

Andante affettuoso.

PIANO.



europeana
sounds

1. So, my Kathleen, you're go_ing to lave me, All a_lone by my_self
2. Och! them English de.ceivers by nature! Tho', may be, you'll thin

al tempo.

sure that you'll nev_er de.ceive me, Oh
say you're a charming, sweet creature



Europeana Sounds Final Public Report

www.europeanasounds.eu

Coordinated by the British Library, the Europeana Sounds project is co-funded by the European Union, through the ICT Policy Support Programme as part of the Competitiveness and Innovation Framework Programme (CIP) http://ec.europa.eu/information_society/activities/ict_psp/



European
Commission

Contents

Introduction	3
WP1 Aggregation	4
WP2 Enrichment and Participation	5
WP3 Licensing Guidelines	6
WP4 Channels Development	7
WP5 Technical Infrastructure	8
WP6 Dissemination and Networking	9
WP7 Project Management and Sustainability	10
Summary of project results	11

March 2017

Authors Maarten Brinkerink, Anouck Girard, David Haskiya, Lisette Kalshoven, Laura Miles, Tom Miles, Vassilis Tzouvaras

Introduction

Europeana Sounds was a three-year European Commission funded Best Practice Network project. Commencing in January 2014 and coming to an end in January 2017 the project work was carried out by 24 project partners, located in 12 different European countries. The Europeana Sounds Project partners were:

1	British Library	UK
2	Netherlands Institute for Sound and Vision	Netherlands
3	Kennisland	Netherlands
4	Europeana Foundation	Netherlands
5	National Technical University of Athens	Greece
6	Bibliothèque nationale de France	France
7	AIT Austrian Institute of Technology	Austria
8	Net7	Italy
9	Historypin	UK
10	Centre national de la Recherche scientifique	France
11	UAB DIZI	Lithuania
12	Deutsche Nationalbibliothek	Germany
13	Music Library of Greece of the Friends of Music Society	Greece
14	Istituto Centrale per il Catalogo Unico delle biblioteche italiane	Italy
15	Irish Traditional Music Archive	Ireland
16	The Language Archive at MPI-PL	Netherlands
17	National Library of Latvia	Latvia
18	Österreichische Mediathek	Austria
19	Rundfunk Berlin-Brandenburg	Germany
20	Sabhal Mòr Ostaig	UK
21	Statsbiblioteket	Denmark
22	Austrian National Library	Austria
23	Institute of Contemporary History, Universidade Nova de Lisboa	Portugal
24	Comhaltas Ceoltoiri Eireann	Ireland

The project objectives as laid out in the project Description of Work were:

Aggregation: provide a critical mass of digital audio tracks and supporting objects through Europeana to meet the needs of public audiences, creative industries and academic researchers.

Enrichment: support discovery and use by improving metadata through innovative methods including semantic enrichment and crowdsourcing.

Access: work with our data providers as well as publishers, the recorded music industry, rights holders and libraries to improve access to out-of-commerce¹ audio content and increase the opportunities for creative re-use of Europeana content.

Channels: enhance the existing Europeana portal by implementing a mechanism for providing channels that enable specific user communities to discover, share and annotate digital audio content and which can be extended to address other communities of interest and media.

Infrastructure: underpin the technical infrastructure required to enable the aggregation of metadata from archived digital content, primarily music and speech audio, including out-of-commerce recordings and crowdsourced content, through the Europeana portal.

Dissemination and networking: expand the work of the Europeana Sounds Best Practice Network among target audiences, acting as a catalyst for the inclusion of a significant quantity of items from collection-holders not yet engaged with Europeana.

Each of these objectives formed a thematic work package (WP) within the project, in addition to a seventh work package of project management. The project work packages and work package leaders were:

1 Aggregation British Library

2 Enrichment & participation Netherlands Institute for Sound and Vision

3 Licensing guidelines Kennisland

4 Channels development Europeana Foundation

5 Technical infrastructure National Technical University of Athens

6 Dissemination and networking Bibliothèque nationale de France

7 Project management and sustainability British Library

The project achieved all of its objectives during the three years. In the following sections of this report, each work package has reported on the successes, results and conclusions drawn from the work carried out.

¹ Out-of-commerce content: content that is still in copyright but no longer available via commercial channels.

Aggregation

One of the main aims of Europeana Sounds was to make the metadata for 500,000 sounds and 225,000 sound-related digital objects accessible on Europeana. These were ambitious targets and they were successfully exceeded.

The project's data providers used a Metadata INgestion Tool (MINT) to convert their metadata to XML files (a file extension for an Extensible Markup Language) according to the Europeana Data Model (EDM). Once transformed and published on MINT, this metadata was harvested by and published on Europeana. All data providers used MINT, which was maintained and run by the National Technical University of Athens (NTUA). All staff from the 18 data providers learnt how to operate MINT and successfully transform their metadata. Data providers' staff were not necessarily from an IT background, but often from a curatorial/cataloguing background and have therefore developed new skills which can be harnessed by their organisation in the future. During the lifecycle of the project, three training workshops were held by NTUA: two in Athens in April 2014 and June 2015 and one hosted in, Dublin in April 2016.

The uploading, mapping and transforming of metadata had to be done accurately and the British Library and staff at NTUA supported the data providers in carrying out this work. NTUA showed great commitment in diagnosing and solving these problems themselves or helping data providers to address them. It is a great credit to all involved – NTUA, the data providers and the Europeana Data Ingestion team – that this partnership resulted in 531,947 metadata records, representing over 900,000 digital objects – being published on Europeana.

The use of directly linked content – in other words, providing links to the actual address of a digital file, rather than the page in which it is embedded – was encouraged throughout the Europeana Sounds project. This ensures that the user can listen or view content from Europeana, without having to navigate back to data provider's web page and the content can additionally be included in the recently developed thematic channel, Europeana Music.

The table below outlines the key performance indicators (KPIs) for work package one. The final publication on Europeana took place at the end of January 2017.

KPI name	Target	Result
Number of audio items aggregated	500,000	616,888
Number of other items aggregated	225,000	311,050
Number of items freely available for re-use	90,000	242,727
Number of data providers using new EDM profile	100%	100%
Number of consortium partners to have made use of training resources	100%	100%

Such an outstanding result has ensured that audio heritage is far more discoverable and freely accessible online. At the start of the project audio heritage was described as 'under-represented' in Europeana; Europeana Sounds has remedied this deficit.



WP2 Enrichment and Participation

Work Package 2 of Europeana Sounds aimed to support the discovery and use of the audio and audio-related metadata in Europeana, by improving metadata through innovative methods including semantic enrichment and crowdsourcing. In order to accommodate this, WP2 designed and implemented mechanisms to improve the quality of existing metadata and contextual information.

At the heart of the system architecture is the Annotations API, which implements a REST² (representational state transfer) interface for management and administration of annotations, a developer console used as documentation and a testing and debugging environment. WP2 deployed a threefold strategy to pursue semantic enrichment of the aggregated data: enrichment during ingestion, enrichment through crowdsourcing and enrichment through alignment. The enrichments strategies informed the development of several core systems as components – namely the Entity API and Annotations API – in the Europeana core infrastructure. Crowdsourcing applications such as Europeana Radio³, Pundit⁴ and Tunepal⁵ connect with the Annotations API to allow their users to add annotations to Europeana objects. The Annotations API aimed to support the application scenario developed by the project (most prominently the crowdsourcing application), but also allowed for experimentation with various ways to align and link the data held by Europeana to external knowledge structures and resources (such as the Musical Instrument Museums Online (MIMO) and Wikidata).



Participants working during an edit-a-thon. (Picture by Brigitte Jansen CC BY)



First Europeana Sounds edit-a-thon in progress at Sound and Vision The Netherlands (Sebastiaan ter Burg CC BY 2.0)

WP2 developed mechanisms for improving four types of existing enrichments:

- Automatic enrichments that are applied during the ingestion of metadata
- Semi-automatic enrichments, resulting from the application of the ontologies by data providers
- Assisted structured enrichments by users, as a result of crowdsourcing micro-tasks
- Manual enrichments by users, as a result of crowdsourcing micro-tasks

The types of annotations developed by the project were deployed in various crowdsourcing applications: Traditional Music Pilot (object linking), Historypin (geotagging), Pundit (semantic tagging, moderation), WITH (semantic tagging, moderation) and Europeana Radio (semantic tagging). The Crowdsourcing Applications were promoted through two crowdsourcing campaigns.

WP2 had many successes during the final year of the project. A highlight enjoyed by data providers were the collaborations with Wikimedia, which resulted in ten edit-a-thons held across Europe, increasing the presence of audio heritage on Wikipedia.

² https://en.wikipedia.org/wiki/Representational_state_transfer

³ <http://www.europeana.eu/portal/en/radio.html>

⁴ <http://thepund.it>

⁵ <http://www.tunepal.org>

WP3 Licensing Guidelines

WP3 focussed on intellectual property and the barriers cultural heritage organisations face when publishing audio heritage online. These barriers include out-of-commerce works, domain-constrained works and other barriers such as those presented by possible moral and ethical rights in audio and audio-related heritage collections.

- Out-of-commerce works are works that are still protected by copyright but are no longer commercially available because the authors and publishers have decided neither to publish new editions nor to sell copies through customary channels of commerce⁶. Potential copyright issues are a big factor in choosing which analogue recordings to digitise; most partners have chosen to submit public domain works to Europeana Sounds and their out-of-commerce works are never taken into consideration. As the works are still protected by copyright, an institution cannot make out-of-commerce works available without permission from the rights holders, who are sometimes difficult to trace and may not have an economic incentive for engaging in licensing negotiations.
- The term domain-constrained works, or sometimes geo-locked works, refer to copyright protected material that has been made available for online access, but only for a specific territory. Since Europeana is intended for worldwide use, it is important that the collections are cleared for publication worldwide.
- Moral and ethical rights often refer to the rights and interests in cultural heritage which may, under national, customary and other laws, reside with the traditional custodians of such materials. An example of situations where moral and ethical rights might be in play is in recordings of religious ceremonies or potentially damaging testimonies by living people.

If cultural heritage institutions want to publish parts of their collections on their websites and afterwards share the metadata with Europeana, institutions need to have permission from the rights holders, if the digital object is still in copyright. Since budgets are limited and copyright clearance is time-intensive and (negotiating) licenses cost money, we see a trend that public domain heritage is favoured for digitisation over protected material.

The Europeana Sounds project partners identified a need for an update to the European copyright and related rights legal framework. During the project guidelines for data providers were produced, in collaboration with expert organisations, which facilitated the process of making out-of-commerce and geo-locked materials available.

While Europeana Sounds has been very successful in getting more audio heritage available online, we still want more and better access to our shared audio and audio-related heritage. The project has carried out research over the past two years in determining where the barriers to online access lay and how we can overcome them. Therefore this work package has spent a considerable amount of time collaborating on the creation of seven recommendations that, when adopted, will result in a significant improvement to online access to audio and audio-related heritage in Europe. The full recommendations can be [read here](#). It is hoped that these recommendations will be adopted by European legislators, collective management organisations and heritage institutions across Europe which will help to improve online access.

⁶ Definition based on out-of-commerce works memorandum of understanding: http://ec.europa.eu/internal_market/copyright/out-of-commerce/index_en.htm

The WP3 working group meeting in Lisbon



WP4 Channels Development

This work package enhanced the Europeana portal by contributing to the implementation of a mechanism for providing thematic collections that enable specific user communities to discover, share and annotate digital audio content, which can be extended to address other communities of interest and other media. This involved technical development and user experience design. Underpinning this is the Search API⁷ which makes the metadata in the Europeana database searchable and retrievable to Europeana Collections and any other client application with a valid API-key. The development efforts in Europeana Sounds have contributed to the development of two new databases and APIs: [Annotations](#) and the [Entity Collection](#).

The [Entity API](#) allows API clients (such as Europeana Collections) to suggest and retrieve known entities such as music genres, musical instruments, musicians and composers. The Entity API, once integrated into Europeana Collections and the Music Collections, will allow for a better and specific search and browse experience.

An important outcome of this work package has been the development of thematic collections in the Europeana portal, which are named Collections. Collections are thematic access points to the data in Europeana. The themes are subsets of data in Europeana and these cross languages, countries and data providers. Individual thematic collections are easily found and navigated on the main Europeana Collections landing page. The home page of the Music Collections was curated by the partners of the Europeana Sounds project. Every month a data providing institution would take over the role of ‘curator’ and highlight and promote content, to a particular theme chosen by the partner, such as ‘Greek folk music’ or ‘Viennese operas’, and Europeana Music Collections would be dedicated to this theme. After the end of Europeana Sounds, curation by contributing partners will be continued.

One of the goals of the project was to make content available to as wide audiences and as many users as possible. To reach this goal the project investigated and tested Spotify and SoundCloud as platforms on which to re-publish whole or parts of their collections. The research carried out during the project recommended that content should not be published on Spotify due to the very strict constraints on copyright that Spotify has set. By contrast, research determined that SoundCloud was an easy

external platform to work with and project partners shared audio and playlists via this site.

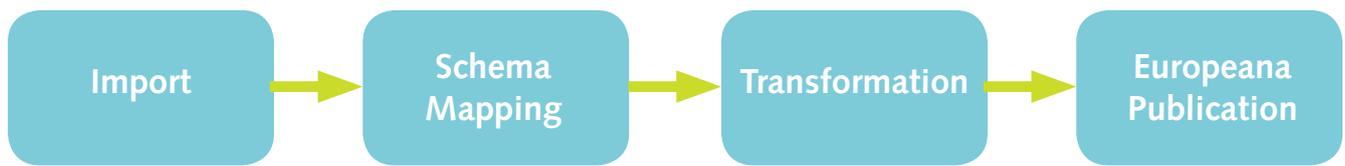
During the final year of the project [Europeana Radio](#) was also developed and launched. This new radio player feature allows a user to listen to almost 200,000 tracks collected from sound archives across Europe. In this application, listeners can browse a wide range of sound recordings (Classical, Folk and Popular music), play them on random mode and even tag the musical genres recognised. The radio player is an opportunity to increase the discoverability of institutions’ content in Europeana.



The Europeana Radio player

⁷ Search API documentation: <http://labs.europeana.eu/api/introduction> and search API source code: <https://github.com/europeana/api2>

WP5 Technical Infrastructure



MINT Ingestion workflow

This work package represented the technical infrastructure for the project. The aggregation of content for Europeana Sounds is a two-phase process. The first phase is handled by the MINT ingestion platform, the stage where data providers get involved in the aggregation process. In the second phase, transformed metadata are delivered to Europeana in the Europeana Data Model (EDM) via an OAI (Open Archives Initiative) repository and are then handled and published on Europeana website. WP5 enabled the metadata aggregation by extending and enhancing the existing Europeana aggregation infrastructure. In order to maximize the quality of delivered metadata, a number of services for metadata enrichment such as cleaning and normalization, were designed, deployed and made available to data providers. The MINT ingestion platform has been used in other Europeana related projects, making it by now one of the mature solutions for metadata handling and reuse. The EDM sounds profile specification, which is the sound enriched version of EDM used as intermediate schema before publishing to Europeana, was designed and released.

The enrichment of aggregated metadata records is approached through three main processes. The first one is that of data cleaning; typographical errors are corrected and it is possible to conform to specific conventions easily. The second methodology is data reconciliation, whereby users can align their metadata to vocabularies. A number of quality checks are performed using the preview and statistics functionalities of MINT: general structure of the provided EDM data, existence of unique and well-formed identifiers and richness of the literal values. The last methodology, which is complementary to data reconciliation, is that of linking the metadata records to external resources. This method is automatic and is available after the metadata are ingested to Europeana servers. The MINT aggregation platform facilitates the ingestion of semi-structured data and offers the ability to establish crosswalks to the reference schemas (EDM Sounds profile, EDM) in order to take advantage of a well-defined, machine understandable model. The underlying data serialization is in XML, while the user's mapping actions are registered as XSL transformations.

The operation of the aggregation was planned and monitored by WP1 in order to coordinate data provision (workflow guide, publication cycles and contingency planning). The technical infrastructure was implemented, monitored and maintained by WP5 in order to support the activities of WP1 and to implement the produced specifications and enable emerging ingestion and publication requirements. During the aggregation process, technical support was constantly provided on a Basecamp platform. The providers sought support through Basecamp records of previous discussions and posted files and documentation, or by posting new questions. All of the provider's requests were answered and resolved successfully by modifying the aggregation tool chain, or by adding new functionalities to it and providing support and guidance individually to every provider request.



WP6 Dissemination and Networking

WP6 managed the communication and dissemination for Europeana Sounds and identified three primary objectives for the Europeana Sounds dissemination and networking activities:

- Highlight and promote access to, and creative reuse of, Europe's sound and music collections
- Engage audiences with the aggregated materials and attract them to Europeana Sounds
- Build a community of interest and practice to ensure the sustainability of sound heritage, ensuring high quality audio content is contributed to Europeana after the end of the project, i.e. foster the creation of a sustainable Best Practice Network of stakeholders.

During the first year of the project, WP6 established the brand recognition attributes of the project (name, graphic identity and tagline) and its online presence (website and social media). A communication toolbox was created with posters, flyers, factsheets, slideshow templates and videos. Networking activities took place both online and offline through participation at events; raising awareness of the project, the partners' collections and the project outputs to come. In the second year, WP6 continued to strengthen the project's online presence and reach, completing the communication toolbox, intensifying networking activities and promoting the first project outcomes – the publication of the first sets of metadata onto Europeana and the alpha and beta version of Europeana Music. The highlight of the second project year was the organisation of the first international public conference, “The Future of Historic Sounds” in October 2015 in Paris.

During the project we also released communication highlights. These included a successful online exhibition about [recording and playing machines](#) and a [sound timeline](#), drawn from the collections of project partners to offer a selection of items that highlight specific significant events and key figures through history and to illustrate the role of sound and music through time. Furthermore, we also developed the [Sounds Collections](#) to provide an easy, single access point to all of the material shared by the Europeana Sounds consortium during the project.

The final year of the project was the most intense in terms of communications actions. Online activities focused on the building of a strong community on social media, as well as on the production of blog posts centred on the project's rich contents and results, including new user-friendly tools such as the Sound Collections and the crowdsourcing possibilities developed by WP2. Offline activities included participation at, and organisation of,

numerous events; particularly the second international conference “Unlocking Sound Collections” which took place in November 2016 in Vilnius – during which Europeana Music was officially launched. A full list of these dissemination activities can be found in [D6.6 Communication plan and evaluation V3](#).

The project partners out-stripped the WP6 key performance indicators, as the project was presented at 155 events and in 1,105 publications over the three years of the project.



The first international public conference, “The Future of Historic Sounds” in Paris



Europeana Sounds information leaflet and project postcards

WP7 Project Management and Sustainability

WP7 was responsible for the management of the Europeana Sounds project, including administrative management, reporting to the EC and performance monitoring. This took the form of providing leadership and guidance for the Best Practice Network (BPN) and evaluating the effectiveness of the BPN, managing priorities and risks and ensuring the quality of deliverables. We also developed an exploitation plan for sustaining the platform and services post-project and managed the ongoing evaluation process for the project. This included surveying project partners, monitoring key performance indicators, measuring statistics and commissioning an external evaluator to evaluate project performance.

WP7 was also responsible for maintaining and developing the Europeana Sounds BPN that represents both technical partners and heritage organisations who are experts in their field. The bringing together of such a wide variety of organisations created an expert knowledge pool and skill sharing has become commonplace between partners. This also encouraged an additional five Associate Partners to join the network. The Associate Partners are heritage organisations who are keen to share their audio collections through Europeana. Knowledge sharing and connections will continue after the project lifetime and for many partners this has been an important outcome of the project.

WP7 managed the planning for the post-project sustainability. The first step of this was to evaluate the key project outputs and ascertain which of these should be sustained after the project lifetime. The key project outputs include:

- Large-scale audio heritage aggregation and improving data quality
- The Europeana Music Collection, which is an excellent way of promoting and showcasing key music-related audio heritage;
- The licensing guideline work carried out by WP3;
- Advocacy including the promotion of the preservation of audio heritage, engagement with copyright work and sharing of best practices.

To continue and preserve these outputs, the International Association of Sound and Audiovisual Archives (IASA) formally agreed to continue the work initiated by Europeana Sounds with the support of a newly-created IASA Europeana Sounds Task Force. The focus for the Task Force is to: investigate sources of funding to sustain the work of project; raise awareness of the richness and threats to Europe's

sound heritage; provide expert training on providing correct metadata for adding audio heritage to the Europeana portal, via conference workshops and webinars; investigate the curation of Europeana Music for music and any other audio-related thematic channels on Europeana; copyright and moral rights relating to providing access to audio heritage via Europeana and information sharing on best practices for online audio access.

Further sustainability plans also include: the creation of a Europeana task force managed by the WP3 working group that will focus on moral and ethical rights; the ongoing availability of crowdsourcing tools to all partners; the monthly curation for Europeana Music during 2017 and onwards; and developing the Best Practice Network. The majority of data providers will continue to add material to the Europeana platform, which they will be able to do using MINT. This reflects the fact that many organisations joined the project in order to increase the visibility and accessibility of their collections through Europeana and, due to the success of the project, plan to continue to do so after the project lifetime.



The Europeana Sounds project team at the final plenary meeting in Leipzig, December 2016 (Picture by Deutsche Nationalbibliothek – Bärbel Kaiser)

Summary of project results

Overall, the Europeana Sounds project met all its objectives and achieved many successful outcomes that can be attributed to the collective effort of all project partners. Highlights of these results include:

- The aggregation of over 600,000 audio files and over 300,000 sound-related digital objects
- Ten edit-a-thons held across Europe demonstrating a powerful collaboration between Europeana Sounds partners and Wikimedia
- The development and launch of Europeana Radio
- The development of two new Europeana databases and APIs: Annotations and the Entity Collection.
- The development and public launch of the thematic collection: Europeana Music
- Mass annotation of metadata to improve the discoverability of sound and related items on Europeana
- Seven policy and licensing recommendations released which, when adopted, will result in a significant improvement to online access to audio and audio-related heritage
- Two public conferences, held in Paris and Vilnius and profile raising in numerous publications worldwide
- Successful social media accounts which have shared project results and encouraged new audiences to explore sound heritage collections
- Five new Associate Partners
- Creation of the IASA Europeana Sounds Task Force to continue the work that was initiated during Europeana Sounds.





BRITISH
LIBRARY

The coordinator of this project, the British Library, would like to take this opportunity to thank all project partners for their hard work during this period. If any further information is required about the project, the British Library can be contacted through:

Richard Ranft, Europeanana Sounds Project Coordinator:
Richard.ranft@bl.uk

Laura Miles, Europeanana Sounds Project Manager

Additional information can be found on the project website:
<http://www.europeanasounds.eu/>

Reports on public deliverables and milestones are on the Europeanana Pro website:
<http://pro.europeana.eu/project/europeana-sounds>

Europeanana Sounds is on Facebook and Twitter: @eu_sounds

Copyright notice

Copyright © Members of the Europeanana Sounds Consortium, 2014–2017. This work is licensed under the Creative Commons CC-BY License:

<https://creativecommons.org/licenses/by/4.0/>