBrainz IDs providing standardised metadata. This database will spur the development of new music technology research and allow music hackers to create new and interesting applications.

Five recent publications related with AudioCommons:

- Ostuni, V. C., Oramas S., Di Noia T., Serra X., & Di Sciascio E. (In Press). A Semantic Hybrid Approach for Sound Recommendation. 24th International World Wide Web Conference (WWW 2015).
- Font, F., Serrà J., & Serra X. (2014). Class-based tag recommendation and user-based evaluation in online audio clip sharing. Knowledge-Based Systems. 67, 131-142.
- Font, F., Serrà J., & Serra X. (2014). Audio clip classification using social tags and the effect of tag expansion. AES 53rd International Conference on Semantic Audio.
- Font, F., Oramas S., Fazekas G., & Serra X. (2014). Extending Tagging Ontologies with Domain Specific Knowledge. International Semantic Web Conference.
- Sordo, M., Gouyon F., Sarmento L., Celma Ò., & Serra X. (2013). Inferring Semantic Facets of a Music Folksonomy with Wikipedia. Journal of New Music Research. 42(4), 346-363.

Previous projects connected to the proposal

CompMusic (ERC Advanced Grant 267583)
CompMusic is a research project funded by the European Research Council and coordinated by Xavier Serra from MTG-UPF. It aims to advance in the automatic description of music by emphasizing cultural specificity. It carries research within the field of music information processing with a domain knowledge approach, taking advantage of semantic technologies and advanced signal processing techniques.

MIReS (FP7 ICT-2011.1.5 Project No. 287711)

[34] http://essentia.upf.edu
[35] http://acousticbrainz.org

MIReS project aims to create a research roadmap of MIR field that will provide a meta-analysis of the MIR discipline, address emergent contexts and major challenges, formulate research evaluation standards for the discipline, contribute to the establishment of music production and digital library management standards, engage a variety of stakeholders from different disciplines of academia and industry and deliver innovative platforms for co-creative workshops focusing on horizon-scanning and technology foresight.

Metaverse1 (ITEA2 07016)
The Metaverse1 Project, "Setting Global Standards among Real and Virtual Worlds", started in 2008 and finished in 2011. It was an European ITEA2 project with the participation of industrial partners (Philips, Alcatel-Lucent, OrangeLabs, Technicolor, other) and research institutions (MTG-UPF, INRIA, Utrecht University…). MTG-UPF's participation focused in the area of analysis and synthesis of environmental sounds, developing a real-time engine for the autonomous generation of soundscapes and an automatic classification for environmental sounds.

SIMAC (FP6-507142 IST-2.3.1.7)
SIMAC was the a project funded by the European Commission and coordinated by the MTG between 2004 and 2006. SIMAC's main task was the development of prototypes for the automatic generation of semantic audio descriptors and the development of prototypes for exploration, recommendation, and retrieval of music. One special feature was the development and use of semantic descriptors. That is, ways to tag music that are close to the user's way of describing its contents. We developed tools to be used by music consumers, music distributors, and music creators.

Significant infrastructure relevant to the proposed work

DTIC houses infrastructure for both engineering work (computing, robotics and sensing, specialised audio-visual equipment) and experimentation with human subjects (experimental rooms with specialised equipment), complemented with links with hospitals and other external infrastructures.

The IT infrastructure covers three main layers: centralized storage and remote backup, supercomputing, and private cloud virtualization.

- Storage services are provided by two NetApp filers (FAS3140 on main site, and FAS2240 on backup site) housing 125TB of primary storage and 85TB on backup, with disaster recovery capabilities on backup site.
- The SNOW Supercomputer gives researchers more than 700 CPUs and 2,7TB RAM over 11 computing nodes, powered by Debian 7 and orchestrated with OpenGrid Scheduler job manager.
- Virtualization services are provided on a private cloud model, with all services housed onsite. Eighth powerful HP Proliant 385 G7, 10Gbit enabled, VMWare ESXi nodes, provide virtualization services to the research groups,who can host, deploy and manage all kind of services, from development and experimentation platforms to high-demanding, 24x7 infrastructure services.

System administration for the three layers is provided by UPF main IT services. This way, a dedicated team of system administrators works closely to the researchers within our department, finding together the best technical solutions for the researchers, taking care of the systems and achieving high levels of efficiency and systems uptime.

4.1.2 Queen Mary University of London (QMUL)

Description of the legal entity

The **Centre for Digital Music (C4DM)** at **Queen Mary University of London** (QMUL) is a world-leading multidisciplinary research group in the field of Music & Audio Technology. C4DM has over 60 members, who work on digital signal processing, music signal processing, music informatics, machine listening, audio engineering, interaction and performance. Research funding obtained since 2001 totals over £19M, primarily from the EPSRC and the EU. Current and recent projects include the £5.2M Fusing Audio and Semantic Technologies for Intelligent Music Production and Consumption (FAST-IMPACt, EP/L019981/1, with University of Oxford and University of Nottingham, 2014-2019) Sustainable Software for Digital Music and Audio Research (EP/H043101/1, 2010-2014), the £2.5M Online Music Recognition and Searching II (OMRAS-2, EP/ E017614/1, with Goldsmiths, University of London, 2007-2010), an EPSRC Platform Grant (EP/E045235/1, 2007-2012), the £5.9M Doctoral Training Centre in Digital Music and Media for the Creative Economy (EP/G03723X/1, 2009-2017), Linked Music Metadata (JISC, 2010-2011), Sustainable Management of Digital Music Research Data (JISC, 2011-2012), Musicology for the Masses (RCUK Digital Economy EP/I001832/1, 2010-2012), MIRes: Roadmap for Music Information Research (EU FP7-ICT-2011-7-287711, 2011-2013) and Networked Environment for Music Analysis (Mellon Foundation, 2008-2010).

QMUL's activity will take place in the context of one of the largest computational music research groups in the world, bringing expertise to the consortium in semantic audio analysis and processing, knowledge engineering and ontological modelling, musicology, music performance, cognitive science, engineering, machine learning and statistical modelling. The Centre for Digital Music (C4DM), which will host the project, consists of over 60 members at different academic levels, working in the fields of Machine Listening, Audio Engineering, Signal Processing, Music Information Retrieval, Semantic Web, Interactive Music Performance, Music Cognition and Computational Creativity.

Main tasks in AudioCommons

Within AudioCommons, QMUL will lead WP2 (ontology/API specification and service orchestration) and WP6 (prototyping and evaluation), and will contribute to WP4 (analysis of musical signals), WP1 (coordination) and WP7 (dissemination). QMUL has substantial experience in audio semantics including the analysis of audio content to extract high-level descriptors and the design of Semantic Web ontologies for the structured representation of music related metadata. An example of this is the Music Ontology which became a de-facto standard for representing music related information on the Semantic Web. This ontology found substantial use outside of academia including the British Broadcasting Corporation (BBC). The applicants within the group also have experience in large-scale deployment of feature extraction technologies, building services and API design. WP2 will build on our experience in ontological modelling and ontology based web applications in the context of Semantic Web and Linked Data. WP6 will utilize our experience in knowledge transfer, including the deployment of audio feature extraction services in a commercial environment, working with industry scale music databases and user evaluation in professional music production (see Making Musical Mood metadata completed in collaboration with the BBC and Computational Analysis of the Live Music Archive in collaboration with the Internet Archive). WP4 will benefit from our experience in audio and music analysis (see Queen Mary Vamp plugin set and Sonic Annotator Web Application).

Key persons

Dr George Fazekas (M) is a Lecturer in Digital Media at the Centre for Digital Music, Queen Mary, University of London (QMUL). He holds a BSc, MSc and PhD degree in Electrical Engineering. His PhD thesis investigates how semantic audio and Semantic Web technologies streamline music production workflows. He has published over 50 academic papers in the fields of Semantic Web, Ontologies, Music

Information Retrieval, Linked Data and Semantic Audio analysis. He participated in research and knowledge transfer projects as researcher, developer and at management level. He was Co-Investigator of research projects worth over €230K, including the JISC funded Shared Open Vocabularies for Audio Research and Retrieval project. He worked with BBC R&D to create mood-based music recommendation systems in the nationally funded Making Musical Mood Metadata project. He was papers co-chair and organising committee leader of the AES 53rd International Conference on Semantic Audio held in London, UK in 2014. He is regular reviewer for international conferences and journals including IEEE Transactions, JNMR and others. He is a member of the IEEE, ACM, BCS and AES and received the Citation Award of the AES for his work on the Semantic Audio Analysis Technical Committee.

Professor Mark Sandler (M) PhD FAES FIET FIEEE FBCS CEng is Founding Director of the Centre for Digital Music, a world-leading research group in audio and music technology with over 60 members. He is also Director of the Centre for Doctoral Training in Media and Arts Technology, a UK government funded special PhD training. Over the period 2010-2014 he was Head of the School of Electronic Engineering and Computer Science. He is currently supervising 6 PhD students with 32 previously awarded, and has examined over 30. He is a recipient of the prestigious Royal Society Wolfson Research Merit Award (2015-19). He has published over 400 papers in conferences and journals.

Dr Mathieu Barthet (M) is a Lecturer in Digital Media at the Centre for Digital Music at Queen Mary University of London. He was awarded a PhD in Acoustics, Signal Processing and Computer Science applied to Music from Aix-Marseille II University and CNRS-Laboratory of Mechanics and Acoustics in 2008 ("From performer to listener: an acoustical and perceptual analysis of musical timbre"). He joined the Centre for Digital Music in 2009 where he has since led research in the fields of Music Informatics, Music Perception, Human Computer Interaction, Computational Musicology and Semantic Web, and as part of EPSRC projects "Online Music Recognition and Searching 2", "Musicology for the Masses", "Making Musical Mood Metadata" and the AHRC project "Digital Music Lab: Analysing Big Music Data". He was general chair of the 9th International Symposium on Computer Music Modeling and Retrieval (CMMR) conference "Music and Emotions" held at Queen Mary University of London in 2012. He regularly reviews for conferences such as ISMIR, CMMR, IEEE ICME, SMAC, ACM RecSys and journals such as JASIST and IEEE Transactions on Multimedia.

Resources provided related to the call content

Relevant software and services:

Sonic Visualiser[36]

Sonic Visualiser is an application for viewing and analysing the contents of music audio files. The aim of Sonic Visualiser is to be the first program to reach for when studying a musical recording rather than simply listening to it. Sonic Visualiser is of particular interest to musicologists, archivists, signal-processing researchers and anyone else looking for a friendly way to take a look at what lies inside the audio file.

Music Ontology[37]

The Music Ontology provides a vocabulary for publishing and linking a wide range of music-related data on the Web. Music Ontology data can be published by anyone as part of a web site or an API and linked with existing data, therefore creating a music-related web of data. The Music Ontology is specified using the Ontology Web Language (OWL), which provides a set of constructs to describe domain models in Resource Description Framework (RDF). These standards and technologies are supported by the W3C and were chosen to tackle the many competing requirements of the music domain with a standalone format. By using

[36] http://www.sonicvisualiser.org
[37] http://musicontology.com

RDF, the Music Ontology gains a powerful extensibility mechanism, allowing Music-Ontology-based data to be mixed with claims made in any other RDF vocabulary.

Vamp Audio Analysis API[38]
Vamp is an audio processing plugin system for plugins that extract descriptive information from audio data — typically referred to as audio analysis plugins or audio feature extraction plugins. Similarly to audio effects plugin (such as a VST), a Vamp plugin is a binary module that can be loaded by a host application and fed audio data. However, unlike an effects plugin, a Vamp plugin generates structured symbolic information. Typical features that a Vamp plugin might calculate include the locations of moments such as note onset times, visualisable representations of the audio such as spectrograms, or curve data such as power or fundamental frequency.

Sonic Annotator Web Application[39]
SAWA demonstrates audio feature extraction technology developed at the Center for Digital Music, our standard plugin API for audio analysis and a Semantic Web ontology based representation for returning the results of the analysis. It is primarily aimed at researchers who wish to become familiar with these technologies. SAWA may be used as a batch feature extractor, for example, for automatically generating a small reference data set in RDF format. The resulting files can be loaded in a suitable RDF database and queried over, they may be browsed using a Semantic Web browser (e.g. the OpenLink RDF browser) or loaded and visualised with Sonic Visualiser a tool for studying the content of audio files.

Five recent publications related with AudioCommons:

- Kolozali, S., Barthet, M., Fazekas, G., Sandler, M. (2013) "Automatic Ontology Generation for Musical Instruments based on Audio Analysis", IEEE Trans Audio, Speech and Language Processing, Vol: 21, No.10, DOI: 10.1109/TASL.2013.2263801 , Page(s): 2207 - 2220
- Fazekas, G., Raimond, Y., Jacobson, K., Sandler, M. (2011). "An overview of semantic web activities in the OMRAS2 project", Journal of New Music Research (JNMR) Vol: 39, No: 4, Page(s): 295-311.
- Raimond, Y., Gängler, T., Giasson, F., Jacobson, K., Fazekas, G., Reinhardt, S., Passant, A. "The music ontology specification", Published online: musicontology.com, 2010-2015.
- Fazekas, G., Sandler, M. (2013) "Describing Audio Production Workflows on the Semantic Web" (invited paper) in Proc. of the 14th IEEE International Workshop on Image and Audio Analysis for Multimedia Interactive Services (WIAMIS) 3–5 July, Paris, France
- Fazekas, G., Barthet, M., Sandler, M. (2013) "The BBC Desktop Jukebox music recommendation system: A large-scale trial with professional users", in proc. IEEE Int. Conf. on Multimedia and Expo (ICME), 15-19 July, San Jose, CA, USA. DOI: 10.1109/ICMEW.2013.6618235

Previous projects connected to the proposal

FAST-IMPACt (EP/L019981/1)
The Fusing Semantic and Audio Technologies for Intelligent Music Production and Consumption project (worth £5.2M, funding: EPSRC, 2014-2019) brings the very latest technologies to bear on the complete music industry, end-to-end, producer to consumer. It aims at making the production process more fruitful, the consumption process more engaging, and the delivery and intermediation more automated and robust. The project addresses 3 premises: (i) that Semantic Web technologies should be deployed throughout the content value chain from producer to consumer; (ii) that advanced signal processing should be employed in the content production phases to extract "pure" features of perceptual significance and represent these in

[38] http://www.vamp-plugins.org/
[39] http://isophonics.net/sawa/

standard vocabularies; (iii) that this combination of semantic technologies and content-derived metadata leads to advantages and new products and services at many points in the value chain, from recording studio to end-user (listener) devices and applications.

OMRAS2 (EP/E017614/1)
The EPSRC funded "OMRAS2: A Distributed Research Environment for Music Informatics and Computational Musicology" project demonstrated (i) the utilities of high-level semantic features of musical audio (including musicological terms and free-form labels such as social tags) in multimedia content management, (ii) the use of low-level audio features and probabilistic statistical models to derive high-level semantic descriptors automatically, facilitating navigation in large online audio collections, (iii) the utilities of the Semantic Web, and Semantic Web technologies in online audio content navigation and delivery, (iv) the use of digital signal processing and machine learning for the manipulation of digital audio content on the semantic level, allowing interaction with notes, chords or performance characteristics such as vibrato by re-synthesising audio from parametrised descriptors. Several computational algorithms were developed for automatic annotation of musical audio, including novel methods for audio transcription, chord recognition, key recognition, tempo and beat detection, structural segmentation and music similarity. Several Semantic Web ontologies were created for describing and publishing music related metadata on the Semantic Web. The Music Ontology, a core framework connecting the OMRAS2 ontologies became a de-facto standard in music-related data publishing.

EASAIER (IST-FP6-033902)
The Enabling Access to Sound Archives through Integration, Enrichment and Retrieval was a two and a half year European research project addressing access to and preservation of cultural and scientific resources. The EASAIER project went well beyond the state-of-the-art in the targeted field of multimedia processing, archiving and accessing. An innovative approach to knowledge representation led to development of the Music Ontology and the Audio Features Ontology, which are now widely used outside the consortium. In the field of automatic feature extraction, QMUL developed new, high performance methods to identify and characterise sound objects (emotion detection, laughter detection, key extraction, tempo identifier...). In the area of presentation of multimedia material, novelty was presented in sound source separation, equalisation and noise reduction algorithms. A key innovation also allowed the video stream to be synchronised with the audio during real-time time and pitch scaling.

M4 (TS/J002283/1)
Making Musical Mood Metadata (M4) was a collaborative technology transfer project with BBC R&D and I Like Music. Over eighteen months, the team developed new and innovative methods of extracting high-level metadata from music contact, including information about the mood and emotional content of tracks. Having access to this information makes it easier for content producers to find the music they are looking for. Large-scale trials involving feature extraction from over 1M tracks were performed with BBC users using an online platform providing recommendations of sounds-like tracks for audio-visual productions. The M4 prototype was exhibited at the Cable and Satellite Exhibition in Dubai in 2013 and created a great deal of interest from both radio users and TV broadcasters.

The following two projects were collaborations with other partners of the consortium:
MIReS (FP7 ICT-2011.1.5 Project No. 287711) - MTG-UPF
SIMAC (FP6-507142 IST-2.3.1.7) - MTG-UPF
Audio Data Exploration: New Insights and Value (EP/M507088/1) - Surrey (description below)

Significant infrastructure relevant to the proposed work

Within the school of Electronic Engineering and Computer Science at QMUL, there is a small CPU/GPU platform which can be used as a cluster that consists of 212 cores and 10 Nvidia Tesla GPU cards. There are VMWare and Openstack/LVM virtualisation platforms. There is a Hadoop cluster making use of the 200 desktop machines. There are several compute servers for stand-alone workloads, which vary from small 48GB machines with 16 cores to half-terabyte machines with 32 cores. There is a shared 162 node heterogeneous cluster with around 2400 compute cores and 10TB of memory.

4.1.3 University of Surrey (Surrey)

Description of the legal entity

The **University of Surrey** is an internationally recognised UK university specialized in science, engineering, medicine and business. University of Surrey's participation in AudioCommons includes members from three research groups: the Institute of Sound Recording, the Centre for Vision, Speech and Signal Processing and University of Surrey's Business School.

The **Institute of Sound Recording (IoSR)** is a leading research group in psychoacoustic engineering, modelling the relationships between acoustic parameters and listener perception of audio, and building these models into systems to control, measure and describe the perceptual attributes of sound. Its work combines elements of acoustics, digital signal processing, psychoacoustics (theoretical and experimental), psychology, sound synthesis, software engineering, statistical analysis and user-interface design, with an understanding of the aesthetics of sound and music.

The **Centre for Vision, Speech and Signal Processing (CVSSP)** at the University of Surrey is one of the largest groups of its type in the UK with a grant portfolio of over £12M. CVSSP contains over 120 active researchers working in the areas of audio, speech, image, video, and graphics, and has received numerous EPSRC and EU awards, as well as supports from industry. Machine Audition Lab within CVSSP has led a number of projects in audio signal processing research including audio source separation, music transcription and spatial audio.

The **University of Surrey's Business School** staff team has decades of experience in startups creation, entrepreneurship and global strategy. Research fields cover computer science; the investigation of "open commercial" software delivery models; software design for media. Such expertise clearly fits with the challenges of AudioCommons.

Main tasks in AudioCommons

Surrey's participation in AudioCommons will include leading WP3 (IP and business models) and WP5 (semantic annotation of non-musical content), and contributing to WP1 (coordination) and WP7 (dissemination). The IoSR has extensive experience in designing and running listener-based experiments, analysing and profiling the resulting data, correlating physical and perceptual parameters, and using these correlations to inform the development of models of both low- and high-level perceptual attributes of audio. These are precisely the stages through which the research in WP5 will proceed, making the IoSR ideally suited to the work in this package. The CVSSP has significant expertise in audio signal processing, machine listening, sparse representations, and the development of audio algorithms and software. They will provide complementary skills in low-level signal processing and feature extraction for timbral description, modelling and tagging and will additionally manage software standards and the interface between the timbral characterisation system and the project's other software elements. The Business School's staff team has decades of experience in startups creation, entrepreneurship and global strategy. Research fields cover computer science; the investigation of "open commercial" and Libre/Open Source software delivery models

and creative commons licensing; software design for media. Such expertise clearly fits with the planned results of WP3.

Key persons

Prof Mark Plumbley (M) is a Professor in Signal Processing in CVSSP where he joined in January 2015. He was Director of the Centre for Digital Music at Queen Mary University of London before moving to CVSSP. He is a leading expert in machine listening and signal processing. He has led a number of EPSRC, EU and industrial projects in the area of audio analysis and processing, including the use of sparse representation and latent variable analysis techniques, such as non-negative matrix factorization, for audio source separation, music transcription, and audio event detection.

Dr Wenwu Wang (M) is a Reader in Signal Processing in the CVSSP and Co-Director of Machine Audition Lab. His research focuses on audio signal processing, machine learning, and sparse representations, the EPSRC- and industry-funded projects he has led and worked on have involved the development of advanced signal processing algorithms for source separation, feature extraction, and sparse representations.

Dr Tim Brookes (M) is Director of Research at the IoSR and has led multiple EPSRC- and industry-funded research projects in psychoacoustic engineering, resulting in over 50 publications. Internationally recognised contributions include new techniques for perceptual testing, profiling and modelling, and for perceptually-enhanced source separation. He was an Executive Board Member in the Digital Music Research Network and CI of the Spatial Audio Creative Engineering Network.

Dr Russell Mason (M) is Senior Lecturer in the IoSR and is Programme Director for the Tonmeister programme in Music & Sound Recording. His research focuses on psychoacoustic engineering and the EPSRC- and industry-funded projects he has led and worked on have involved the analysis of audio descriptors, development of models of auditory perception, and implementation of novel subjective experimentation techniques.

David Plans (M). More than 15 years experience running startups in the technology sector. He helped foster the integration of an EU-wide consortium of Open Source SMEs, and has helped NHS Primary Care Trusts implement mobile health strategies. With a background in software design for media, he is currently running a successful startup (BioBeats) and teaching entrepreneurship and innovation in both undergraduate, postgraduate, and MBA-level modules in the Surrey Business School.

Carla Bonina (F) is Lecturer in Entrepreneurship and Innovation at Surrey Business School. Her research interests lie in the intersection of technology innovation, entrepreneurship and policy. Her current projects revolve around the critical implications of big and open data, value creation, and new business models in the digital economy. She has performed consultancy roles for the government of Mexico and Kazakhstan, telecom operators in Latin America and Europe. She has also provided advice to international organizations such as the OECD, the IDRC, the World Bank and the ITU.

Resources provided related to the call content

Five recent publications related with AudioCommons:

- Francombe J, Mason R, Dewhirst M, Bech S. (2014) '*Elicitation of attributes for the evaluation of audio-on-audio interference*'. Journal of the Acoustical Society of America, United States: 136 (5), pp. 2630-2641.
- Conetta R, Brookes T, Rumsey F, Zielinski S, Dewhirst M, Jackson P, Bech S, Meares D, George S. (2014) '*Spatial Audio Quality Perception: Impact of Commonly Encountered Processes / A Linear Regression Model*'. Journal of the Audio Engineering Society, 62 (12), pp. 831-860.

- A. Alinaghi, P. Jackson, Q. Liu, and W. Wang, "*Joint Mixing Vector and Binaural Model Based Stereo Source Separation*", IEEE/ACM Transactions on Audio Speech and Language Processing, vol. 22, no. 9, pp. 1434-1448, 2014.
- M.G. Jafari and M. D. Plumbley. *Fast dictionary learning for sparse representations of speech signals*. IEEE Journal of Selected Topics in Signal Processing 5(5): 1025-1031, Sep 2011.
- Liebenau JM, Elaluf-Calderwood SM, Bonina CM. (2014) '*Modularity and network integration: Emergent business models in banking*'. Proceedings of the Annual Hawaii International Conference on System Sciences, pp. 1183-1192. doi: 10.1109/HICSS.2014.153

Previous projects connected to the proposal

S3A: Future Spatial Audio for an Immersive Listener Experience at Home (EP/L000539/1)
S3A is a major five-year UK research collaboration between internationally leading experts in 3D audio and visual processing, the BBC and UK industry. Its goal is to enable listeners to experience the sense of "being there" at a live event, such as a concert or football match, from the comfort of their living room through delivery of immersive 3D sound to the home. The IoSR's role in S3A is to: (i) ascertain which attributes of reproduced spatial audio are most important to listeners; (ii) identify any important attributes missing from previous studies; (iii) determine the relationships between the important attributes and listener preference; (iv) model overall spatial quality in terms of the important perceptual attributes; and (v) model these perceptual attributes in terms of their physical correlates. The methods employed in determining the relationships between physical parameters and multiple aspects of perception will be directly applicable to the timbral modelling tasks for the semantic annotation of non-musical audio content in AudioCommons.

Audio Data Exploration: New Insights and Value (EP/M507088/1)
This project, funded by Innovate UK and EPSRC, is a collaboration between Audio Analytic Ltd., QMUL and Surrey. The project is undertaking advanced audio data analysis and modelling techniques in Automatic Environmental Sound Recognition, to create value across a variety of applicative domains. As well as current application areas such as Professional Security and Home Security, a range of novel markets can be developed in relation to Multimedia Database Indexing, Environmental and Industrial Monitoring, the Internet of Things and more. The project will gather the newly developed audio analysis and modelling techniques into a demonstrator instantiated as a "Personal Audio Space Indexer". The techniques for analysing and visualizing everyday sounds will be particularly relevant for WP5.

Musical Audio Repurposing using Source Separation (EP/L027119/2)
This EPSRC-funded project aims at developing a new approach to high quality audio repurposing (upmixing and remixing), based on high quality musical audio source separation. To achieve this, University of Surrey's CVSSP will combine new high resolution separation techniques with information such as musical scores, instrument recognition, onset detection, and pitch tracking. In parallel, perceptual evaluation measures for source separation, remixing and upmixing will be investigated, and new diagnostic evaluation techniques tailored to measure different aspects of the repurposed outcome will be developed. The outcomes of this project will allow music consumers to enjoy their favourite songs in interactive remixing apps and games, even where the original separate "stems" are not available. It will also allow music companies, broadcasters and sound archive holders to provide high quality upmixed versions of their large archive content, for an increasing generation of listeners with surround sound systems in the home. The combination of source separation technologies and audio repurposing, plus the link between signal processing and perceptual models, will inform the work of both WP4 and WP5 (see section 3.1).

QESTRAL (EP/D041244/1)
This project, in partnership with Bang & Olufsen, DK, and BBC R&D, UK, developed a set of test signals and related physical measures that can be used to evaluate the performance of spatial audio processing

systems in relation to a range of perceptually important spatial quality attributes of a reference system, and modeled the reproduced sound field generated by selected transducer arrays at a number of monitoring positions within a listening space. The processes of identifying key perceptual attributes, refining and expanding on existing metrics for these attributes, and building them into a comprehensive sound characterisation system, are very closely related to those that will be employed in AudioCommons.

POSZ: Perceptually Optimised Sound Zones (EPSRC)
This project, funded by Bang & Olufsen, DK, and EPSRC, combined engineering (to create the sound zones) and psychoacoustics (to evaluate and predict the perceived quality). CVSSP developed methods to create sound fields where the audio is concentrated on the corresponding sound zones, with minimal spill into other zones. IoSR determined the most appropriate perceptual factors when listening interfering signals in a sound zone, and developed perceptual models that can be used in the cost function of the system optimisation. This synergy between psychoacoustics and engineering is precisely what will drive part of the research in AudioCommons.

ENDuRE : European Network of Design for (u)Resilient Entrepreneurship
(554337-EPP-1-2014-1-IT-EPPKA2-KA)
This project, in partnership with iVeridis, Blue Ocean Robotics, IDEA, CEDIT and the Universities of Pisa, Surrey and Southern Denmark is looking at emergent business models and methodologies for inter european entrepreneurship. Part of the remit of ENDuRE is to address how new economic models in the digital economy can survive the pressures of market change, and take advantage of cross-subsidisation from other industries, such as music resold/used within games for health, which will be particularly pertinent for AudioCommons research.

Significant infrastructure relevant to the proposed work

The IoSR has three professional recording studios with extremely well equipped control rooms and live rooms built to strict acoustic standards, together with an ITU-R BS.1116 critical listening room. These facilities will be used to record and process test stimuli and to run the controlled listening experiments that will allow the connections between acoustic parameters and human perception of timbre to be established.

CVSSP has an Anechoic Chamber and a Media Engineering Lab with state-of-the-art acoustic capture and analysis facilities enabling research into audio source separation, music transcription and spatial audio, and computing infrastructure comprising a cluster of over 20 servers each with hyper-threaded quad or dual core CPUs and between 32-128MB RAM running Ubuntu Linux. These facilities will complement those in IoSR and will be used for running advanced signal processing algorithms and software to be developed in this project.

4.1.4 Jamendo SA (Jamendo)

Description of the legal entity

Jamendo is one of the biggest platforms for independent music in the world. The goal of the website is to bring an entirely free service for all music lovers to discover music, to provide the most simple ways for an artist to share its music online, and to offer a variety of music for professionals to license for commercial at a fair price. Based on Creative Commons licenses, the music catalog is entirely free for users to stream, download and share. For artists, it's a very easy solution to find a wider audience internationally by simply uploading music under Creative Commons licenses, allowing them to decide what usage can be made for each work shared on the platform. Jamendo helps artists monetizing their content thanks to Jamendo Licensing, a commercial program accessible by professionals who need to license music for commercial purposes (advertising, film, TV, in-store music, etc.). Jamendo directly shares 50% of revenue with its artists.

Main tasks in AudioCommons

Jamendo as a content provider will collaborate in the definition of the ontology and API (WP2), in providing datasets for the semantic annotation of music pieces (WP4) and also in the creation of a web interface for accessing AudioCommons music pieces (WP6). Furthermore, Jamendo will bring their business model experience for carrying out the tasks related to intellectual property management and business models (WP3), and contribute in the dissemination of the project (WP7). Jamendo.com managed to attract 35,000 artists from all over the world, who contributed more than 470,000 tracks since 2005. Thanks to the open nature of the Creative Commons licenses applied to each and every one of those tracks, Jamendo is an ideal source for anybody who need specific music for their working projects (videomakers, directors, ad agencies, other musicians, etc.): tracks can be used for free for private use, and the rights are pre-cleared for commercial use, only requiring a one-time license purchase. Thanks to Jamendo's open API, the whole catalog is easy to access and connect to any third-party program or application, and this possibilities will be extended by integrating AudioCommons technologies and offering Jamendo content in the ACE.

Key persons

Martin Guerber (M) has been in the music industry for almost 10 years, working for independent labels, booking agencies and artist management companies. Since 2011 he is the music ear of Jamendo, in charge of the music content, working closely with artists and partners who are contributing new music and enriching the catalog on a daily basis.

Resources provided related to the call content

Jamendo[40]
Jamendo is the biggest free music platform under Creative Commons licenses. It currently hosts 470,000 tracks from 35,000 unique artists, featuring 2.3 million members and 50,000 unique visitors/day.

Jamendo Licensing[41]
Jamendo Licensing (20,000 clients in 2014) is an extension of Jamendo.com, offering all the Jamendo music that is available for commercial use (use in film, TV, advertising, in-store background music) thanks to an opt-in option for all artists. Jamendo Licensing has almost 200,000 tracks (40% of the whole Jamendo

[40] http://www.jamendo.com
[41] http://licensing.jamendo.com

catalog), the ones that the Jamendo artists have chosen to add to the program in order to license and make money.
Jamendo Licensing offers 3 services:

- Catalog: the full library with an advanced search engine to help anyone find the right song for a project.
- In-Store: a solution for all public places and businesses (stores, hotels, restaurants, etc.) needing to play some music. 21 radios are available, covering different genres, moods and themes.
- Composition: an on-demand service for clients wishing to have a custom-made track based on very specific needs. Jamendo crowd sources those demands and ask our community of artists to submit creations in response.

Previous projects connected to the proposal

UGC Music Discovery project: Development of innovative music discovery systems in User Generated Content music websites
In the free content world (based on free license), we cannot base the development of a company on the ownership of the content as the provider was only hosting the content. Jamendo music can be proposed on many other websites. The value of the company is based on the ecosystem that the company can build around the content it includes:

- The ability to provide a reliable service to discover/listen/download the music for users and customers
- The quality of the relationship and management of customers
- The competitiveness of the pricing model
- The capacity of integration with other music business partners
- The additional data and information the company can create and aggregate around the music is the key value and a true advantage on the competitors.

The aim of the project is to increase the level of quality and reliability of all the services that we can build around the music. It should create a virtuous circle where music discovery would encourage users to visit Jamendo, so artists know that they can find more audience thanks to Jamendo. Knowing that Jamendo music has found a public, customers would be keener to use Jamendo music for their business.

Relevant publication:

- Bazen, S., Bouvard, L., & Zimmermann, J. B. (2014). 'Jamendo: The Heartbeat of Free Music!': Musicians and the Creative Commons. Musicians and the Creative Commons.

4.1.5 AudioGaming (AudioGaming)

Description of the legal entity

AudioGaming is a company developing next generation audio tools for sound designers and video games developers. AudioGaming proposes innovative solutions based on cutting-edge audio research developed both internally and in partnership with internationally recognized laboratories. Thanks to AudioGaming's procedural audio engine D.O.S.E. (Dynamic Object Sound Engine), their products adopt a new and refreshing point of view always giving control over the sound generation or sound processing in real-time, with extensive use of sound manipulation and sound synthesis techniques to produce realistic controllable sounds and effects.

AudioGaming's goal is to create meaningful audio tools for next generation audio needs. Products are focused on increasing efficiency and quality of sound designers' work as well as crafting ergonomic tools offering new creative possibilities. AudioGaming seeks to bring help for tedious and repetitive tasks by offering smart tools able to understand and process sound depending on its intrinsic properties. AudioGaming's procedural audio technology enables a dramatic change in the way sound designers can create and sound traditional and interactive media.

Main tasks in AudioCommons

Within AudioCommons, AudioGaming will mainly participate in WP6 (prototyping and evaluation), but will also contribute in WP2 (ontology/API specification and service orchestration), and in WP7 (dissemination). In WP2, AudioGaming will give feedback and overviews of the creative processes used in the video game industry to assess the necessity of the scientific developments to be carried out in WP4 and WP5. In WP6, AudioGaming will research & develop a sound design plug-in with a direct interfacing with AudioCommons non-musical content (sound effects). This plugin will allow end users to efficiently get access to and upload sounds in the plug-in along with its licence. They will then be able to use the sound or part of the sound to create new sound fxs by layering them with other sounds or transforming them in real-time using specific effects developed by AudioGaming. This plug-in will be compatible with most of today's industrial standard DAW like Protools, Nuendo, etc. Furthermore, AudioGaming will research & develop a demonstrator for its future interactive sound design tools creation platform. The demonstrator will showcase the integration of AudioCommons contents in an interactive plug-in creation platform. This will serve as an example showing the seamless integration of AudioCommons content in future tools. Through its activities AudioGaming is acting both as a tool creator and an interactive content user. We have an ideal position to formalise a good overview of the industrial pipelines used both in the audio post-production market and in the interactive industries. Since we are also market leaders in terms of procedural audio tools & innovative interactive sound design approaches, we are ideally placed to create demonstrators showcasing the power hidden in large scale open content initiatives like AudioCommons.

Key persons

Amaury La Burthe (M) - CEO / Product Design. After a few years in the computer industry, Amaury La Burthe reoriented to his true passion: he followed a master in acoustics and signal processing applied to music at IRCAM in Paris. He then worked as assistant researcher for Sony and as lead audio designer for video games company Ubisoft. Through his professional experiences he developed a dual expertise: scientific and artistic in the field of sound design and specifically in the creation of interactive material. For SONY, Amaury worked on various music browser prototypes and recently for interactive music prototypes. For Ubisoft, he worked as audio designer on *Prince of Persia: Warrior Within* & *Splinter Cell: Chaos Theory* and as lead audio designer for *Splinter Cell: Double Agent* & *Far Cry 2*. He recently completed entrepreneurship programs at Babson College in Boston and EMLyon. Since 2008, Amaury La Burthe is an

adjunct lecturer on interactive audio at Master level for French video game school ENJMIN. He has also been involved in Canadian research project ARC-PHONO with UQAM (Prof. Mario Coté) and Hexagram for which he created interactive audio software. He his co-author of several scientific publications and has given talks on signal processing (ISMIR 2002, Paris, CBMI 2003, Rennes, WedelMusic 2003, Leeds) and interactive audio (SAT Montreal 2005: "Sound space in the context of video games", FutureGameOn-ACMSiggraph 2010, Paris: "Innovative audio technologies for creative gameplay using sound", MIGS Montreal, 2010, 2011, GDC 2013 San Francisco).

Chunghsin YEH (M) - Main researcher. During his college years at National Taiwan University, Chunghsin Yeh spent much time in local studios and concert halls learning recordings and digital audio production. He found his passions in digital audio signal processing and therefore, he went to IRCAM in Paris to pursuit his Master and Ph.D. His award-winning algorithm of multiple fundamental frequency estimation has recently been integrated into commercial products. He his 4 times winner of the Mirex (Music Information Retrieval Evaluation eXchange) competition for polyphonic signal transcription. Since 2012. he joined AudioGaming as Main Researcher to develop innovative analysis/synthesis technologies.

Resources provided related to the call content

Since mid 2012, AudioGaming has released 4 innovative products: AudioWind, AudioRain, AudioMotors, AudioSteps. These products are mainly based on mathematical modelings and generate real-time sounds for audio post-productions. They have no equivalent on the market at this day. These products have been used by companies like LucasFilm, Ubisoft or Sounddelux on projects like Assassin's creed licence or Django Unchained for the most famous ones. AudioGaming is an official laureate of the Tribeca New Media Fund in 2015 for it's original interactive application Notes On Blindness. This application to be released in 2016 depicts the perception of blind people and how they interact with the world. AudioGaming's engine technology is used by the next official World Rally Championship video game to be released at the end of 2015. AudioGaming has contributed to the design and implementation of gamified audio content for the Jumping Sumo drone created by the French company Parrot. Through its former CTO, AudioGaming has participated in the elaboration of Google's project Cardboard. 500k+ Cardboards have been shipped until today. AudioGaming has done all the interactive audio technology and content of the serious game Type:Rider. Type:Rider is the first videogame created by French national television Arte. It's been selected as iOS pick of the week by starbuck which distributed 1 million copies of the game. The game has been featured in FastCompany and won several awards.

Previous projects connected to the proposal

For SONY, Amaury La Burthe worked on various music browser and interactive music prototypes. This work has been featured in several scientific publications/communications:

- La Burthe A., 2005. L'espace sonore dans le contexte du jeu vidéo. Conférencier invité dans le cadre du cycle de conférences '*Son numérique, maîtrise de l'espace sonore*', Société des Arts Technologique (S.A.T.), Montréal, Canada. (Vidéo).
- La Burthe A., Pachet F., and JJ. Aucouturier, 2003. *Editorial metadata in the cuidado music browser: between universalism and isolationism*. Proceedings of the WedelMusic Conference, Leeds, UK.
- Pachet F., La Burthe A., and Aucouturier JJ., 2003. *The Cuidado Music Browser: an end-to-end electronic music distribution system*, Third International Workshop on Content-Based Multimedia Indexing (CBMI03), Rennes, France.
- Peeters G., La Burthe A., and Rodet X., 2002. *Toward Automatic Music Summary Generation from Signal Analysis*. Third International Conference on Music Information Retrieval (ISMIR2002), Paris, France.
- Pachet F., La Burthe A., Aucouturier JJ. and Beurivé A., 2005. *Editorial Metadata in Electronic Music Distribution Systems: Between Universalism and Isolationism*. Journal of New Music

Research, 34(2):173-184.
- Pachet F., La Burthe A., Zils A., and Aucouturier JJ., 2004. *Popular Music Access: the Sony music browser*. Journal of the American Society for Information Science and Technology, 55(12): 1037-1044.

For Ubisoft, Amaury La Burthe worked as audio designer on Prince of Persia: Warrior Within & Splinter Cell: Chaos Theory and as lead audio designer for Splinter Cell: Double Agent & Far Cry 2.

4.1.6 Waves Audio Ltd. (Waves)

Description of the legal entity

Waves Audio is the world's leading developer of audio DSP solutions for the professional, broadcast, and consumer electronics audio markets. Since its start in the early '90s with the introduction of the Q10 equalizer plugin, Waves has gone on to develop a comprehensive line of over 200 audio plugins, including industry standards like the L1 and L2 Ultramaximizers, popular vintage console models, and innovative mixing tools like Vocal Rider and the Artist Signature Series. For its accomplishments, Waves received a Technical GRAMMY® Award in 2011, and the Q10 was selected as an inductee into the TECnology Hall of Fame.

Waves technologies are now used to improve sound quality in virtually every sector of the audio market, from recording, mixing, mastering and post-production to broadcast, live sound, and consumer electronics. Waves has over 20 years of expertise in the development of psychoacoustic signal processing algorithms that leverage knowledge of human auditory perception to radically improve perceived sound quality. Waves' award-winning processors are utilized to improve sound quality in the creation of hit records, major motion pictures, and top-selling video games worldwide. Waves offers computer software and hardware-plus-software solutions for the professional and broadcast markets.

Main tasks in AudioCommons

Within AudioCommons, Waves will participate in WP6 (prototype and evaluation), WP2 (ontology/API specification and service orchestration) and WP7 (dissemination). In WP2, Waves will collaborate in the definition of requirements and by providing their experience in product licensing. In terms of WP6, Waves provides 2 major contributions. Firstly, Waves will integrate AudioCommons results into Ardour (an open source hard disk recorder and digital audio workstation application for professional use) and Tracks Live (a multitrack recording software solution designed to capture live performances of any kind, from the smallest gig to the largest production). In addition, Waves will also research and develop a sampler plugin that will connect the end user and music professional to the AudioCommons content. This connection will allow the user to locate, access and obtain the license for the contents published in the ACE. Waves is the world leader in developing audio DSP for the professional audio market. As Waves products are being used by the majority of audio professionals around the world, Waves plugins are the natural choice for incorporating AudioCommons in the workflow of the sound pro. Waves, as the leading developer of audio DSPs, has a large team (15) of researchers in the field, being able to both provide the platform for offering AudioCommons to the world, and to develop the necessary technology to make it happen. Waves has also developed digital audio platforms and musical instruments (Tracks Live, MultiRack) , and based on the expertise obtained, is the natural candidate to develop the link between content and the tools to work on it.

Key persons

Meir Shashua (M) - CTO and Co-Founder. Mr. Shashua co-founded Waves in 1991 and presently serves as Chief Technology Officer responsible for the development of all products and technologies. Prior to founding the Company, he worked together at Audio Animation, where he was responsible for the development of all the DSP algorithms for a digital audio processor for FM radio stations. Mr. Shashua was part of the Talpiot program in the IDF, serving as a R&D officer for five years. During his military service he was active in developing DSP algorithms and control software for an experimental radar system, and later served as the head of the RPV section. He holds a B.Sc. from the Hebrew University in Mathematics and Physics, and also studied at Tel Aviv University's Department of Communication Control and Computers.

Yuval Levi (M) Senior software developer/architect with strong and rich technical and managerial skills. Software architect - Audio Commons. Yuval has, in the past, Co-founded and was COO, Software architect leader development team in various startups, such as: Robust strategic software Systems Ltd , Cerylion Inc, Eyepic ltd ,FLAIR Technologies ltd. He holds a B.Sc. from the Hebrew University in Computer science and Mathematics and also studied towards BMus. in Piano/Jazz performance at the Rubin Academy for music and dance in Jerusalem.

Rami Shapira (M) Product Manager. Creator and Producer of electronic music since 1988. A pioneer of electronic music in Israel, with extensive experience in musical arrangement and production. Rami was also a lecturer in various schools of music in Israel, music consultant and Waves tech support expert.

Resources provided related to the call content

Tracks Live
Tracks Live is a multitrack recording software solution designed to capture live performances of any kind, from the smallest gig to the largest production.

Waves Plugins
Waves have developed over 200 audio plug-ins which are used to improve sound quality in virtually every sector of the audio market, from recording, mixing, mastering and post-production to broadcast, live sound, and consumer electronics. As Waves products are being used by the majority of audio professionals around the world, Waves plugins are an unbeatable choice for incorporating AudioCommons in the workflow of sound professionals.

Waves License Center
Wave License Center (WLC) is a client server system for managing Waves licenses. The system allows the customer to independently manage his licences, (on his computers or DOK). It also allows for off-line work.

VSSD: Waves VSSD
Virtual stock software distribution is an internet-based fulfilment system that manages the sale and delivery of downloadable software products to the worldwide dealer/distribution network. The system allow the users (dealers, stores) to sell licences without having to hold inventory (and at up-to-date terms). VSSD is password-protected, dealer-specific, and includes parameters for margins, products, and payment terms. It allows dealers and distributors to order products and manage their contact with Waves via an online account they open on Waves IT system.

On-line store
Waves sells licenses and products at its online store at www.waves.com.

Significant infrastructure relevant to the proposed work

Waves' web services are hosted in a DataCenter in the USA: 90% of Waves' services there are based on VMWare ESXi. Waves website (Based on Microsoft IIS + SQL Server), includes content + online store which is protected with Verisign certificates. (We have average of 800,000 Session per month – not including peak times such as the Black Friday weekend). Vssd.waves.com – Online store for Waves Dealers – Based on Microsoft IIS + SQL Server protected by Verisign certificate. Waves License Server - Based on Microsoft IIS + SQL Server protected by Verisign certificate. Waves' also hosts services on Amazon AWS: Waves Online Installer – online server that host our installers to download and install specific features from our software are hosted on Amazon EC2. Offline installers are hosted on Amazon S3 Buckets, linked from our sites to download our full installer directly from Amazon Cloud.

4.2 Third parties involved in the project (including use of third party resources)

Does the participant plan to subcontract certain tasks (please note that core tasks of the project should not be sub-contracted)	N
Does the participant envisage that part of its work is performed by linked third parties	N
Does the participant envisage the use of contributions in kind provided by third parties (Articles 11 and 12 of the General Model Grant Agreement)	N

Section 5: Ethics and Security

5.1 Ethics

5.1.1 Involvement of human participants in AudioCommons tasks

There are two main type of activities foreseen in the AudioCommons project that deal with the collection of data coming from end users. Firstly, there are some tasks in the workplan which aim at deploying, validating and gradually improving the base technologies through their exposure to the user base available to our partners. Thus, various software prototypes resulting from the technology development activities in the AudioCommons project will be made available to end users for **beta testing** through different platforms and users will be able to voluntarily use those platforms (without prior identification nor recruitment). All the Beneficiaries are involved in the testing of prototypes.

Second, within WP5, University of Surrey (Beneficiary nr 3) is carrying out a series of **perceptual tests involving human participants** is considered in order to model the relationships between acoustic parameters and listener perception of audio, measure those subjective attributes of sound, assessing their usefulness and building standard vocabularies around those concepts.

The target users to be involved in both the beta testing activities and perceptual studies are adult healthy volunteers. There is no interest in usage data of children, minors or adults who are unable to give informed consent. The ethical standards and guidelines of Horizon2020 will be rigorously applied, regardless of the country in which the research is carried out.

5.1.2 Detailed information about the procedures implemented for data collection and its handling

Beta testing of the AudioCommons prototypes
The main goal behind collecting usage data is to better understand which features/parts of the prototypes need to be further improved. We plan to log the user's behaviour when using the prototypes in an anonymous manner. The data to be collected includes only information such as: user actions that lead to a failure of the application, the most common usage patterns of the prototype and so on.

No personal or private information of the users will be collected, no access to records of personal or confidential information or IP addresses will be needed. The users will be asked if they agree to share their usage data in an anonymous manner and be used for improving the AudioCommons technological outcomes. Users will have an option to opt out from sharing such logged activity both during the initial registration and anytime after while using the application.

In order to track this activity data, the user will be requested to provide a nickname or alias and an email address, which alone are insufficient to identify individuals. During the user registration stage, and prior to the activity tracking, the users will be assigned an ID, thus disassociating this information from users' identifying data.

All collected information and data will be used during the AudioCommons project lifecycle and exclusively for the purpose stated (testing the application, no research purposes).

Perceptual tests involving human participants
No personal data will be collected during the perceptual tests. Participants will listen to audio excerpts and will be asked to provide data relating to their perception of those excerpts. For example, participants might be asked to rate each excerpt in terms of its richness or brightness or clarity on a scale from 0 to 100.

Participants will typically give their responses via a graphical user interface to a piece of computer software. All data will be captured and stored anonymously.

Prior to the test, each participant will be briefed verbally and in writing, and will be given an opportunity to ask questions before signing a consent form (see Annex 1 attached to this document) to confirm that they understand the nature of the test and what will be required of them, that they agree to take part, and that they consent to the use and possible publication of the captured anonymous data.

All participants agreeing of their own accord to participate will be asked to sign an Informed Consent form which will be written according to the EU guidance notes. The ethical standards and guidelines of Horizon2020 will be rigorously applied, regardless of the country in which the research is carried out.

Participants will be able to withdraw from the session at any time. Contact information for non-project staff will be provided to participants as a way to establish complaint procedures.

The participant's name, the name of a witness to their signature, and the date, will be the only pieces of personal information entered on the consent form. The forms will include no participant-related reference numbers. It will not be possible from the forms, nor from the captured test data, nor from the combination of the two, to relate any element of the test data to any particular participant.

5.1.3 Gender aspects

In terms of gender analysis, AudioCommons will take all measures necessary to promote equal opportunities between men and women in the implementation of the action. To the extent possible, a gender balance will be ensured at all levels of personnel assigned to this project, including the activities addressed to product validation through focus groups and interviews involving potential early adopters and end users.

5.1.2 Data protection

Beta testing of the AudioCommons prototypes
The project will conform to privacy and confidentiality guidance from the EU guidance notes, "Data protection and privacy ethical guidelines" and to the Data Protection Directive (Directive 95/46/EC, http://ec.europa.eu/justice/data-protection/index_en.htm)

Secure Access Policy: Data will be encrypted and password protected. Only members of the team directly working with the data ("need to know") will have authorisation to access the data.
Secure Storage: Location and Hardware. All personal data will be stored in digital hard disks on computers that are not connected to WAN Internet. Removable storage will be large capacity hard drives that will be kept in locked closets.
Monitoring of Data Transfer. The data will therefore not be transferred outside the AudioCommons consortium.
Compliance with institutional, national, and European legislation. The protection of data collected during the research studies will comply first locally with the participants' policies and Horizon2020 guidelines.

Perceptual tests involving human participants
Regarding the collection of data during the perceptual tests, and accordance with the University of Surrey's Research Data Management guidelines:

(i) the anonymous test data will be stored digitally on the University of Surrey's computer systems, with access restricted to project team members; at the end of each project phase the data generated during that phase will be made available publicly via a recognised Research Data Repository, and will be preserved for a

minimum of 10 years from that time or, if others have accessed the data, from the last date on which access to the data was requested;

(ii) the completed consent forms will be stored for a period of 10 years in their original paper format, in a secure location at the University of Surrey, separate from the anonymous test data; at the end of that period the forms will be destroyed by the Data Archival Officer.

5.2 Security

Please indicate if your project will involve:

- Activities or results raising security issues: **NO**
- 'EU-classified information' as background or results: **NO**

References

[Aslam12] Aslam, N., Ullah, I., Rohullah, B. S., Akram, T., & Shabir, M. Tracking the Progression of Multimedia Semantics: from Text Based Retrieval to Semantic Based Retrieval. World Applied Sciences Journal, 20, 4 (2012), 549-553.

[Basu01] Basu, S., Fernald, J. G., and Shapiro, M. D.Productivity growth in the 1990s: technology, utilization, or adjustment?. Carnegie-Rochester Conference Series on Public Policy, Elsevier, vol. 55, 1 (2001), 117-165.

[Bischoff09] Bischoff, K., Firan, C. S., Paiu, R., Nejdl, W., De, L. S., Laurier, C., and Sordo, M. Music Mood and Theme Classification - A Hybrid Approach. In Information Retrieval (2009), 657–662.

[Bogdanov13] Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O., Roma, G., Salamon, J., Zapata, J., and Serra, X. ESSENTIA: an Audio Analysis Library for Music Information Retrieval. In International Society for Music Information Retrieval Conference (2013), 493–498.

[Caclin05] Caclin, A., McAdams, S., Smith, B.K., and Winsberg, S. Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. Journal of the Acoustical Society of America, 118 (2005), 471-482.

[Cannam11] Cannam, C., Jewell, M. O., Rhodes, C., Sandler, M., and D'Inverno, M. Linked Data And You: Bringing music research software into the Semantic Web. Journal of New Music Research 39, 4 (2010), 313–325.

[Cannam10] Cannam, C., Landone, C., and Sandler, M. Sonic Visualiser : An Open Source Application for Viewing , Analysing , and Annotating Music Audio Files. In Proceedings of the ACM Multimedia 2010 International Conference (2010), 1467-1468.

[Carroll09] Carroll, E. A., Latulipe, C., Fung, R., Terry, M., and Cheriton, D. R. Creativity Factor Evaluation: Towards a Standardized Survey Metric for Creativity Support. In Proceedings of the Seventh ACM Conference on Creativity and Cognition (2009), 127–136.

[Casey02] Casey, M. General Sound Classification and Similarity in MPEG-7. Organised Sound 6, 2 (2002), 153–164.

[Csikszentmihalyi97] Csikszentmihalyi, M. Creativity: Invention, Flow and the psychology of discovery and Invention. Harper Perennial, New York, 1997.

[Davis13] Davis, N., Winnemöller, H., Do, E. Y.-l., and Tech, G. Toward a Cognitive Theory of Creativity Support. In Proceedings of the 9th ACM Conference on Creativity & Cognition (2013), 13-22.

[Fazekas11] Fazekas, G., Raimond, Y., Jakobson, K., Sandler, M. An overview of Semantic Web activities in the OMRAS2 Project, Journal of New Music Research, Vol. 39. Issue 4, (2010) pp. 295–311. DOI: 10.1080/09298215.2010.536555.

[Fazekas11b] Fazekas, G., and Sandler, M. B. The Studio Ontology Framework. In 12th International Society for Music Information Retrieval Conference (2011).

[Fazekas13] Fazekas, G., Barthet, M., Sandler, M. (2013) The BBC Desktop Jukebox music recommendation system: A large-scale trial with professional users, in proc. IEEE Int. Conf. on Multimedia and Expo (ICME), 15-19 July, San Jose, CA, USA. DOI: 10.1109/ICMEW.2013.6618235

[Fielding00] Fielding, R. T. Architectural Styles and the Design of Network-based Software Architectures. PhD thesis, 2000.

[Font14] Font, F., Oramas, S., Fazekas, G., and Serra, X. Extending tagging ontologies with domain specific knowledge. In Proceedings of the International Semantic Web Conference (2014).

[Font13] Font, F., Roma, G., and Serra, X. Freesound technical demo. Proceedings of the 21st ACM International Conference on Multimedia (2013), 411–412.

[Gerhard03] Gerhard, D. Pitch extraction and fundamental frequency: History and current techniques. Technical report, Dept. of Computer Science, University of Regina (2003).

[Gomez06] Gómez, E. Tonal Description of Polyphonic Audio for Music Content Processing. INFORMS Journal on Computing 18, 3 (2006), 294-304.

[Gouyon05] Gouyon, F., and Dixon, S. A Review of Automatic Rhythm Description Systems. Computer Music Journal, 29, 1 (2005), 34-54.

[Herrera03] Herrera, P., Peeters, G., and Dubnov, S. Automatic classification of musical instrument sounds. Journal of New Music Research, 32, 1 (2003), 3–21.

[Klapuri08] Klapuri, A. Multipitch analysis of polyphonic music and speech signals using an auditory model. IEEE Transactions on Audio, Speech and Language Processing, 16 (2008), 255–266.

[Klapuri06] Klapuri, A. P., Eronen, A. J., and Astola, J. T. Analysis of the meter of acoustic musical signals. In IEEE Transactions on Audio, Speech and Language Processing, 14 (2006), 342–355.

[Kolozali13] Kolozali, Ş., Barthet, M., Fazekas, G., and Sandler, M. Automatic ontology generation for musical instruments based on audio analysis. IEEE Transactions on Audio, Speech and Language Processing, 21 (2013), 1–14.

[Kudumakis14] Kudumakis, P., Sandler, M., Anadiotis, A. C. G., Venieris, I. S., Difino, A., Wang, X., Tropea, G., Grafl, M., Rodríguez-Doncel, V., Llorente, S., and Delgado, J. MPEG-M: A digital media ecosystem for interoperable applications. Signal Processing: Image Communication, 29 (2014), 150–166.

[Lakatos00] Lakatos, S. A common perceptual space for harmonic and percussive timbres. Perception & psychophysics 62 (2000), 1426–1439.

[Levy08] Levy, M., and Sandler, M. Structural segmentation of musical audio by constrained clustering. IEEE Transactions on Audio, Speech and Language Processing, 16 (2008), 318–326.

[Mauch10] Mauch, M., and Dixon, S. Simultaneous estimation of chords and musical context from audio. IEEE Transactions on Audio, Speech and Language Processing, 18 (2010), 1280–1289.

[Milner99] Milner, R. Communicating and Mobile Systems: the Pi-Calculus. Cambridge university press, 1999.

[Nem14] NEM Vision & SRIA, Position paper published by the New European Media (NEM). http://nem-initiative.org/wp-content/uploads/2014/10/NEMVisionSRIA-PositionPaper-2014.pdf (2014)

[Paulus10] Paulus, J., Müller, M., and Klapuri, A. Audio-based music structure analysis. In International Society for Music Information Retrieval Conference (2010), 625–636.

[Peeters03] Peeters, G. A large set of audio features for sound description (similarity and classification) in the CUIDADO project. CUIDADO IST Project Report 54 (2004), 1–25.

[Raimond07] Raimond, Y., Abdallah, S., Sandler, M., and Giasson, F. The Music Ontology. In International Society for Music Information Retrieval Conference (2007), 417–422.

[Roma10] Roma, G., Janer, J., Kersten, S., Schirosa, M., Herrera, P., and Serra, X. Ecological acoustics perspective for content-based retrieval of environmental sounds. Eurasip Journal on Audio, Speech, and Music Processing, 2010 (2010).

[Salamon12] Salamon, J., and Gomez, E. Melody extraction from polyphonic music signals using pitch contour characteristics. IEEE Transactions on Audio, Speech and Language Processing, 20 (2012), 1759–1770.

[Serra13] Serra, X., Magas, M., Benetos, E., Chudy, M., Dixon, S., Flexer, A., Gómez, E., Gouyon, F., Herrera, P., Jorda, S., Paytuvi, O., Peeters, G., Schlüter, J., Vinet, H., and Widmer, G. Roadmap for Music Information ReSearch. 2013.

[Schneidermann07] Shneiderman, B. Creativity Support Tools: Accelerating Discovery and Innovation. Commun. ACM 50, 12 (2007), 20–32.

[Tao10] Tao, R., Li, Z., and Ji, Y. Music Genre Classification Using Temporal Information and Support Vector Machine. In ASCI Conference (2010).

[Terasawa05] Terasawa, H., Slaney, M., and Berger, J. Perceptual Distance in Timbre Space. In Eleventh Meeting of the International Conference on Auditory Display, 5 (2005), 61–68.

[Weyde14] Weyde, T., Cottrell, S., Benetos, E., Wolff, D., Tidhar, D., Dykes, J., Plumbley, M. D., Dixon, S., Barthet, M., Gold, N., Abdallah, S., Mahey. M. Digital Music Lab - A framework for analysing big music data. In European Conference on Data Analysis (2014).

[Wilmering15] Wilmering, T., Fazekas G., Dixon S., Bechofer, S., Page K., Automating Annotation of Media with Linked Data Workflows, 3rd International Workshop on Linked Media, co-located with the World Wide Web conference (2015).

[Yeh10] Yeh, C., Roebel, A., and Rodet, X. Multiple fundamental frequency estimation and polyphony inference of polyphonic music signals. IEEE Transactions on Audio, Speech and Language Processing, 18 (2010), 1116–1126.

[Zhang99] Zhang, T. Z. T., and Kuo, C.C. Classification and retrieval of sound effects in audio-visual data management. Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers, 1 (1999), 730-734.

[Zwicker07] Zwicker, E., and Fastl, H. Psychoacoustics: Facts and Models, 3rd edition, Springer (2007).