OPENING THE SOUND ARCHIVES OF EUROPE:
THE EUROPEANA SOUNDS PROJECT

ELISA SCIOTTI

Istituto Centrale per il Catalogo Unico
delle biblioteche italiane e per le informazioni bibliografiche (ICCU)
Viale Castro Pretorio, 105
00185, ROME, Italy

UDK / UDC: 026.06:004
DOI: http://doi.org/10.21857/mnlqgc0qky

Abstract

Europeana Sounds is a project co-funded by the European Commission within the ICT Policy Support Programme as part of the Competitiveness and Innovation Framework Programme. The project aimed to enhance the total amount of the sound contents available through Europeana. Currently, more than 600,000 audio items and more than 300,000 audio-related contents have been aggregated thanks to the activities carried out by Europeana Sounds. The project has achieved other important objectives such as the enrichment of the metadata related to the digital objects, the realisation of a set of Policy Recommendations, the launch of the first thematic channel of Europeana entirely devoted to the music named Europeana Music and a Radio that makes 200,000 music tracks available online. Europeana Sounds will continue to pursue its objectives through the Task force established within the International Association of Sound and Audiovisual Archives (IASA).

Keywords: Europeana, Europeana Sounds, Music, Sounds, Sound heritage, Edit-a-thons, Europeana music, Metadata, Digital service infrastructure, Data enrichment

Introduction

Europeana is a digital service infrastructure designed with the aim of gathering and providing access to a wealth of European cultural heritage. Over the course of several years, this large platform for cultural heritage, which has been
available for online consultation since 2008, has managed to vastly increase the resources it makes available to the public and at the same time has improved methods for searching its contents.

With the large-scale involvement of over 3,500 cultural institutions from thirty-six countries, the portal has gone from the 2,000,000 resources available upon its launch in 2008 to over 51,000,000 today: images, sounds, videos, 3D and text resources which display themselves to the users like a vast treasure chest overflowing with books, works of art, videos, music and sources of oral heritage along with much more besides.

In this brief presentation we will look at the Europeana Sounds project, which is Europeana’s fifth domain aggregator after APEX for archives, EUscreen for television, European Film Gateway for film, and TEL for libraries.

The Europeana Sounds project began in February 2014 with the intention of making as-yet unknown European sound heritage more visible and accessible, as stated in the project’s initial factsheet: »Hundreds of thousands of sounds dating back to the invention of the first audio recorders are waiting to be discovered and listened to. By creating the European online sounds library we will take the pulse of Europe!«

Under this initial mission, a consortium made up of twenty-four partners from twelve European countries was formed. Guided by the British Library, the

---

4 The EUscreen portal offers free online access to audiovisual heritage. It was built by a consortium of European audiovisual archives, public broadcasters, academic and technical partners. See http://euscreen.eu/ (access 30 November 2017).
6 The European Library (TEL) was launched in 2004 as the union catalogue of European national libraries and has since become a web portal and open data hub for national library data in Europe. TEL has been the biggest provider to Europeana bringing in more than 11 million records from European libraries. Due to the change in individual national libraries’ needs, TEL services concluded by the end of 2016. See http://www.theeuropeanlibrary.org/tel4/ (access 30 November 2017).
heterogeneous consortium consists of libraries, archives, research centres, government bodies, universities and companies.

Since Europeana Sounds’ beginning, several key points for the project, which correspond to the same number of stated objectives, have been explicitly listed. First of all, it aimed to increase considerably the amount of audio content available on Europeana while making it easier to access and enriching the descriptive metadata associated with the resources. At the same time, a clear desire was expressed to create a thematic channel that would make sound and other contents on Europeana easily accessible.

Of the initiative’s various proposed objectives, particular attention has been paid to the issue of constraints on the use of resources, and attempts have been made to identify best practices and methods to facilitate access to out-of-commerce audio content, working with music publishers and rights holders to achieve this goal.
Europeana Sounds has also devoted attention to the exciting idea of the reuse of those sound resources in creative and innovative ways.

By creating a network of cooperation between the various professionals in the field, Europeana Sounds has consistently attempted to ensure that the project is broad-based and that the accessibility of the available content is constantly increasing.

**The sound collections**

So what has the Europeana Sounds project collected over the course of its history? What kind of resources have been made available through this platform? To connect sound only with musical content would be overly simplistic.

The world of sound can be split into countless categories, and in the case of the more than 600,000 resources collected by Europeana Sounds, these have been divided into five main groups: Music, Spoken word recordings, Environment recordings, Radio programmes and Sound effects recordings.
The journey starts from the aforementioned and fundamental category of *Music*, which includes in turn a wide range of genres going from classical or art music through traditional and folk music all the way to popular music.

The collection is also enriched by the inclusion of the *Spoken word recordings* category, which includes historical speeches, interviews, stories, languages and dialects which have an absolute value in their ability to communicate a vast cultural heritage across time. We should also remember the artistic performances such as theatrical pieces and recited poems that the category contains.

Nature and the environment also contribute to this enormous wealth of material, and they can be found in the *Environment recordings* section, which collects both the sounds of animals and the ambient sounds of natural environments or big cities.

Moreover, there is the category of *Radio Programmes* which includes news, documentaries and programmes which were conceived and produced for radio broadcast and finally the *Sound effects*.

The range covered is extremely broad, running from speeches made to a crowd, an aria from an opera or the song of a bird to the telling of a story, and including also the sound of a dive into water and recited verses. To those mentioned above we must also add more than 300,000 audio-related items: printed scores, music manuscripts, pictures and much more which are hosted in direct correlation with the sound documents.

The aggregation

At this point it is helpful to look at the workflow and the technical infrastructure which underlies the aggregation of the data made available through the Europeana portal. The work package which handled the organisation and management of the various phases of aggregation was coordinated by the British Library, while the group which dealt with the technical infrastructure necessary for the collection, management and publication of data was coordinated by the National Technical University of Athens (NTUA).

The platform which allows the acquisition of data and its being made available on Europeana is called MINT - Metadata Ingestion Tool\(^{10}\) - and is managed by NTUA.

Data is uploaded onto the MINT platform by data providers, who then transform the data into the Europeana Data Model (EDM)\(^{11}\) format required by European-
ana. The data is then sent to Europeana through an OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) protocol. Transformation into EDM format is necessary for publication in the portal.

In the specific case of this project, the EDM format was adapted by creating a specific EDM profile for sound, enriched specifically to adequately represent the individual properties of the audio and its related objects.

In addition to the aggregation of audio resources and related content, the project has carried out in-depth studies of the legal issues involved with the aim of making the audio heritage more easily and consciously accessible online. The work package that handled this aspect (licensing guidelines) was coordinated by Kennisland from the Netherlands. Many factors and issues need to be considered before making an audio resource freely available online: content may, for example, no longer be commercially available but may still be under copyright, and resources whose use is limited in some countries may be accessible in others and vice versa. In the study of these issues, various questions regarding copyright and domain constraints were considered, such as the moral aspects and the protection of privacy in relation to the dissemination of a resource.

Over two years of work and discussions led to the publication of the Europeana Sounds Policy Recommendations for improving online access to audio and audio-related heritage in Europe, which was announced at the Europeana Sounds: Unlocking Sound Collections conference in Vilnius in November 2016 and which is currently available online. The document is divided into three sections and offers seven recommendations. The first section tackles the topic of the legislative framework in Europe in relation to copyright and provides three recommendations for this issue. The second section is devoted to the need to obtain – in an easier, clearer way – information on copyright holders and on the legal status of the material, and provides two recommendations in this regard. Finally, in the third section, the issue of access to sound resources in relation to extra-legal factors is addressed, and in this regard, the final two recommendations are provided.
The results of the Europeana Sounds project aimed to provide an improvement in the accessibility of resources available online.

**The promotional activities**

In addition to the aforementioned activities, over these three years Europeana Sounds also undertook increasing amounts of promotional initiatives aimed at raising awareness of the work carried out by the project and highlighting the ongoing sharing of content with all users. The promotional activities (work package dissemination and networking), coordinated by the Bibliothèque nationale de France (BnF), made it possible to engage with the public and raise awareness of the project primarily through an active blog that has grown, week by week, thanks to its posts and its emphasis on unique sound content, stories and events. The blog was joined by social initiatives, the creation of virtual exhibitions which can be visited online, the organization of events and participation in conferences, meetings and workshops throughout Europe, as well as the release of publications in various countries.

These dissemination activities gradually increased in number over the course of the project, peaking last year, and played an important role in the promotion of
one of Europeana Sounds’ most important achievements: the Europeana Music collection, the thematic channel dedicated to its music collections.\textsuperscript{15}

The work package, led by the Europeana Foundation, oversaw the creation of the first Europeana thematic channel, the definitive version of which was launched on November 4,\textsuperscript{16} 2016 (after two beta versions in August 2015\textsuperscript{17} and December 2015\textsuperscript{18}).

Europeana Music allows easy and intuitive access to the music collections hosted in Europeana and allows users to explore audio and video resources, and texts and images from various content providers related to the overall theme of music.

In addition to the already vast range of music on offer there is also editorial content curated by the various partners of the Europeana Sounds Consortium and rotated on a monthly basis since the channel’s opening. As stated in the Europeana Music launch press release: »it brings together unique music content - from some of Mozart’s earliest recordings to Smashing Pumpkin concerts,«\textsuperscript{19} a brief statement which nonetheless summarizes perfectly the diverse range of materials available.

Over the months, there have been various points of focus, including the European composers proposed by the British Library, opera arias by the National Library of Latvia (NLL), Gustav Mahler and the Vienna Court Opera and Giuseppe Verdi, proposed respectively by Österreichische Mediathek (OeM) and the Istituto Centrale per il Catalogo Unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU), musical manuscripts with particular regard to those of Salieri offered by the Österreichische Nationalbibliothek (ONB), cylinder phonographs by the Deutsche Nationalbibliothek (DNB), Greek folk music from Thrace and Macedonia by the Music Library of Greece of the Friends of Music Society (FMS), French folk music from the Bibliothèque nationale de France and the connection between music, sound and anthropology proposed by the Centre National de la Recherche Scientifique (CNRS), to name but a few and to show the transversal nature of the issues examined.

The institution in charge of the curation of the month chooses a theme to highlight a particular aspect of its own collection. Every month the music channel changes the header image, which immediately identifies the monthly topic, providing entry points to guide users through the contents available on Europeana and editorial contents to explain the theme.

\textsuperscript{19}See footnote 12.
Finally, in January 2017, a new application for users was launched: Europeana Radio. A collection of 20,000 tracks from various institutions which can be explored or listened to at random to discover the treasures possessed by Europeana’s participating partners. Besides listening to the audio tracks, users can also actively participate by tagging recognized music genres and helping to improve the retrieval of content already on Europeana Radio. In this regard it is worth mentioning the #TagDayThursday campaign promoted on social media, especially on Twitter (@EuropeanaMusic). Europeana Radio can also be embedded and made available through other websites.


In addition, a graphical representation that allows you to explore the five categories identified by the Europeana Sounds project (Music, Spoken word recordings, Environment recordings, Radio Programmes, Sound effects) even more quickly has been developed as part of the project: by clicking on each category, users can access the options available easily, making it much simpler to access the desired section.

Europeana Sounds has also arranged for direct contact activities with users: in addition to the two major conferences, »Europeana Sounds 2015: The Future of Historic Sounds« and 2016’s »Unlocking Sound Collection«, held respectively in Paris and Vilnius, various other events aimed at involving the public in a range of ways have been organized.

The (Re)discovery events have been designed for local audiences and aimed to turn the spotlight on collections or single items of particular value which each of the consortium partners has made available on Europeana Sounds. These events tackled a range of different themes but had as a common goal direct engagement of the public for the promotion of sound heritage.

In all, five of these events have so far taken place: the ONB focused on two manuscripts signed by Beethoven and Liszt, the BnF examined traditional French music of the past and present, the ICCU based its event on the collection that the Istituto Centrale per i Beni Sonori ed Audiovisivi (ICBSA) made available through Europeana Sounds, while the Netherlands Institute for Sound and Vision exhibited rare musical items from its collection and the British Library offered an evening of live music and an exhibition of old recording equipment.

Data enrichment

Another kind of event is the edit-a-thon. This activity is managed within the work package devoted to enrichment and participation coordinated by the Netherlands Institute for Sound and Vision.

Among the many aforementioned initiatives of promotion and public involvement undertaken to this end by the project, the ten edit-a-thons which have taken place over the years and which were realized thanks to cooperation between Europeana Sounds and Wikimedia are particularly worthy of mention.

The main purpose of the edit-a-thons was that of promoting the direct involvement of users with the aim of content enrichment - in other words, cooperation

26 For a definition of a edit-a-thon, see the English language Wikipedia page: https://en.wikipedia.org/wiki/Edit-a-thon, furthermore is useful to remember the definition provided during the second plenary meeting of Europeana Sounds by Sebastiaan ter BURG of Wikimedia Nederland «scheduled gathering, in person or online, where volunteers collectively write Wikipedia articles and/or improve
between institutions, Europeana Sounds and Wikipedia which made it possible to edit, improve and expand the entries in the online encyclopedia through new content - in this case, principally audio content coming from Europeana.

For example, at ICCU, an edit-a-thon was held about the well-known Italian composer Giuseppe Verdi as its central theme. The ICBSA, a content provider for the Europeana Sounds project, provided recordings of Verdi’s work for uploading to Wikimedia Commons so that they could later be used to enrich the entries on Verdi’s work in the online encyclopedia with audio material.

Over the working day, Wikipedians, music experts, librarians, scholars of this composer and volunteers met to take part and obtain mutual benefit from this day of cooperation.

The process that this kind of synergy triggers allows the enrichment of encyclopedia pages and the enrichment of the audio content itself, and this becomes part of an information context which increases its value. The initiative offers the participants - who share their respective competences - the opportunity to familiarize themselves and learn how to interact with the entries of the famous online encyclopaedia.27

This experience has been repeated many times, tackling a range of topics from various institutions, including musical events such as the ones held at the National Library of Latvia, which organized an event dedicated to Latvian musicians and composers, and at the Music Library of Greece (FMS), which devoted an edit-a-thon to traditional Greek songs.28

Enrichment consists of ensuring that the descriptive metadata linked to a digital object is as integrated and complete as possible, thus making the information relative to the object more comprehensive and exhaustive. The semantic enrichment of metadata also facilitates the retrieval of resources.

Enrichment can be carried out at different times and in different ways: directly, during the ingestion of the metadata through - for example - the use of thesauri, controlled vocabularies or data sources, in crowdsourced participatory campaigns, or through alignment, which consists of semi-automated operations of a more experimental nature.29


29 For a detailed discussion of the enrichment of metadata, the various methods of activation and the description of the activities carried out by the work group, see: Maarten BRINKERINK – Giulio ANDREINI – Breandán KNOWLTON – Hugo MANGUINHAS – Vassilis TZOUVARAS – Sergiu...
Above, we spoke of crowdsourced methods of enrichment. The key systems for semantic enrichment are Annotation API and Entity API, which make it possible to use applications like Pundit, WITH, Historypin, Europeana Radio (mentioned above) and others which are devoted to crowdsourcing.

The annotation system which the use of the above-mentioned applications allows is critical in the active participation of users who, thanks to the tools made available, can add information about audio content and context and can tag using controlled vocabularies or can georeference resources.

In this respect, we should mention the crowdsourcing campaigns which were promoted through social media with the aim of encouraging as much user involvement as possible: the first organized campaign was launched in the summer of 2016 and was aimed initially at using the Historypin geotagging application, and then continued with the promotion of Pundit which allows the recognition and tagging of musical instruments using MIMO’s controlled vocabulary on the contents of the WITH platform.

The second campaign, which began in late 2016, accompanied the project to its conclusion and saw once again the promotion of Historypin and WITH, concluding with an invitation, through the #TagDayThursday campaign, to listen to the newborn Europeana Radio and to tag all the musical genres recognized.

In light of the various project initiatives listed above, it is clear that Europeana Sounds has created a large, complex network of initiatives and products which should be supported and preserved: the vast mass of sound resources and related content available online, the Europeana Music channel, the brand new web radio, the legal study and the large-scale promotional activities and the to-
ols that allow users to enrich the data available online should all achieve sustainable continuity. 38

Conclusion

The Europeana Sounds project, co-funded by the European Commission through the ICT Policy Support Program, part of the Competitiveness and Innovation Framework Program (CIP), ended in January 2017, but this did not stop the initiative’s momentum.

The working group that has so far handled the management and sustainability of Europeana Sounds, led by the British Library, has continued to do so, and the work carried out over the last few years is being continued thanks to the Euro-

peana Sounds task force created within the International Association of Sound and Audiovisual Archives (IASA)\(^39\) and the project Europeana DSI-3.\(^40\)

Finally, the Europeana Music channel is searching for new curators outside its consortium to offer different point of view regarding the world of music.\(^41\)

Europeana Sounds was an attractive and productive project that has highlighted the relevance of the European sound heritage. It has valorised the European institutional collections that have taken part in this initiative, making their access easier and increasing the visibility of the contents provided. Europeana Sounds promoted the sound heritage but has also analysed the rights issue in order to properly use the resources and has allowed the creative re-use of the digital content in different contexts. Finally, thanks to the activities carried out within the project, experts in different fields related to the sounds heritage have come into contact creating a network of skills.

Since its beginning, Europeana Sounds has invited users to »stay tuned« to the project and even three years later, the call is still to »keep listening« in order to discover the evolution which this initiative can undergo again in the future.


\(^{40}\) https://pro.europeana.eu/project/europeana-dsi-3 (access 30 November 2017).

Sažetak

OTVARANJE ZVUKOVNIH ARHIVA EUROPE: PROJEKT EUROPEANA SOUNDS