EUROPEANA SOUNDS

Project Number: 620591

D1.1 Content Selection Policy

Document Identifier:	EuropeanaSounds-D1.1-Content-Selection-Policy-v1.2.docx
Document link:	http://pro.europeana.eu/web/europeana-sounds/documents
Date:	30/07/2014

Abstract

This document is the *Europeana Sounds Content Selection Policy* defined as Deliverable D1.1 in the *Description of Work*. This deliverable builds on the document produced for Milestone *MS1 Initial Selection Policy* and provides expanded information based on updated information from data providers.

Dissemination level		
P	Public	X
C	Confidential, only for the members of the Consortium and Commission Services	
I	Internal, only for the members of the Consortium	

Co-funded by the European Union
Europeana Sounds is coordinated by the British Library

I.   COPYRIGHT NOTICE

II.   REVISIONS

Version	Status	Author	Partner	Date	Changes
0.1	ToC	Andra Patterson	BL	13/6/2014	ToC
1.0	Draft	Andra Patterson, with input from Aude Da-Cruz-Lima, Sara Tandar, Véronique Ginouvès, Wibke Weigand, Johannes Theurer, Ute Sondergeld, Alexander König	BL	26/6/2014	First draft
1.1	Second draft	Andra Patterson, with input from Ute Sondergeld, Johannes Theurer, Alexander König, Wibke Weigand	BL	10/7/2014	Second draft, integrating comments from reviewers Richard Ranft and Henning Scholz
1.2	Final	Andra Patterson	BL	30/7/2014	Integrating comments from PMB

III.   DELIVERY SLIP

	Name	Partner/WP	Date
Document Author	Andra Patterson andra.patterson@bl.uk	BL / WP1	30/07/2014
Reviewed by	**Reviewers:** Richard Ranft Henning Scholz	BL / WP7 EF /WP4	10/07/2014
Approved by	**Coordinator & PMB**		30/07/14

IV.   DISTRIBUTION

No.	Date	Comment	Partner / WP
1	30/07/2014	Submitted to the European Commission	BL / WP7
2	30/07/2014	Posted on Europeana Pro website	BL / WP7
3	30/07/2014	Distributed to project consortium	BL / WP7

V.   APPLICATION AREA

This document is a formal output for the European Commission, applicable to all members of the Europeana Sounds project and beneficiaries. This document reflects only the author's views and the European Union is not liable for any use that might be made of information contained therein.

VI.   DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the authors named in the Delivery Slip.

VII.   TERMINOLOGY

A complete project glossary is provided at the following page:
http://pro.europeana.eu/web/guest/glossary

Further terms are defined below as required:

TERM	DEFINITION
AB	Advisory Board
APEX	Archives Portal Europe network of excellence
EC-GA	Grant Agreement (including Annex I, the Description of Work) signed with the European Commission
GA	General Assembly
PC	Project Coordinator
PI	Performance Indicator
PM	Project Manager
PMB	Project Management Board
PSO	Project Support Officer
TEL	The European Library
TD	Technical Director
UAP	User Advisory Panel
WP	Work Package

VIII.   PROJECT SUMMARY

Europeana Sounds is Europeana's 'missing' fifth domain aggregator, joining APEX (Archives), EUscreen (television), the Europeana film Gateway (film) and TEL (libraries). It will increase the opportunities for access to and creative re-use of Europeana's audio and audio-related content and will build a sustainable best practice network of stakeholders in the content value chain to

aggregate, enrich and share a critical mass of audio that meets the needs of public audiences, the creative industries (notably publishers) and researchers. The consortium of 24 partners will:

- Double the number of audio items accessible through Europeana to over 1 million and improve geographical and thematic coverage by aggregating items with widespread popular appeal such as contemporary and classical music, traditional and folk music, the natural world, oral memory and languages and dialects.
- Add meaningful contextual knowledge and medium-specific metadata to 2 million items in Europeana's audio and audio-related collections, developing techniques for cross-media and cross-collection linking.
- Develop and validate audience specific sound channels and a distributed crowd-sourcing infrastructure for end-users that will improve Europeana's search facility, navigation and user experience. These can then be used for other communities and other media.
- Engage music publishers and rights holders in efforts to make more material accessible online through Europeana by resolving domain constraints and lack of access to commercially unviable (i.e. out-of-commerce) content.

These outcomes will be achieved through a network of leading sound archives working with specialists in audiovisual technology, rights issues, and software development. The network will expand to include other data-providers and mainstream distribution platforms (Historypin, Spotify, SoundCloud) to ensure the widest possible availability of their content.

For more information, visit http://pro.europeana.eu/web/europeana-sounds and http://www.europeanasounds.eu.

IX.   STATEMENT OF ORIGINALITY

This document contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

X.   EXECUTIVE SUMMARY: CONTENT SELECTION POLICY

This document is the *Europeana Sounds Content Selection Policy* defined as Deliverable D1.1 in the *Description of Work*. This deliverable builds on the document produced for Milestone *MS1 Initial Selection Policy* and provides expanded information based on updated information from data providers.

It includes a section on Breadth of Materials, developed using information gathered in the Europeana Sounds Rights & Metadata Ingestion Survey [Appendix], sent to data providers in April 2014. The survey asked data providers to give detailed information about the content in Table 0 (Underlying content) and Table c (Content with restricted access) in the Description of Work.

In addition to outlining the content selection criteria, this document also provides guidance for the practical application of the policy in terms of minimum technical standards and metadata requirements, including guidance on locating information regarding the use of shared ontologies, appropriate rights statements and the Europeana Data Model (EDM).

TABLE OF CONTENTS

1   INTRODUCTION

This *Content Selection Policy* document builds on the work done for *MS1 Initial Selection Policy[1]*, which outlined the initial selection policy and the aggregation process. It integrates information gathered from the *Europeana Sounds Rights and Metadata Ingestion Survey* [Appendix], undertaken by data providers in April 2014. The survey asked data providers to give detailed information about the content listed in Table 0 (Underlying content) and Table c (Content with restricted access) in the *Description of Work[2]*.

All Europeana Sounds content will be accompanied by appropriate metadata records containing controlled vocabulary terms, mapped across to the Europeana Data Model. Therefore, in addition to detailing the content selection criteria, this document also provides guidance for practical application in terms of metadata requirements, including guidance on locating information regarding the use of shared ontologies, appropriate rights statements and the Europeana Data Model (EDM).

By making available 720,000 audio and audio-related objects via the Europeana portal with EDM-compliant, enriched metadata, Europeana Sounds will make a significant improvement in the quality and quantity of content accessible via Europeana. We will aggregate content in several thematic areas with significant appeal for the general public and creative industries. For academic research it is necessary to have direct access to primary resources and through this project we will provide easy access to audio and audio-related content which is vital for research, especially in the arts and social sciences. This Content Selection Policy specifies how we will make a critical mass of audio and audio-related content available via Europeana, balancing quality and quantity of content described by good quality descriptive metadata.

The primary aim is to add 500,000 sound recordings and 220,000 associated objects on Europeana, bringing the full number of audio items available via the portal to well over a million. Although it is essential to achieve this target, it is also important to ensure diversity of subject matter, genre, language and geographical spread in content selection, so that sounds and languages of Member States are represented as comprehensively as possible. Experience from members of the consortium shows that diversity of material is important for ensuring that all audiences are satisfied, even within a category which may be very broad in itself, for example classical music. The quality of the recordings is also of high importance, both in terms of format and audio quality, and also in terms of the cultural and historical significance of the material.

To ensure breadth of material in terms of genre and subject representation, we concentrate on the following main themes:

---

[1] *MS1 Initial Selection Policy* http://pro.europeana.eu/documents/2011409/23b8a083-07ee-46c3-af20-f8ea543e5869 [REF 1]

[2] *EC-GA including Annexe I* ("Description of Work") http://pro.europeana.eu/documents/2011409/8d0e9833-4608-494e-af77-681e68f8a8c8 [REF 2]

- Music
    - Classical / Art music
    - Traditional / Folk music
    - Popular music
- Spoken word
    - Oral memories
    - Languages and dialects
    - Spoken word performances
- Radio
- Environment
    - Natural sounds
    - Soundscapes

As well as audio content, this document also covers audio-related content such as music scores, which should as far as possible match available sound recordings, text files such as interview transcripts, and photographs of people or objects associated with audio content, for example musicians and audio equipment.

Section 2 *Content selection* in this document provides guidance to data providers selecting content for inclusion. However, this guidance should not be viewed as all-encompassing, as there may be categories of content that become available during the lifetime of the project which have not been anticipated, in which case a data provider should discuss the possible inclusion of this content with the Metadata Manager. Another reason for providing guidance rather than strict criteria is that this document could serve as a guide for future contributors of metadata describing audio content to Europeana after the completion of this project. If that is the case, it is highly likely that categories of material which have not been anticipated for this project will be added in the future.

The aggregation process for Europeana Sounds is divided into four main tasks: content selection, metadata preparation, metadata ingestion, and metadata curation and enrichment. This was outlined in section 3 of *MS1 Initial Selection Policy* and will be detailed in the forthcoming *Europeana Sounds Manual for Data Providers*. The infrastructure supporting the aggregation process is being developed in WP5, with *MS23 Revised aggregation design available[3]* providing details of the developments.

1.1 *Focus on audio aggregation*

This Content Selection Policy chiefly focuses on the medium of sound recordings in order to remedy a content gap. Audio content represents only 2% of the 32 million objects currently in Europeana, yet Europeana's web statistics show these are accessed ten times more than other Europeana content types such as images and texts. Prior to this project, although some national library data

---

[3] *MS23 Revised aggregation design available* http://pro.europeana.eu/documents/2011409/1e94c275-22e9-4467-98cd-3398e41b69c8 [REF 3]

providers have contributed metadata to Europeana via The European Library (TEL), they have sometimes been unable to contribute metadata for audio content to date, because of the special conditions for metadata profiling required and/or lack of resources. This project enables those institutions to overcome those barriers. For example the British Library (BL) has contributed fewer than 1,000 audio tracks to Europeana before this project, but during the course of Europeana Sounds it will add 76,000 tracks.

In the Europeana *2013 Business Plan*[4], audiovisual media - sounds and moving images - are identified as requiring special emphasis. Two related aggregation projects for Europeana, namely EUscreenXL and the European Film Gateway, focus on television and film respectively. As the current audio aggregator for Europeana, Europeana Sounds addresses the audio content gap, completing the range of audiovisual media to remedy that gap. By the end of the project, Europeana Sounds will have more than doubled the amount of audio currently accessible via Europeana.

1.2   Aggregating related content

In addition to audio content, audio-related content that contextualises and enriches audio content will also be made available via Europeana Sounds. Metadata for texts, videos and images that are related to audio content will be aggregated by some data providers. For example the Irish Traditional Music Archive (ITMA) is contributing not only audio recordings but a wide range of audio-related items of Irish traditional music covering song and dance such as books, booklets, sheet music, video recordings of Irish traditional musicians, singers and dancers, drawings, photographs and paintings.

Some providers will not contribute such a diverse range of material but will contribute items with an important relationship to other data providers' contributions: for example the Netherlands Institute for Sound and Vision (NISV), the Austrian National Library (ONB) and the British Library (BL) will provide images of printed and manuscript classical music scores. Many of the manuscript scores are in the composer's handwriting, allowing users to investigate original source material. Having multiple resources available via the Europeana portal for the same musical work will enable users to compare performance styles in different countries and track changes in interpretations and playing styles over time, while comparing with the music score.

Images of music scores made available via Europeana Sounds will also support task *T2.4.1 Linking music to scores*, a pilot focusing on the creation of links between segments of audio content and related image content (sheet music) and/or digital scores.

Many Europeana Sounds data providers will make available images of sound recording labels and covers, providing information as found on the original resource. Photographs of historical recording and playback equipment are included to provide information about the recording processes used and to display the method of playing recordings over time.

---

[4] Europeana *2013 business plan*
http://pro.europeana.eu/documents/858566/9d4632d3-3f6d-4162-ba29-27a9a739946d [REF 4]

Doc ID: EuropeanaSounds-D1.1-Content-Selection-Policy-v1.2.docx   **Date:** 30/07/2014

1.3 Complementing existing content in Europeana

Recordings of music performances will also be linked to existing visual content in Europeana, including song-sheets, scores, photographs and paintings of composers, performers and concert venues. Such contextualisation will enhance the appreciation and value of items that until now have been presented as isolated objects. Linking with existing visual content will also enhance our understanding of music performances, by showing us information such as the size and shape of a concert hall, the musical instruments used and the age of the performers.

1.4 Extensive network of data providers

The project brings together a range of institutions from across Europe that hold audio collections: national libraries, research institutes, broadcasters, music libraries and media centres. Together these will act as centres of expertise for other audiovisual repositories in Europe, stimulating future aggregation and as catalysts for aggregating additional content, through training activities under WP1, T1.5.

Among the data providers in this project are five of the ten 'Tier 1' countries highlighted in the Europeana content development plan 2013 where the gap between target and actual content provision exceeds 50%: Greece, Italy, Latvia, Portugal and the UK[5].

As well as the items coming in from Consortium members, over the course of the project Europeana as a partner will bring in additional audio content from around 15% (40) of the other providers to Europeana that are not partners in this project. This should generate at least 200,000 further items, according to the research undertaken for the Europeana V2 deliverable, *D2.1, Europeana Partner Strategy and Development Plan[6]*, in 2012. It will use the experience, knowledge and new systems generated by the Consortium partners to make sure that this data can integrate with Europeana Sounds.

Europeana and The European Library between them already provide a base of audio-related content of around 499,000 audio items and 223,000 audio-related items. This content will be updated by Europeana's ingestion and scientific coordination teams during the project to enrich the data and to comply with the requirements of Europeana Sounds as an audience channel[7].

All together, the amount of audio material available via Europeana, through Europeana Sounds plus other Europeana projects, is expected to increase from 0.5M to 1.23M items. During the course of

---

[5] Europeana Version 2, D3.8: *Collections and Metadata Analysis, Strategy and Plan 2013*
http://pro.europeana.eu/documents/866067/4b7a8158-277e-4ec3-8f7b-9913edb3077a [REF 5]
[6] *D2.1 Europeana Partner Strategy and Development Plan*
http://pro.europeana.eu:9580/documents/866067/983565/D2.1+Europeana+Partner+Strategy+and+Develop ment+Plan [REF 6]
[7] *EC-GA including Annexe I* ("Description of Work"), Part B p. 17 (source: Email from Executive Director Europeana Foundation, April 2013) http://pro.europeana.eu/documents/2011409/8d0e9833-4608-494e-af77-681e68f8a8c8 [REF 2]

Europeana Sounds   © Members of Europeana Sounds   PUBLIC   10 / 50
EC-GA 620591   Consortium

the project, measures will be put in place to ensure continued growth as the Best Practice Network (BPN) expands.

During the creation of the project consortium and from the start of the project itself we have been in touch with numerous potential data providers who do not form part of the current consortium for Europeana Sounds. As part of this project we will explore how further audio material can be incorporated into Europeana, to extend and sustain the BPN.

2   CONTENT SELECTION

2.1   *Mandatory Selection Criteria*

The following criteria must be met when selecting content for Europeana Sounds:

- Digital objects must be freely available online, worldwide, without password
- Metadata must be provided under the Creative Commons CC0 waiver[8]
- Metadata must contain link(s) to the digital objects being described

We particularly encourage inclusion of material that is made available for re-use with minimal or no conditions, specifically those objects labelled Public Domain, CC0, CC-BY and CC-BY-SA. Additional information about criteria for inclusion of digital objects in Europeana's Content reuse framework can be found in Europeana Sounds *MS2 Rights labelling guidelines[9]*.

2.2   *Underlying content*

Broad details of the underlying content selected for Europeana Sounds have been described in the Europeana Sounds DoW, Part B, Table 0 (pages 22-27)[10], and are reproduced below, with minor adjustments to the labels of some columns:

---

[8] Creative Commons CC0 waiver http://creativecommons.org/about/cc0 [REF 7]

[9] Europeana Sounds *MS2 Rights labelling guidelines* http://pro.europeana.eu/web/europeana-sounds/documents/-/document_library_display/0Hv5/view/2089168 [REF 8]

[10] *EC-GA including Annexe I* ("Description of Work"), Part B, p. 22-27 http://pro.europeana.eu/documents/2011409/8d0e9833-4608-494e-af77-681e68f8a8c8 [REF 2]

Table 1: Underlying Content (Table 0 in Dow)

Provider	Quantity & type	Subject matter	Format	Current IPR (content)	Current use outside of Europeana	Current status of existing metadata	Metadata language
BL	79,000 audio files	Classical music, traditional music, language / dialect, oral memory, natural sounds, including 'Endangered Archives' collections	MP3	Rights cleared for free online access; 10% for download and re-use	50,000 files online, remainder to be added in 2014-2015. General public worldwide: 1,200,000 page views / year; UK academic users: 270 UK universities and colleges	MARC21-lite (XML exportable)	English
BL	36,000 images	Classical music manuscripts	JPG	Rights cleared for free online access	Widely used accessible online to scholars and general public at Early Music Online -5,000 monthly visits	50% in AACR2 MARC21 format; 50% ISAD(G)	English
NISV	6,170 audio files	Classical music, nature sounds, oral memory	WAV, MP3	Rights cleared for free online access		100% in proprietary format, mappable to Dublincore	Dutch

Provider	Quantity & type	Subject matter	Format	Current IPR (content)	Current use outside of Europeana	Current status of existing metadata	Metadata language
NISV	50 video files	Oral memory / newsreels	MPEG	Rights cleared for free online access	None, will be made available for the first time through Europeana Sounds.	100% in proprietary format, mappable to Dublin core	Dutch
NISV	10,000 images	Music-related stills	JPG	Rights cleared for free online access and download	Partly available through the Sound and Vision wiki. One million page impressions every year.	100% in proprietary format, mappable to Dublin core	Dutch
NISV	80,000 images	Popular live music recordings from the Internet Archive's Live Music library via the Dutch national aggregator operated by NISV.	Lossless downloadable	Rights cleared for free online access and download	Available through Archive.org. Several million page impressions annually. Registration is not required.	100% in proprietary format, mappable to Dublin core	English

Provider	Quantity & type	Subject matter	Format	Current IPR (content)	Current use outside of Europeana	Current status of existing metadata	Metadata language
NISV	4,000 text files	Digitised music scores from Musopen.org via the Dutch national aggregator operated by NISV.	PDF	Rights cleared for free online access	Available through Musopen.org	100% in proprietary format, mappable to Dublin core	English
BnF	10,000 audio files	Classical Music, Traditional Music, Language / Dialect	MP3	Rights cleared for free online access	Some available via Galicia	100% Dublin Core	French
BnF	10,000 image files of disc labels	Classical Music, Traditional Music, Language / Dialect	JPG	Rights cleared for free online access	Some available via Galicia	100% Dublin Core	French
SMO	38,000 audio files	Traditional Music, Language/Dialect, Stories, Oral Memory	MP3	Rights cleared for free online access	General public worldwide: approx 84,000 views/year	100% in proprietary format, mappable to Dublin core	English, Scots, Gaelic

Provider	Quantity & type	Subject matter	Format	Current IPR (content)	Current use outside of Europeana	Current status of existing metadata	Metadata language
ITMA	1,200 audio files	Audio recordings of Irish traditional musicians & singers, 1890s–2010s	MP3	Rights cleared for free online access	Users: General public, national & international. ITMA current online annual footprint is 240,000.	100% Dublin Core	English & Irish
ITMA	1,200 texts	Printed items of Irish traditional music, song & dance: books, booklets, sheet music, ballad sheets, etc., 1800s–2010s	PDF	Rights cleared for free online access	Users: General public, national & international. Current online use as above.	100% Dublin Core	English & Irish
ITMA	1,200 video files	Video recordings of Irish traditional musicians, singers & dancers, 1960s–2010s	MP4	Rights cleared for free online access	Users: General public, national & international. Current online use as above.	100% Dublin Core	English & Irish

Europeana Sounds
EC-GA 620591

© Members of Europeana Sounds
Consortium

PUBLIC

16 / 50

Provider	Quantity & type	Subject matter	Format	Current IPR (content)	Current use outside of Europeana	Current status of existing metadata	Metadata language
ITMA	1,200 image files	Drawings, photographs, paintings etc. of Irish traditional musicians, singers, dancers, 1500s–2010s	JPG	Rights cleared for free online access	Users: General public, national & international. Current online use as above.	100% Dublin Core	English & Irish
DNB	500 audio files	Classical Music, Traditional Music, Language/Dialect from phonograph cylinders or shellacs	WAV	Rights cleared for free online access	Currently none, being in preparation for online access mid-2014	MARC21	German
NLL	2,000 audio files	Classical Music, Traditional Music, popular, dance and film-music of the first half of the 20th century	MP3	Rights cleared for free online access	None (planning to launch the collection at the end of October 2013)	100% Dublin Core	Latvian

Provider	Quantity & type	Subject matter	Format	Current IPR (content)	Current use outside of Europeana	Current status of existing metadata	Metadata language
NLL	2,000 image files	Classical Music, Traditional Music, popular, dance and film-music of the first half of the 20th century	JPG	Rights cleared for free online access	None (planning to launch the collection at the end of October 2013)	100% Dublin Core	Latvian
MPG	13,563 audio files	Traditional Music; Language/Dialect, Oral Memory	WAV	Rights cleared for free online access		XMLbased IMDI format (moving to CMDI)	English (some Spanish, Portuguese, Russian)
DIZI	10,000 audio files	Traditional Music	MP3	Rights cleared for free online access		CC0 licence. Metadata format: in house, possible to map to DC or ECE	Lithuanian

Provider	Quantity & type	Subject matter	Format	Current IPR (content)	Current use outside of Europeana	Current status of existing metadata	Metadata language
FMS	2,500 audio files	Classical Music, Traditional Music	WMA, WAV	Rights cleared for free online access 84% (16% free online access only for the metadata)	25,000 users for the 16% of the content; 250 limited subscribed users for the 16% of the content	CC0 licence: proprietary 84%, 16%Mods schema	Greek (some English)
FMS	3,300 digitised photos	Traditional Music	JPG	Rights cleared for free online access	250 limited subscribed users	CC0 licence: proprietary 100%	Greek (some English)
FMS	500 video files	Traditional Music	WMV, MPEG	Rights cleared for free online access	250 limited subscribed users	CC0 licence: proprietary 100%	Greek (some English)
FMS	2,830 sheet music files	Traditional Music	RTF	Rights cleared for free online access	250 limited subscribed users	CC0 licence: proprietary 100%	Greek (some English)

Provider	Quantity & type	Subject matter	Format	Current IPR (content)	Current use outside of Europeana	Current status of existing metadata	Metadata language
CNRS	37,600 audio files	Traditional Music, Language/Dialect, Natural Sounds, Oral Memory	MP3	Rights cleared for free online access (35%)	250 subscribed users; General public worldwide: approx 34,000 users/year	Dublin Core	French
SB	11,900 audio files	Classical Music, Traditional Music, Uncategorised (other music)	WAV, MP3	Rights cleared for free online access for parts of the collection	None	Dublin Core	Danish
OeM	12,000 audio files	Classical Music, Traditional Music, Oral Memory including radio news	MP3	Rights cleared for free online access		Dublin Core	German
ICCU	80,000	Classical Music, Oral Memory, Uncategorised		Rights cleared for free online access		100% MAG metadata (Italian library standard)	Italian

Provider	Quantity & type	Subject matter	Format	Current IPR (content)	Current use outside of Europeana	Current status of existing metadata	Metadata language
RBB	105,000	Music: 30% classical; 30% World and traditional; 20% Jazz; 20% alternative Rock / Pop (30-second clips)	MP3	Rights cleared for free online access		ESE or EDM	English
ONB	150,000 image files	Music manuscripts from composers from 17-19th century (Mozart, Salieri, Haydn, Schubert, imperial court music chapel etc.)	JPG	Rights cleared for free online		MAB2 (or DC or simple MARCXML)	German

Provider	Quantity & type	Subject matter	Format	Current IPR (content)	Current use outside of Europeana	Current status of existing metadata	Metadata language
FCSH	39,000 Audio files	Classical music, traditional music, language and dialects, oral memory, uncategorised sounds	MP3	Rights cleared for free online access		UNIMARC (XML exportable)	Portuguese
FCSH	2,700 texts	Classical music, traditional music	PDF	Rights cleared for free online access		UNIMARC (XML exportable)	Portuguese
CCE	15,800 audio and video files	Irish Traditional music	MP3 audio and H.264 video	Rights cleared for free online access		Dublin Core based	English / Irish

2.3   Content with restricted access

In addition to the content outlined in Table 0 in the previous section (2.2) that will definitely be added during the project, several data providers have significant collections of digital content that are online with access currently restricted, but may become accessible through Europeana if access conditions can be altered. These include collections that are online but only on-site to listeners within the premises of their institutions (DNB, BL), online to national users only (SB, DIZI), or behind password protected services (SB, BL and others). As part of the work being undertaken in WP3 we expect to be able to make available some of the collections identified in the table below. An example is the music collections of SB which could deliver around 250,000 additional tracks of contemporary music to Europeana Sounds via the website BibZoom as 30-second snippets free to play for everyone. Together these additional restricted collections amount to over 1.5m items. We expect much more 'hidden' content of this kind will be revealed during the project. The WP3 Rights Workshop held in June 2014[11] provided a forum to begin to investigate the reasons behind access restrictions for audio content and data providers will work together during the project to investigate possible solutions.

Table 2: Examples of restricted access content from data providers[12]

Partner	Quantity of audio & associated content	Subject	Current constraints
DNB	6, 500 digitised photographs; 400,000 digitised text documents; 2,250 digital audio files	Archive of German Democratic Republic's Association of Composers and Musicologists 1949-1990	Out-of-commerce
DNB	450,000 digital audio files	All music published in Germany since 2010	In commerce
SB	500,000 digital audio files	Classical Music, Traditional Music, Uncategorised (other music)	In commerce / out-of-commerce
SB	1m hours	Radio broadcasts	Geo-locked: Danish listeners only
CCE	13,285 digital audio files	Traditional music	Geo-locked: audible watermarks in certain countries

---

[11] http://www.europeanasounds.eu/news/discussing-our-barriers-to-online-access [REF 9]
[12] *EC-GA including Annexe I* ("Description of Work"), Part B, p. 18
http://pro.europeana.eu/documents/2011409/8d0e9833-4608-494e-af77-681e68f8a8c8 [REF 2]

CCE	300 digital audio files	Oral memory	Geo-locked: audible watermarks in certain countries
BL	2,700 digitised audio files	Classical music 78 rpm recordings	Geo-locked: EU access only
BL	40,000 digital audio files	Traditional music	Password protected: UK university access only
BL	40,000 digital audio files	Classical and traditional music, spoken word, oral memory	Out-of-commerce and orphan works
DIZI	10,000 audio files	Lithuanian traditional music	Geo-locked to Lithuania
NISV	14,000 digital audio files	Classical and traditional music	Out-of-commerce
TLA	3,900 digital audio files	Language files	Password protection for academic access only

2.4 *Europeana Sounds Rights & Metadata Ingestion Survey*

In April 2014, WP1 and WP3 Leads sent out a combined survey *Europeana Sounds Rights & Metadata Ingestion Survey* [Appendix], which asked data providers about their rights clearance practices, use of identifiers, issues relating to the Europeana Licensing Framework and the contents of their digital collections. Of particular relevance to content selection, the survey asked data providers to give detailed information about the content in Table 0 (Underlying content) and Table c (Content with restricted access) in the *Description of Work*, asking data providers to assign a current rights statement, a predicted final rights statement and a projected metadata ingest set number for each collection they will contribute.

The completed surveys provided the Europeana Sounds Metadata Manager with valuable information for planning the three phases of metadata ingestion, confirming that the figures specified in the Performance monitoring table on page 91 of Part B in the *Description of Work* will be reached or exceeded.

As well as providing valuable information for planning the aggregation management for task *T1.4 Aggregation Management*, which begins in August 2014, information from the surveys has been

integrated into this document, with section *2.6 Breadth of materials* containing specific examples of content selected for Europeana Sounds.

2.5   Quality of recordings

In total 70% of the sound recordings currently available via Europeana are previews of longer recordings[13]. Most of the recordings which will be made available via Europeana Sounds will be full-length recordings, including individual tracks lasting in some cases for more than one hour. While it is difficult to estimate the duration of existing and new content, this project is expected to more than quadruple the number of hours of audio currently available via Europeana. Although it is desirable to provide access to full-length content, this is not always possible due to licensing restrictions. For content in these categories, in particular for recent recordings, some data providers have negotiated access to short clips, or "snippets", and these are included in Europeana Sounds in order to achieve diversity of content - if audio content is in copyright it is preferable to include snippets rather than not including the content at all.

We recommend that audio recordings be made available at high quality at a minimum sample rate of 44.1kHz (44,100 samples per second), a bit-depth of 16 bits or more, and in mono or stereo according to the source recording, in one of the following formats:

- Lossy compressed MP3 files at a minimum bit-rate of 128 kbps
- Fully uncompressed, 'mastering quality' lossless file formats such as WAV, FLAC or APE

Although the digital objects must be able to be accessed freely by anyone worldwide, not all digital objects will be available for re-use due to licensing restrictions. For digital objects that can be made available for re-use via Europeana Sounds, it is important to note that the technical parameters specified in the previous paragraph are also the minimum required to comply with Europeana's Content Re-Use Framework that will be implemented in late 2014-2015, based on the requirements in Europeana Creative *D3.1 Specifications for Implementing the Content Layer of the Extended Europeana Licensing Framework*[14]. For more information on this, see Europeana Sounds *MS2 Rights labelling guidelines*[15]. Practical information about the technical aspects of how reusable objects should be provided will be detailed in a separate Europeana Sounds Manual for Data Providers as this project progresses.

---

[13] EC-GA including Annexe I ("Description of Work"), Part B, p.19 (source: 2013 statistics provided by Europeana) http://pro.europeana.eu/documents/2011409/8d0e9833-4608-494e-af77-681e68f8a8c8 [REF 2]

[14] Europeana Creative *D3.1 Specifications for Implementing the Content Layer of the Extended Europeana Licensing Framework* http://pro.europeana.eu/documents/1538974/62080954-342a-4731-b592-69abc199b5fb [REF 10]

[15] Europeana Sounds *MS2 Rights labelling guidelines* http://pro.europeana.eu/web/europeana-sounds/documents/-/document_library_display/0Hv5/view/2089168 [REF 8]

2.6 Breadth of materials

Diversity of material is important in ensuring that all audiences are satisfied, even within a category which may be very wide in itself e.g. classical music. The quality of the recordings is also of high importance, both in terms of format and audio quality, and also in terms of the cultural and historical significance of the material.

Content made accessible via Europeana Sounds includes items which represent the breadth and richness of Europe's cultural capital. Some examples of iconic sounds from the collections were gathered in a series of concerts which formed part of the Kick-off event for the project in February 2014 and illustrate the broad range of subjects and musical genres that will be made accessible during the project. These are tabulated below:

Table 3: Examples of content showcased at the Kick-off meeting

Title	Description	Provider
Dance songs from Petrota, a prefecture of Evros, Thrace, Greece.	A rare piece of traditional music, representative of the music tradition of this acritic region of Greece, and unique because of these geographical parameters. It is performed by a team of women that sing and dance dressed in traditional costumes. The field recording was made in 2000 in the context of the research programme "Recording, study and promotion of the Thracian heritage in music and dance", conducted by Music Library of Greece "Lilian Voudouri" of the Friends of Music Society.	Video recording (Selection from the Friends of Music Society)
Gull in a Gale https://soundcloud.com/the-british-library/gull-in-a-gale	You can't help but sympathize with the young Herring Gull featured in this recording. The typical calls we associate with this confident seaside regular have been replaced by some rather pathetic cries as the individual is constantly buffeted by an unrelenting wind. Conditions must have been absolutely atrocious!	Recorded in England by Lawrence Shove. (Selection from the British Library)

Sora ancestor song https://soundcloud.com/the-british-library/music-from-india-sora-ancestor	The musical structure of the performance is interesting: the lead singer begins the song with the second singer following just seconds behind. This results in an 'echo' effect that seems unique to the Sora, though a similar style is used in forms of both North and South Indian classical music where the accompanist (sarangi or violin for example) follows the singer.	Gorjang village, Orissa. (Selection from the British Library)
Tututū, abełėle	Three female singers from Slavičiškis village, Kupiškis district, Lithuania. Recorded in 1972.	Traditional (Selection from DIZI)
Biržiečių untyte (penkiese)	Wooden multi-pipe whistles played by five male players, recorded in Biržai town, Lithuania 1973. Both samples are typical of Sutartines polyphonic traditional music kind, which is acknowledged by UNESCO as a proclaimed masterpiece of Intangible Cultural Heritage.	Traditional (Selection from DIZI)
Zéphirin Catellon, "siblaïre" from the South of France	Zéphirin Castellon was singer, siblaïre (piper), bell ringer and songwriter, born in 1926 in the village of Belvedere, in the valley of Vésubie (hinterland of Nice). The collection includes occitan music, polyphonic songs, songs of conscripts, farandoles, etc.	Traditional (selection from CNRS - MMSH)
Oral memory of the Armenian neighborhoods of Marseille and Bouches-du-Rhône	120 interviews with the families of descendants of Armenians living in 18 communes of the Bouches-du-Rhône since 1920	Oral memory (selection from CNRS - MMSH)
Melanholiskais valsis (Melancholic waltz)	This is the only surviving work of orchestral music by Latvian composer Emils Dārziņš, who has composed also many choral and solo songs. He destroyed his other symphonic works because he was accused of plagiarism, and ended his life when he was just 34, apparently throwing himself under a train.	Emīls Dārziņš, 1875-1910 (Selection from the National Library of Latvia)

Latvian folk song "Pūt, vējiņi" (Blow, wind, blow) https://soundcloud.com/europeana/put-vejini-blow-wind-blow	Arranged by Andrejs Jurjāns for 4-voice mixed choir a capella, this is a truly iconic song in Latvian culture. During the Soviet occupation this song served as unofficial anthem of Latvia, and has been a staple in the choir repertoire and in the Latvian Song and Dance Festival throughout the years.	Traditional (Selection from the National Library of Latvia)
Documentation of Tauwema village https://soundcloud.com/europeana/in-tauwema-village	Ambient sounds and some small pieces of conversation in the language Kilivila. This is a selection from the original recording which is about an hour long, made by Gunter Senft, as part of his language documentation project about the language Kilivila, spoken on the Trobriand Islands (located near Papua New Guinea).	Filmed by Gunter Senft (Selection from the Max Planck Gesellschaft)
Kilivila language	An adult talking to school children about their future education in the Kilivila language. This was also recorded by Gunter Senft in 2003 as part of the Kilivila language documentation project.	Filmed by Gunter Senft (Selection from the Max Planck Gesellschaft)
Stanley Robertson talks about changes in detail	This recording comes from the School of Scottish Studies Archive at the University of Edinburgh, the source of the majority of the recordings that can be heard on Tobar an Dualchais. As a Scots recording given by a Traveller, it also represents the substantial amount of material in that genre. It was recorded by Alan Bruford on reel in 1978.	Recorded by Alan Bruford (Selections from Tobar an Dualchais)
Kate MacMillan sings "Latha Dhomh 's mi Buain a' Choirce"	From the Canna collection held by the National Trust for Scotland. It is a traditional Gaelic song recorded by John Lorne Campbell on wire in 1949 and represents the thousands of Gaelic songs that can be heard on Tobar an Dualchais.	Recorded by John Lorne Campbell (Selections from Sabhal Mor Ostaig)

The sounds and sound-related material selected for this project embrace a range of popular genres and subjects: classical music and contemporary performances with timeless and universal appeal;

traditional and folk music and storytelling; sound effects, environmental sounds and noises of the natural world; languages, accents and dialects and oral recollections, all with particular resonance in different regions. Together these reflect the diverse cultures, histories, languages and creativity of the peoples of Europe over the past 130 years.

Sections 2.6.1-2.6.5 provide content selection guidance for each category of material, with examples of material already selected for inclusion in Europeana Sounds. This guidance should not be viewed as all-encompassing as there may be categories of content that become available during the lifetime of the project which have not been anticipated, in which case a data provider should discuss the possible inclusion of this content with the Metadata Manager.

2.6.1   Music recordings

Classical / Art music

Classical or Art music includes western classical/art music as well as classical/art music from non-western cultures. It is art music following long-established principles or rule systems, often supported by patrons[16].

This category includes content that exemplifies a data provider's national classical/art music tradition, content that exemplifies European scholarly endeavours in collecting classical/art music of non-western cultures and content that demonstrates the technical development of sound recordings.

Western Classical / Art music

Western classical/art music includes music from the early medieval period to the present, comprising secular as well as liturgical music. It includes modern and contemporary music such as serialism, electronic art music, experimental music and minimalist music. Western classical music may include music which crosses traditional boundaries of categorisation. This is particularly true of contemporary music, which might have attributes of several music genres.

Free access to full-length audio recordings of classical music will be provided via Europeana Sounds when a recording is free of licensing restrictions, so there will be a high proportion of recordings of older and established composers and performers, in which the copyrights have expired. Due to copyright, performing and recording rights issues, some classical music made accessible via Europeana Sounds will be provided as audio snippets.

The range of western classical music includes in-depth examples from major European composers and performers as well as works derived from this tradition regardless of their geographical origin.

Some examples of western classical/art music already selected for Europeana Sounds are:

---

[16] Grove Music Online *India, §I: The region: cultural context and musical categories* [REF 11]

- The Ruben cylinder collection, comprising Denmark's oldest sound recordings, consists of 160 phonograph cylinders, among the oldest in the world, recorded in Copenhagen between 1889 and 1897 and represent unique recordings from the period with great Danish actors and singers of the day such as Peter Schram (with possibly the very first recording from Mozart's *Don Juan*), Louise Phister, Frederik Jensen and Niels Juel Simonsen (SB)
- The significance and pan-European appeal of some of Europe's primary composers and performers is highlighted with, for example, early recordings of Guiseppe Verdi's operas, notably from 1952 of Maria Callas in *La Traviata* (NISV) and of the ever popular Enrico Caruso (ICCU)
- Digital audio files made from wax cylinders and piano rolls, such as the compilation *Der unsichtbare Pianist* (DNB)[17]
- Digital audio files made from vinyl LPs, such as a recording of Clara Haskil with the Winterthur Symphony Orchestra playing the Rondo from Beethoven's Piano Concerto no. 3 in C minor, op. 37 (NISV)[18]

Non-western Classical / Art music

Classical or art music from non-western cultures is included in Europeana Sounds. Some of this content exemplifies European scholarly endeavours in collecting and studying classical/art music of non-western cultures.

Examples of non-western classical/art music already selected for Europeana Sounds include:

- Recordings of Indian tablā performances from the James Kippen North Indian collection (BL)
- A 1958 field recording Birmanie. *Musique classique et populaire* published as a 45rpm vinyl disc in 1961 (CNRS)

Traditional / Folk music

Content selected for Europeana Sounds includes a particularly rich collection of traditional/folk music from around the world. We define traditional/folk music as a generic term for music which has a strong collective dimension with social and ritual functions within a group of peoples.

For performers studying and practising traditional/folk music, it is essential to listen to source material in the form of field recordings because traditional and folk musicians rely heavily on imitation and oral memory. The provision of recordings in this category supplies a rich and vital resource for musicians active in this area.

---

[17] http://d-nb.info/1000273466 [REF 12]
[18] http://www.eclap.eu/portal/?q=nl#axoid=urn:axmedis:00000:obj:d6c32cf4-d4da-4ead-acdd-cc0c4ac74cd5&section=search_multilang_filtered_result&cd=4&qid=76ec8eab829f65068c8b8de99dcc6522&n=4088 [REF 13]

Traditional/folk music creates a feeling of membership for particular groups of people and supports national identity. This genre has a strong oral tradition; generally it is not notated music, and hence it does not include art music, such as Indian classical music. Since the content is collective and has evolved through the ages until the present, it is usually difficult to identify a composer or an author of a traditional or folk music work.

Sometimes recordings in this category are created in the context of the original environment during field trips, or they could be recorded out of context for specific uses such as preservation, research, for exhibitions or for concerts. Because of this, these recordings not only document the richness and diversity of Europe's cultural heritage but also provide a view of European exploration, scientific, business and scholarly endeavours over the past 130 years.

Some examples of traditional/folk music already selected for Europeana Sounds are:

- The earliest recordings and related material from the Outer Hebrides, dating from 1937 (TAD)
- Contemporary recordings of Acritic songs from Thrace in Greece that point back to the Byzantine Empire of the 9th century (FMS)
- Recordings from the earliest British scientific expedition to the Torres Strait in 1898 that at once established the scholarly discipline of anthropology in the United Kingdom and documented sacred and secular songs from the islands that have long since ceased to be performed (BL)
- Iconic recordings of Irish traditional music, such as the McDonagh Brothers of Ballinafad collection and early recordings of piping legend Willie Clancy. Unique to the Comhaltas collection, these recordings also follow the evolution of traditional tunes to the present day, showing the contemporary practice of Irish performers in detailed videos that have attracted millions of online viewers (CCE)
- Recordings and documents relating to the Bellaccord Electro Company that operated from 1930 to approximately 1944, informing us not only of Latvian folk music traditions but also of important record business activity. Within this collection, the Latvian song, 'Kara laiku brūte' (War-time bride), is a characteristic soldier's song of the first half of the 20th century, reminding us of Latvia's involvement in both World Wars. These recordings also highlight Europe's past by conjuring up memories of major, pan-European events (NLL)
- Recordings of the Exposition Coloniale Internationale which took place in France in 1931. The collection gives an overview of the history of Europe's colonial past, covering a wide geographic area (BNF)
- Material from the audio archives of the Research Centre of Ethnomusicology, containing field recordings from all continents. Highlights include one of the most important collections of wax cylinders from this institution, the records of the Exposition Universelle which took place in Paris in 1900, which document the linguistic and ethnographic richness of the early 20th century from around the world. Additionally, material from the audio archives of the

Phonothèque Maison méditerranéenne des sciences de l'homme will be made available, which focuses on the south of France and the Mediterranean area. Highlights include recordings of songs and oral traditions from the Cévennes and from the region of Nice, made in the 1970s (CNRS)

Popular music

We define popular music as a generic term for music of all eras and from all cultures that appeals to popular tastes.

It includes the genre pop music and its various subgenres. Since the 1960s, pop music labels have produced an internationally established variant of African-American music that originated in the context of adolescent subcultures, is electro-acoustically processed and is disseminated by the mass media. The definition is constantly expanded and challenged by innovations in the practice of pop music production. Pop music hosts many artistic trends and is often attributed to generations (Beat Generation, Punk generation, HipHop generation, etc.) rather than to narrowed-down genres. It is in a constant flow of diversification and renewal.

Europeana Sounds will contribute popular music from a variety of genres to the pop music cluster in Europeana. It will allow an open classification grid for more genres and diversifications than will be represented by the Europeana Sounds partners in the project's lifespan.

Some examples of popular music already selected for Europeana Sounds are:

- Chansons from well-known stars such as Édith Piaf and Georges Brassens (BNF)
- Recordings of the 1920s, with Marlene Dietrich and Margo Lion as most famous cabaret singers of the period (DNB)
- Greek 'rebetiko' urban folk music performers of the 1930's and 1940's, Markos Vamvakaris and Yiorgos Batis (FMS)
- Recordings of contemporary music including jazz, rock/pop, and electronic music from many artists. A variety of regional mainstream styles with Asian, African, Oceanian or South American affiliation will be included (RBB)
- Recordings of revue artists Louis Davids and Willy Derby, and, in collaboration with the Internet Archive, a collection of live pop and rock music recordings (NISV)

2.6.2 Spoken word recordings

Spoken word audio includes recordings of people who are well-known public figures, people who are not well-known and people who choose to remain anonymous or whose names have not been recorded.

Field recordings of spoken word material can serve to illustrate fields poorly covered by conventional sources, can enhance standard sources by offering a local point of view and can

provide information from people who were present to witness an event or who have acquired local knowledge.

Although spoken word recordings can be transcribed, a transcription cannot replace live recorded voices with tone and emotion, accents and styles of delivery: hesitations, silences, bursts of laughter, inflections and emphasis.

Oral memories

This category includes oral memories of national events, interviews with people from all walks of life about daily life from times gone by, or interviews with national figures.

The development of techniques, skills and other types of knowledge has come from experiences that have been passed down from generation to generation, often by word of mouth. Examples of knowledge and skills acquired by word of mouth include handicrafts, cooking recipes, medical formulae, breeding methods and building skills. For a long period of time in our modern society, the prevailing cultural model has not been based on direct exchange and cultures that still use the oral medium to pass on knowledge and skills are becoming marginalized. Oral memory sources give an additional function to sound archives, which is to back up a transmission channel which may soon disappear.

Oral memories give evidence of past and present societies, of what life is made of: customs (rhythms, rites, work, merrymaking, entertainment) and also upheavals (immigration, drifting from the land, setting up of industrial areas, wars in the 20th century). They tell of professional and entertaining experiences, forms of social behaviour, myths and beliefs, all through personal testimonies. Through oral memories, different people can contribute their experience of the same theme, providing new insight and information. Oral memory collections often include tales, sayings and jokes recorded within an interview which provide insight into the life and community of the speaker.

Some examples of oral memories already selected for Europeana Sounds are:

- A 1912 Paris recording of Alfred Dreyfus speaking (BNF)
- The Jubilee disc released for the event of the silver throne jubilee of the German emperor Wilhelm II in 1913 (DNB)
- Margaret Thatcher discussing the moral dimension of politics in 1969 in London (BL)
- The pioneering oral history collections of George Ewart Evans in England (BL)
- Thousands of recordings covering daily news, culture and politics, an essential source for the political, social and cultural development, public opinion and lifestyle of 1960s life in Denmark (SB)

Languages and dialects

This category includes languages, dialects and examples of accents from around the world. Many of these language collections contain also ritual music, oral history, and recordings of special cultural events.

Some examples of languages & dialects recordings already selected for Europeana Sounds are:

- Accents, dialects and languages from the UK, Ireland and most parts of the English-speaking world and of non-native speakers speaking English (BL)
- Samples of world languages, including the Documentation of Endangered Languages collection, with many retellings of old legends, descriptions of rituals or traditional music, all of it in languages that are endangered and spoken by few people alive today (TLA)
- Recordings reflecting linguistic and cultural identity from the audio archives of the Phonothèque Maison méditerranéenne des sciences de l'homme, such as recordings of southern French dialects (CNRS)

Spoken word performances

This category includes languages, dialects and examples of accents from around the world. Many of these language collections contain also ritual music, oral history, and recordings of special cultural events.

Some examples of languages & dialects recordings already selected for Europeana Sounds are:

- Accents, dialects and languages from the UK, Ireland and most parts of the English-speaking world and of non-native speakers speaking English (BL)
- Samples of world languages, including the Documentation of Endangered Languages corpora, with many retellings of old legends, descriptions of rituals or traditional music, all of it in languages that are endangered and spoken by few people alive today (TLA)
- Recordings reflecting linguistic and cultural identity from the audio archives of the Phonothèque Maison méditerranéenne des sciences de l'homme, such as recordings of southern French dialects (CNRS)

2.6.3   Radio recordings

Many radio recordings comprise material included in other categories, such as music, interviews, etc. In some cases, radio broadcast materials do not fit into these categories as they have been produced as radio recordings in their own right, including radio news, documentaries as well as performances of radio art projects specifically composed for radio transmission.

Some examples of radio recordings already selected for Europeana Sounds are:

- Yannick Dauby's *Arches* - an atmospheric sound piece built around his Winter 2010 recording of the wolves of Lozère, France (BL)[19]
- Recording of retired miner and union official Joe Kenyon, born 1915, recorded in Carlton, Barnsley, South Yorkshire in 1998 for the *Millennium Memory Bank* (BL)[20]

2.6.4 Environment recordings

Natural Sounds

Sounds of the natural world include a wide variety of birdsongs and animal sounds, providing evidence of European scientific exploration and collecting around the world.

The British Library's unrivalled natural sounds collection includes recordings of great scientific, cultural and historic interest. Made for the most part in nature reserves and wild locations, the recordings often have special significance in documenting ancient habitats that are gradually vanishing. Some of the species featured have already become extinct - the red-backed shrike from Norfolk, for example, and the Worcestershire marsh warbler.

Some examples of natural sounds already selected for Europeana Sounds are:

- Lions roaring in a zoo in Arnhem (NISV)[21]
- Cows grazing in a field (NISV)[22]
- Sounds made by dolphins at the Dolfinarium in Harderwijk (NISV)[23]
- Recording of a spawning male haddock courting a female (BL)[24]

Soundscapes

This category includes location ambience such as contemporary street sounds in town and cities, natural ambience sounds such as rainstorms, weather and wild places, sounds of war, sounds from industrial environments, machines such as engines, clocks, sound effects, etc.

Sounds have a powerful way of bringing past events back to life and audio content in this category may reflect everyday life or may reflect landmarks in history.

Some examples of soundscapes already selected for Europeana Sounds are:

- Victorian street sounds and steam railway sounds from London and around the British Isles (BL)
- General ambience in a polder with a barking dog (NISV)[25]

---

[19] http://sounds.bl.uk/Arts-literature-and-performance/Touch-Radio/028M-C1428X0064XX-0100V0 [REF 14]
[20] http://sounds.bl.uk/Accents-and-dialects/Millenium-memory-bank/021M-C0900X14543X-4200V1 [REF 15]
[21] https://soundcloud.com/beeldengeluid/het-grommen-en-brullen-van-een [REF 16]
[22] https://soundcloud.com/beeldengeluid/het-grazen-van-koeien-in-de [REF 17]
[23] https://soundcloud.com/beeldengeluid/dolfinarium-kreten-en-sonar [REF 18]
[24] http://sounds.bl.uk/Environment/Listen-to-Nature/022M-LISTNAT00110-0001V0 [REF 19]

- Recordings held at the Phonothèque Maison méditerranéenne des sciences de l'homme made by anthropologist Christian Bromberger of the ambient sounds of football stadiums, as part of comparative study on football in Marseilles, Naples and Turin (CNRS)

2.6.5 Audio-related materials

Audio-related materials that contextualise and enrich audio content will also be made available via Europeana Sounds. These include printed and manuscript music scores, texts such as interview transcripts, videos and images that are related to audio content. Many contributors will make available images of sound recording labels and covers, providing information as found on the original resource. Photographs of historical recording and playback equipment are included to provide information about the recording processes used and to display the method of playing recordings over time.

Some examples of audio-related materials already selected for Europeana Sounds are:

- Digital copies of precious musical manuscripts by composers from the 17th to 19th century. The manuscripts represent a wide range of musical styles and eras as well as musical life at the Austrian imperial court in the baroque period (ONB)
- Digitised manuscript scores including works by Bach, Mozart, Purcell and others (BL)
- Scores printed in the 16th-century including vocal part books containing motets, madrigals and chansons as well as instrumental music for renaissance instruments such as the lute (BL)
- Books, booklets, sheet music, video recordings of Irish traditional musicians, singers and dancers, drawings, photographs and paintings (ITMA)
- Transcripts of oral history interviews (BL)

2.7 Multilingual aspects of the metadata and content

Europeana Sounds will provide enriched access to content from 18 sound collections from 12 Member States. Although in many cases the dominant language of the content will be the national language of the data provider, many providers have content in languages other than their national language or presented in multiple languages. Examples of the European languages covered are Danish, Dutch, English, Flemish, French, Gaelic (Irish), Gaelic (Scots), German, Greek, Italian, Latvian, Lithuanian and Portuguese.

The Europeana portal provides a multilingual interface accommodating around 30 languages, with search results translatable using APIs. Multilingual labels resulting from vocabulary-based metadata enrichments (such as for genre and subjects) will be indexed for search and will be presented to end-users in a language-contextual manner.

---

[25] https://soundcloud.com/beeldengeluid/algemene-sfeer-in-een-polder [REF 20]

2.8 Geographical spread of the content origins

Europeana Sounds content will include material covering a wide geographical spread. Traditional music, natural sounds, oral history and language material included in Europeana Sounds will document the diversity of Europe's cultural heritage as well as providing a view of European exploration, scientific, business and scholarly endeavours around the world. This will include cultural heritage objects originating in the UK, Ireland, Netherlands, France, Germany, Latvia, Lithuania, Greece, Denmark, Italy, Austria and Portugal, as well as material originating in Asia, Africa, Oceania, the Americas and the Middle East.

The inclusion of geolocation metadata such as coordinates is important for searching by place and this metadata should be added as part of the enrichment process, either by semantic enrichment building on existing metadata for places, or by crowdsourcing. This should be investigated in Europeana Sounds WP2 Enrichment & participation.

3   METADATA REQUIREMENTS

This section provides brief guidance on the practical application of the Europeana Sounds content selection criteria in terms of metadata requirements, including guidance on locating information regarding the use of shared ontologies, appropriate rights statements and the Europeana Data Model (EDM).

A *Europeana Sounds Manual for Data Providers* containing full information about metadata requirements will be created before the first ingestion training course for data providers in October 2014 and will be updated during the project. The Manual will be made available publicly as part of our strategy to widen the community of data providers of audio items.

All content selected for Europeana Sounds will be described by appropriate metadata records, mapped to the Europeana Data Model[26], which is in the process of being updated under task T1.3 to accommodate requirements for audio content. The outcome of this work will be presented in Deliverable D1.4 EDM Profile for Sound, due for completion in September 2014.

All metadata records will be made available under the Creative Commons CC0 waiver[27].

Metadata records will contain one or more links to openly accessible digital objects hosted at the site of the data provider or hosted by a 3rd party on behalf of a data provider. Where possible, the links will be direct, permanent links to a digital object. Although the digital objects must be able to be accessed freely by anyone, not all digital objects will be available for re-use[28] due to licensing restrictions.

3.1   Metadata quality

Good descriptive metadata supports the identification and retrieval of relevant objects by users and supports navigation between objects by collocating related material. It informs users about what they can do with a digital object through clear rights statements. It supports interoperability, which is crucial in an environment such as Europeana, where metadata originates from many different sources. NISO's *A framework of guidance for building good digital collections* (3rd ed., 2007)[29] contains a set of fundamental metadata principles:

- Metadata Principle 1: Good metadata conforms to community standards in a way that is appropriate to the materials in the collection, users of the collection, and current and potential future uses of the collection.
- Metadata Principle 2: Good metadata supports interoperability.

---

[26] Europeana Data Model documentation http://pro.europeana.eu/edm-documentation [REF 21]
[27] Creative Commons CC0 waiver http://creativecommons.org/about/cc0 [REF 7]
[28] Definition of "Feely Re-usable Content", Europeana Pro Glossary http://pro.europeana.eu/glossary [REF 22]
[29] NISO A framework of guidance for building good digital collections
http://www.niso.org/publications/rp/framework3.pdf [REF 23]

- Metadata Principle 3: Good metadata uses authority control and content standards to describe objects and collocate related objects.
- Metadata Principle 4: Good metadata includes a clear statement of the conditions and terms of use for the digital object.
- Metadata Principle 5: Good metadata supports the long-term curation and preservation of objects in collections.
- Metadata Principle 6: Good metadata records are objects themselves and therefore should have the qualities of good objects, including authority, authenticity, archivability, persistence, and unique identification.

One of the aims of the Europeana *Metadata Quality Task Force*[30] established In December 2013 was to document a set of best practice guidelines for the creation of good metadata and to look at the restrictions facing institutions in the provision of good quality metadata. The forthcoming *Europeana Sounds Manual for Data Providers* will provide a link to these guidelines when they become available later in 2014.

3.2   Rights statements

Each metadata description will contain a rights statement relating to the use of the digital object(s). This rights statement will conform to one from the list of *Available Rights Statements* on Europeana Pro[31]. Further information is given in the Europeana Sounds *MS2 Rights labelling guidelines*[32] and in the forthcoming Europeana Sounds document *D1.2 Rights labelling guidelines & D3.1 Guidelines for contributing audio content into Europeana*, due at the end of August 2014. If proven necessary during the project, further explanation and relevant examples for audio and audio-related material will be provided in the *Europeana Sounds Manual for Data Providers*, which will be updated during the project.

3.3   Metadata enrichments

Data providers will carry out enrichments to descriptive metadata for Europeana Sounds, including assigning terms from shared vocabularies (see 3.4 below). Additional enrichment will be achieved through semantic enrichment and crowdsourcing, as outlined in Work Package 2 in the Description of Work[33]. The exchange policy for end-user contributions will be defined in Europeana Sounds *D2.1 Crowdsourcing infrastructure and exchange policy*, due for completion in October 2014.

---

[30] Europeana Meatdata Quality Task Force blog http://pro.europeana.eu/pro-blog/-/blogs/new-metadata-quality-task-force;jsessionid=850E6BB189DCC5527A8F1DF04DAB898F [REF 24]

[31] Europeana *Available Rights Statements* http://pro.europeana.eu/available-rights-statements[REF 25]

[32] Europeana Sounds *MS2 Rights labelling guidelines* http://pro.europeana.eu/documents/2011409/49dae80c-2631-4d43-8e28-571004e8bde7 [REF 8]

[33] EC-GA including Annexe I ("Description of Work") http://pro.europeana.eu/documents/2011409/8d0e9833-4608-494e-af77-681e68f8a8c8 [REF 2]

3.4 Shared vocabularies

Use cases gathered for Europeana Sounds task *T1.3 EDM Profile* show that Europeana data providers use a wide range of controlled vocabularies in their metadata. In the Europeana portal, using shared, multilingual vocabularies in descriptive metadata will enable users to search and navigate across the different metadata sets, enhancing discoverability.

Europeana Sounds task *T1.2 Ontologies* will identify mandatory and recommended multilingual vocabularies for metadata describing digital audio resources. This will be detailed in Deliverable *D1.3 Ontologies for Sound*, due for completion in August 2014 and will also be referenced in the forthcoming *Europeana Sounds Manual for Data Providers.*

NTUA will develop a SKOS[34] thesaurus as part of Work Package 5 to support the use of the Europeana Sounds music genre terms within a hierarchical thesaurus structure.

Data providers will add at least one Europeana Sounds broad genre thesaurus term (Music / Spoken word / Radio / Environment) to their metadata records during ingestion, which will allow Europeana Sounds material to be grouped into broad genre categories. Metadata records may also contain more specific genre terms, as specified in Deliverable *D1.3 Ontologies for Sound*.

3.5 Europeana Data Model

Data providers make a wide variety of cultural heritage objects available via Europeana and use a range of different metadata standards and languages to describe these objects. Mapping these various metadata standards to the Europeana Data Model (EDM) enables the data to appear in a meaningful way in this cross-cultural, multilingual context[35].

The main source of EDM documentation is the Europeana Pro site[36]. A presentation by Valentine Charles, *Introduction to the Europeana Data Model (EDM)* is also available[37].

The EDM offers a list of properties which can be selected for describing any cultural heritage object. Some of these properties are mandatory and must be provided to Europeana. Additional EDM properties for sound will be detailed in Deliverable *D1.4 EDM Profile for Sound*, due for completion in September 2014. Relevant examples for digital audio objects will be provided in D1.4 and in the forthcoming *Europeana Sounds Manual for Data Providers.*

---

[34] Wikipedia entry for SKOS http://en.wikipedia.org/wiki/Simple_Knowledge_Organization_System [REF 26]
[35] Europeana Data Model Factsheet http://pro.europeana.eu/documents/900548/f495317b-4557-4a60-9326-723f4618b44c [REF 27]
[36] Europeana Data Model documentation http://pro.europeana.eu/edm-documentation [REF 21]
[37] *Introduction to the Europeana Data Model (EDM)*
http://www.slideshare.net/ValentineCharles/edm-introduction-tpdlfinal [REF 28]

4 SUMMARY

This *Europeana Sounds Content Selection Policy* provides guidance to data providers selecting content for inclusion in the project, outlining the mandatory content selection criteria, defining categories of material and providing some examples of material already selected for inclusion in Europeana Sounds.

In addition to this document, a separate *Europeana Sounds Manual for Data Providers* containing full information about metadata requirements and the ingestion process will be created before the first ingestion training course for data providers in October 2014. This document will point to other relevant project deliverables such as *D1.2 Rights Labelling Guidelines, D1.3 Ontologies for Sound and D1.4 EDM Profile for Sound* and will be updated during the project.

5   REFERENCES

REF 1	*MS1 Initial Selection Policy* http://pro.europeana.eu/documents/2011409/23b8a083-07ee-46c3-af20-f8ea543e5869
REF 2	EC-GA including Annexe I ("Description of Work") http://pro.europeana.eu/documents/2011409/8d0e9833-4608-494e-af77-681e68f8a8c8
REF 3	*MS23 Revised aggregation design available* http://pro.europeana.eu/documents/2011409/1e94c275-22e9-4467-98cd-3398e41b69c8
REF 4	Europeana *2013 business plan* http://pro.europeana.eu/documents/858566/9d4632d3-3f6d-4162-ba29-27a9a739946d
REF 5	Europeana Version 2, D3.8: *Collections and Metadata Analysis, Strategy and Plan 2013* http://pro.europeana.eu/documents/866067/4b7a8158-277e-4ec3-8f7b-9913edb3077a
REF 6	*D2.1 Europeana Partner Strategy and Development Plan* http://goo.gl/gSrBpE
REF 7	Creative Commons CC0 waiver http://creativecommons.org/about/cc0
REF 8	Europeana Sounds *MS2 Rights labelling guidelines* http://pro.europeana.eu/documents/2011409/49dae80c-2631-4d43-8e28-571004e8bde7
REF 9	Blog post on WP3 Rights Workshop http://www.europeanasounds.eu/news/discussing-our-barriers-to-online-access
REF 10	Europeana Creative *D3.1 Specifications for Implementing the Content Layer of the Extended Europeana Licensing Framework* http://pro.europeana.eu/documents/1538974/62080954-342a-4731-b592-69abc199b5fb
REF 11	Grove Music Online *India, §I: The region: cultural context and musical categories*

REF 12	*Der unsichtbare Pianist* (DNB)
	http://d-nb.info/1000273466
REF 13	Recording of Clara Haskil with the Winterthur Symphony Orchestra playing the Rondo from Beethoven's Piano Concerto no. 3 in C minor, op. 37 (NISV)
	http://goo.gl/eMF28I
REF 14	Yannick Dauby's *Arches* - an atmospheric sound piece built around his Winter 2010 recording of the wolves of Lozère, France (BL)
	http://sounds.bl.uk/Arts-literature-and-performance/Touch-Radio/028M-C1428X0064XX-0100V0
REF 15	Recording of retired miner and union official Joe Kenyon, born 1915, recorded in Carlton, Barnsley, South Yorkshire in 1998 for the *Millennium Memory Bank* (BL)
	http://sounds.bl.uk/Accents-and-dialects/Millenium-memory-bank/021M-C0900X14543X-4200V1
REF 16	Lions roaring in a zoo in Arnhem (NISV)
	https://soundcloud.com/beeldengeluid/het-grommen-en-brullen-van-een
REF 17	Cows grazing in a field (NISV)
	https://soundcloud.com/beeldengeluid/het-grazen-van-koeien-in-de
REF 18	Sounds made by dolphins at the Dolfinarium in Harderwijk (NISV)
	https://soundcloud.com/beeldengeluid/dolfinarium-kreten-en-sonar
REF 19	Recording of a spawning male haddock courting a female (BL)
	http://sounds.bl.uk/Environment/Listen-to-Nature/022M-LISTNAT00110-0001V0
REF 20	Recording of general ambience in a polder with a barking dog (NISV)
	https://soundcloud.com/beeldengeluid/algemene-sfeer-in-een-polder
REF 21	Europeana Data Model documentation
	http://pro.europeana.eu/edm-documentation

REF 22	Definition of "Feely Re-usable Content", Europeana Pro Glossary
	http://pro.europeana.eu/glossary
REF 23	NISO A framework of guidance for building good digital collections
	http://www.niso.org/publications/rp/framework3.pdf
REF 24	Europeana Metadata Quality Task Force blog
	http://pro.europeana.eu/pro-blog/-/blogs/new-metadata-quality-task-force;jsessionid=850E6BB189DCC5527A8F1DF04DAB898F
REF 25	Europeana *Available Rights Statements*
	http://pro.europeana.eu/available-rights-statements
REF 26	Wikipedia entry for SKOS
	http://en.wikipedia.org/wiki/Simple_Knowledge_Organization_System
REF 27	Europeana Data Model Factsheet
	http://pro.europeana.eu/documents/900548/f495317b-4557-4a60-9326-723f4618b44c
REF 28	*Introduction to the Europeana Data Model (EDM)*
	http://www.slideshare.net/ValentineCharles/edm-introduction-tpdlfinal

6   APPENDIX

6.1   *Appendix: Europeana Sounds Rights and Metadata Ingestion Survey*

Rights & Metadata Survey

Europeana Sounds – Rights & Metadata Ingestion Survey

In this survey we ask you to give detailed, up-to-date information about the rights and metadata for the separate collections that make up the content in Table 0 (Underlying content) and Table c ('Locked' content) in the Description of Work. This is part of the work of task 1.4 of WP1 and task 3.1 of WP3.

Rights

Having correct rights labels in the Europeana Portal provides end users with clarity on what they can and cannot do with the available digital objects. The rights questions in this survey gather information on the rights status of your collections, your rights clearing practices and your feedback on the Europeana Licensing Framework. This information is paramount for the further work of WP3. After we have collected all survey responses we will analyse the information and produce extended guidelines for the integration of audio content within Europeana. These guidelines will be validated within the consortium during a workshop to be held on 12 June 2014.

Metadata

The metadata questions in this survey gather information about the metadata format and about the use of identifiers. This will assist with developing the ingest schedule for Europeana Sounds, in which the metadata will be divided into three metadata ingest sets to enable reports on progress based on the Performance monitoring table in the DoW. This survey asks you to indicate which metadata ingest set you predict might be appropriate for your collections. Each data provider will need to make some metadata available for the 1st metadata ingest set, which will be used during the 1st training session in October 2014.

The deadline for returning the survey is two weeks from now: Wednesday 16 April 2014

If you have any questions about the survey, please contact:

Rights: Lisette Kalshoven lk@kl.nl
Metadata: Andra Patterson andra.patterson@bl.uk

Rights & Metadata Survey

(1) Please state your name:

(2) State your institution (content provider):

Rights & Metadata Survey

Europeana Licensing Framework

(3) Currently the Europeana Licensing framework offers 13 different rights statements that can be used to communicate the rights status of digital objects that are made available via Europeana.

Do these 13 rights statements allow you to properly label all the works that you are planning to make available via Europeana as part of the Europeana Sounds project?

YES / NO

(3.1) if answer to (3) is 'No' Please describe the rights status(es) of the works that you are planning to make available that that is not covered by the 13 available rights statements

(4) Are there any (other) problems that you have with the Europeana Licensing Framework?

Rights & Metadata Survey

(5) Do you have any other questions related to the framework?

Rights & Metadata Survey

Rights Clearance

(6) Does your institution have a rights clearance policy or procedure?

YES / NO

(6.1) if answer to (6) is 'Yes' please describe your rights clearance policy (please include for which types of uses you normally clear rights e.g. publication on website, use in exposition, etc.)

(6.2) if answer to (6) is 'No' are you planning to develop a rights clearance policy during the Europeana Sounds project?

YES / NO

(6.2.1) if answer to (6.2) is 'yes' please describe what you expect from such a policy:

Rights & Metadata Survey

(7) Do you have criteria for determining when a work/recording is in the public domain?

YES / NO

(7.1) if answer to (7) is 'Yes' please describe the criteria you are using:

(8) How do you record the rights status of works/recordings in your collection? (please include in your answer the type of collection management software you use, e.g. AdLib, DigiTool, etc.)

Rights & Metadata Survey

Use of Identifiers

(9) What types of identifiers do you use for works in your collection? Please specify the type of identifier (e.g. ISAN / ISRC / ISWC / DOI / ISNI) and the type of work they are used with:

(10) What types of identifiers should Europeana support? Please specify the type of identifier and the type of work they should apply to:

Rights & Metadata Survey

Metadata ingestion training sessions

(11) Each of the 18 content partners will send **one** person to 3 metadata ingestion training sessions in October 2014, July 2015 and April 2016. Please specify who this **one** person will be from your organisation:

(11.1) Name:

(11.2) Email address:

Rights & Metadata Survey

Status of your collections

(12) Please complete the following table providing detailed information on the separate collections that make up the content listed in Table 0 (pages 22-27, Part B, DoW). Please add more lines if necessary. The first 2 lines are examples.

Title and/or brief description of collection or group of digital objects	Number and type of digital objects sound, image, film/video, text, other	Current rights status of digital objects & percentage of this collection	Predicted final rights statement for digital objects & percentage of this collection	Direct links provided to digital objects? Yes/No/ Unsure	File format & technical quality (e.g. sample rate, resolution)	Predicted metadata ingest set Set 1 (Oct 2014-Jan 2015) Set 2 (Feb 2015-Jan 2016) Set 3 (Feb 2016-Jul 2016)	Metadata format(s)	Comments
AR Gregory Kenyan bird recordings	4300 sound	rr-f 25% cc-by 75%	cc-by 100%	Yes	MP3; 41100kHz	Set 3	In-house/XML	
Digitised printed scores	15,000 image	cc-by 100%	cc-by 100%	Yes	JPG; 1135x865	Set 1	MARC/XML	

Rights & Metadata Survey

(13) Please complete the following table providing detailed information on the separate collections that make up the content listed in Table c (page 18, Part B, DoW). These are the collections you might bring into the project, but there could be a (rights related) issues with it. Please add more lines if necessary. The first line is an example.

Title and/or brief description of collection or group of digital objects	Number and type of digital objects sound, image, film/video, text, other	Current rights status of digital objects & % of this collection	Current constraints (e.g. password protected, geo-locked, etc.)	Direct links provided to objects? Yes/No/ Unsure	File format & technical quality (e.g. sample rate, resolution)	Predicted metadata ingest set *Set 1 (Oct 2014-Jan 2015) Set 2 (Feb 2015-Jan 2016) Set 3 (Feb 2016-Jul 2016)*	Metadata format(s)	Comments
Digitised classical music 78rpm recordings	2700 sound	rr-f 100%	Geo-locked, EU access only	Yes	MP3; 41100kHz	Set 3	In-house/XML	

Rights & Metadata Survey

(14) Please comment here if you have anything to add to the survey in terms of information we might need:

(15) If you have any feedback on the survey, it would be very welcome:

The deadline for returning the survey is two weeks from now: Wednesday 16 April 2014

As a reminder, the rights part of this survey will be used to create a document by WP3 that will be validated during a content partner workshop in Copenhagen on June 12th. Please let us know who from your institution will be joining us via this google form: http://goo.gl/pKvku1.

If you have any questions about the survey, please contact:

Rights: Lisette Kalshoven lk@kl.nl
Metadata: Andra Patterson andra.patterson@bl.uk

Thank you very much