Deliverable

Project Acronym: Europeana Creative

Grant Agreement Number: 325120

Project Title: Europeana Creative

D1.3 – Inventory of Content for Experimentation

Revision: 1.0

Authors:

Christina Holm, SAT

Frank Thinnes, PLURIO.NET

Henning Scholz, EF

James Morley, EF

Jiři Frank, NMP

Kati Hyyppä, AALTO

Lisa Mammitzsch, MfN

Lizzy Komen, NISV

Mareike Hirschfeld, MfN

Steven Stegers, EUROCLIO

Susanne Tremml, ONB

Tom Miles, BL

Ulrike Sturm, MfN

D1.3 – Inventory of Content for Experimentation

This document gives an overview of content that is available via Europeana for creative re-use and experimentation. It brings together the work that was done for the three content inventory milestone documents that elaborate on the available content to support the five themes in Europeana Creative. This deliverable also summarises the lessons learned during the content sourcing process and makes a proposal for further developing the Europeana content base.

Revisions

Version	Status	Author	Date	Changes
0.1	Document structure	Henning Scholz, EF	October 14, 2014	Document structure
0.2	Draft	Henning Scholz, EF; Lizzy Komen, NISV; Tom Miles, BL; Mareike Hirschfeld, MfN; Ulrike Sturm, MfN; Lisa Mammitzsch, MfN; Susanne Tremml, ONB; Jiri Frank, NMP, Frank Thinnes, Plurio.net; Kati Hyyppä, Aalto; Christina Holm, SAT; James Morley, EF; Steven Stegers, EUROCLIO	November 17, 2014	Draft of all chapters
0.3	Draft	Lizzy Komen, NISV; Breandan Knowltan, HP	November 21, 2014	Reviewers comments
0.4	Final Draft	Henning Scholz, EF	November 25, 2014	Comments integrated
0.5	Final Draft	Elisabeth Stricker, ONB	November 27, 2014	Layout edits, minor changes
0.6	Final Draft	Susanne Tremml, ONB	November 28, 2014	Minor changes
1.0	Final	Max Kaiser, ONB	November 28, 2014	Minor changes

Distribution

Version	Date of sending	Name	Role in project
0.1	October 14, 2014	Henning Scholz, EF	T1.4 Lead
0.2	November 17, 2014	Lizzy Komen, NISV Breandan Knowltan, HP	Reviewers
0.4	November 25, 2014	Max Kaiser, ONB	Project Coordinator
1.0	November 28, 2014	Marcel Watelet, EC	Project Officer

Approval

Version	Date of approval	Name	Role in project
1.0	November 28, 2014	Max Kaiser, ONB	Project Coordinator

Table of Contents

Tables

Figures

Executive Summary

A key task in Europeana Creative is to source and prepare content for the five Pilot themes of the project. Europeana Foundation lead this task and supported the Pilot Leads to identify suitable content for their prototypes and applications as well as making as much open and re-usable content as possible easily accessible to other creative industry partners involved in the Europeana Creative Challenges.

This deliverable is giving an overview of the content that is available in Europeana for re-use and specifies the content that is available for the five Pilot themes: History Education, Natural History Education, Tourism, Social Networks and Design. As this was the first larger scale project on content re-use involving both cultural heritage institutions (content holders) and creative industry partners, the lessons learned during the content sourcing are summarised in this document too, to help future projects with similar goals. Based on the lessons learned, the content sourcing process as such was reviewed continuously during the project and the up to date approach is described below. Having data providers in the project consortium helped a lot to mobilise content for re-use and make it available in Europeana. This is summarised at the end of the deliverable.

Sourcing, preparing and mobilising content for creative re-use was not only a task designed for the project. It will become a long term task for Europeana and the lessons learned from Europeana Creative will be incorporated for the future content sourcing work, to improve the Europeana repository and to increase the amount of re-usable content in Europeana.

1. Background

Europeana Creative will support and promote the creative re-use of cultural resources available via Europeana. Europeana generally does not hold content, but metadata that refer back to content holders. This presents an obvious barrier to creative re-use, as the development of rich applications clearly requires access to original digital files. A series of milestone documents has developed a process to remove this barrier: MS1 (Content Inventory to Support History Education and Natural History Education Themes) in M9 of Europeana Creative, MS2 (Content Inventory to Support Tourism and Social Networks Themes) in M17 and MS3 (Content Inventory to Support Design Theme) in M22. This deliverable will benefit from the lessons learned during the content sourcing for the five themes and specify the agreed longer-term mechanisms to facilitate the access to content via Europeana for creative re-use.

The outcome of this Europeana Creative task (T1.4 in WP1 of Europeana Creative) is an inventory of content available for creative re-use building on top of the Content Re-use Framework (see D3.1 of Europeana Creative[1]). The amount of content provided for the Pilots does not need to be very large in order for the Pilots to be successful. However, a larger pool of content is highly desired to build on top of the Pilot concepts. This is particularly important for the Challenge events: Content restrictions should not reduce the creative potential of the web and app developers participating in Challenges related to the themes.

In this context, content is used to refer to a Digital Object that is a digital representation of an object that is part of Europe's cultural and/or scientific heritage. Metadata refers to textual information and hyperlinks that serve to identify, discover, interpret and/or manage content.

For more definitions of terms used within this document, please see the Europeana Glossary[2].

---

[1] http://pro.europeana.eu/documents/1538974/62080954-342a-4731-b592-69abc199b5fb; accessed November 24, 2014.

[2] http://pro.europeana.eu/glossary; accessed November 24, 2014.

2. Content Inventory

2.1 Overview of Re-usable Content in Europeana

As of October 2014, the Europeana database contains 10.9 million records referring to freely re-usable objects (Fig. 1). This number is the total amount of records that are accessible via europeana.eu when applying the 'Can I use it?' filter using the value 'Yes, with attribution', which includes all items labelled PDM, CC0, CC BY, CC BY-SA. Extending the filter to include 'Yes, with restrictions' adds another 4.9 million records to the amount of records that are compliant with the Content Re-use Framework based on the rights statement.

Being in-line with the Content Re-use Framework, the content has to comply with several other requirements in order to be fully re-usable. These requirements include the availability of direct links, a high technical quality of the digital objects and no watermarks. A direct link in Europeana directly points to the digital object, not just a website where the object is available and embedded in context or in a viewer (e.g. a bookviewer). For images, the best indicator of the availability of a direct link is the lightbox that opens once the preview on the object page is clicked. Direct links can be provided to Europeana in edm:isShownBy (which would display the lightbox) and edm:object (which is the URL to an object used to generate the preview of the object in Europeana).

Currently it is still very difficult to give precise figures of objects that have a direct link, are of high quality and have no watermarks. This still requires a lot of manual work. In order to make it easier for creative industries to find high quality content for re-use, Europeana Foundation (EF) is making suitable datasets available via Europeana Labs, a playground for remixing and using the cultural and scientific heritage: http://labs.europeana.eu/data/. As of today (November 24, 2014) almost 60 handpicked datasets with openly licensed, directly linked and high quality (e.g. 800px for images) content are available (and 16 more datasets are in final preparation to be added soon). A media file checker[3] is under development that will make it a lot easier to surface re-usable content in Europeana.

---

[3] http://pro.europeana.eu/documents/1538974/62080954-342a-4731-b592-69abc199b5fb; accessed November 24, 2014.
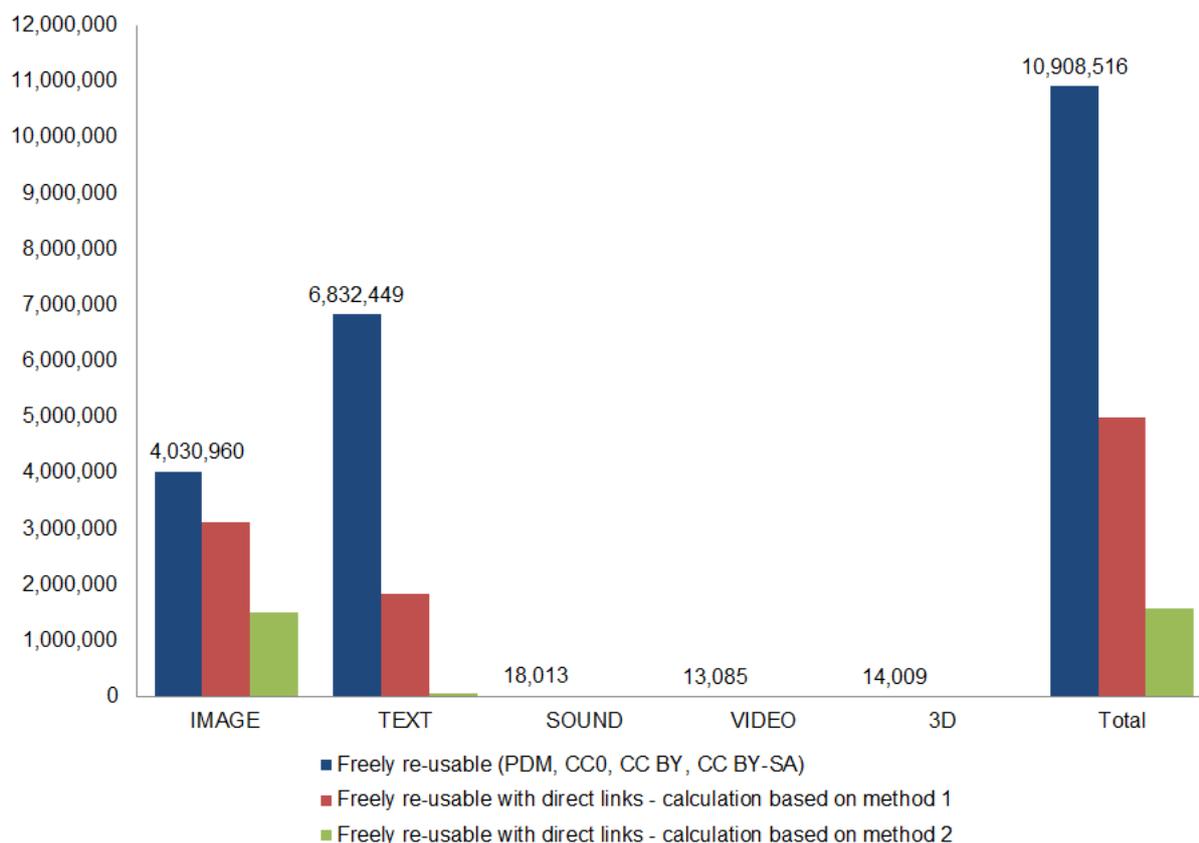
Fig. 1: Content available via Europeana that is freely re-usable and has direct links. The calculation methods are explained in the text

The media file checker will also give better statistics for re-usable content in Europeana. For now EF works with different methods in order to at least give a rough estimate of what is potentially available. A conservative measure of directly linked and re-usable objects is based on the total amount of records for which a value was provided as edm:isShownBy and for which the value provided as edm:isShownBy ends with a specified string representing the potential extension of the media file (calculation method 2, Fig. 1). The value provided with the edm:isShownBy field in the metadata should be a URL, pointing to a media file on the side of the data provider. In many cases, this URL has an extension indicating the type of the media file (jpg, pdf, mp3, etc.). For IMAGE type, the checked extensions are jpg/jpeg/png; for TEXT type, the checked extension is pdf; for SOUND type, the checked extensions are flac/mp3; for VIDEO type, the checked extension is mp4; for 3D type, the checked extension is pdf. Please note that not only links ending with these patterns are direct links, as direct links can also come without a file extension in the URL. Please also note that more extensions are suitable according to the Content Re-use Framework, but in order to keep the search queries manageable for the calculation, the number of extensions was limited. Taking this into account,

this method yields an underestimated number of direct links to media files. Not checking for extensions and only looking for records for which a value was provided as edm:isShownBy yields an overestimated number of direct links to media files (calculation method 1, Fig. 1). All links provided as edm:isShownBy are links to media files, but not necessarily direct links, e.g. linking to a viewer as it is often the case with books. Both methods also don't take into account media files linked directly via edm:object. The correct figures are considered to be lower than those obtained by method 1 and higher than those obtained by method 2. The current development of a media file checker will allow giving more precise figures in the future.

The methods above work for direct links, but not for technical quality and other aspects. As images are a very popular content type for re-use, a good sample of the images in Europeana was manually analysed to understand how many objects comply to all re-use requirements. EF analysed datasets of some big providers of images and a total of 6.2 million images out of a total of 19.2 million images that were accessible in Europeana in September 2014, when this analysis was done.

EF looked at the following criteria in order to analyse the re-usability of the content:

- Re-usable: This refers to the rights statement of the objects. All images labelled PDM or with one of the Creative Commons licenses (incl. CC0) were included.

- Potential direct links: A potential direct link is first of all a direct link to the digital object given to Europeana with the metadata of that object. It is also a direct link that is available on the website of the data provider, but not given to Europeana with the metadata. In this case it was necessary to work with the data provider to understand how to get this link also in the metadata in order to directly link to the object from Europeana.

- Actual direct link: The actual direct link is the one given to Europeana with the metadata of the object, either as edm:isShownBy or edm:object.

- Watermark: Watermarks on objects can be very diverse and can range from bars with the institution information added to the image to a logo or textstring across the entire image.

- Image resolution: It was possible for most of the 6.2 million images to extract resolution information (4.5 million out of 6.2 million images, see Fig. 3). In order to get a technical quality overview, the resolution was labelled in categories: <300px, 300-600px, 601-800px, >800px.

The team working on this task looked at all the criteria separately first to get an idea how many images have a rights statement that allow re-use, have direct links, etc. (Fig. 2). In a second step, they combined all criteria one by one to see how many images eventually remain that comply with all of the above criteria. Out of 6.2 million images in total, only 1.1 million images comply with all of the criteria (Fig. 4). Out of these 1.1 million images, more than 90% of the images are below 800px wide (Fig. 5) and thus not suitable for Europeana Labs, for example. Based on this analysis it is also very obvious, that the lack of an open rights statement is still the biggest barrier for increasing the amount of re-usable content (Figs. 2, 4).

These results are confirmed by another analysis that was done on images in Europeana using API calls to extract information in a similar way the media file checker would extract them (Fig. 6). This time, only freely re-usable images were included (PDM, CC0, CC BY, CC BY-SA). Adding all CRF requirements successively and then going from low to high resolution images continuously reduces the number of available items in Europeana.

These results can only give an idea how re-usable the Europeana content is and how many items the media file checker will surface when finally implemented. It is difficult to extrapolate these results to other parts of the Europeana repository. It is likely to assume that the remaining images not analysed for this purpose behave in a similar way. It is unlikely, however, that the patterns for text, sound, video or 3D are similar in any way. The raw data used for these analyses will be evaluated internally in more depth and these results will inform the plan to improve the Europeana repository in order to increase the amount of re-usable content over the next years.

As it needs to be a priority to increase the amount of openly licensed data in Europeana, it needs to be investigated what prevents cultural heritage institutions from opening up their collections for re-use. It is known, for example, that institutions are either worried or not even allowed (because of certain laws) to open collections for commercial re-use. Attribution is another problem that sometimes is enforced with restrictive licenses. Also the lack of expertise is in some cases responsible for choosing a restrictive license to avoid any copyright infringements. The case study 'Democratising the Rijksmuseum'[4] elaborates on related aspects in more depth, but it shows that a lot of work is needed in order to open up data for re-use.

Direct links are a requirement for an easy re-use of cultural heritage content via Europeana. While discussing this with data providers that offer direct links on their own website but not giving them to Europeana with the metadata, it was indicated that technical problems are a barrier. In some cases it is simply not possible for the technical infrastructure set up on the side of the data provider to expose these direct links for including them into the metadata.

The technical quality of the content is another requirement to make it attractive for creative industries. Technical problems are the most likely barrier to make high quality available. In some cases, objects were digitised some years ago in a lower quality. Making high quality available would mean to digitise everything again. In other cases, technical infrastructures would not allow to expose higher quality objects to the web. While high quality is made for preservation, the set-up of the systems is not considering the generation of high quality derivatives for the web.

---

[4] http://pro.europeana.eu/documents/858566/858665/Democratising+the+Rijksmuseum; accessed November 24, 2014.
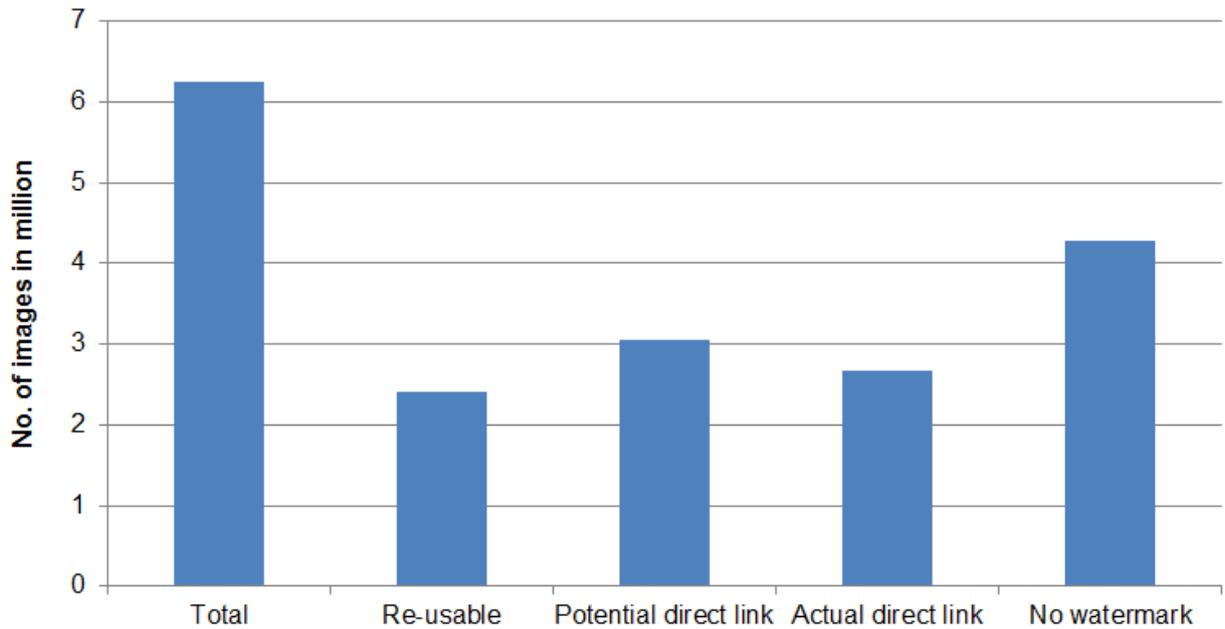
Fig. 2: Re-use criteria separately applied to 6.2 million images from the Europeana repository
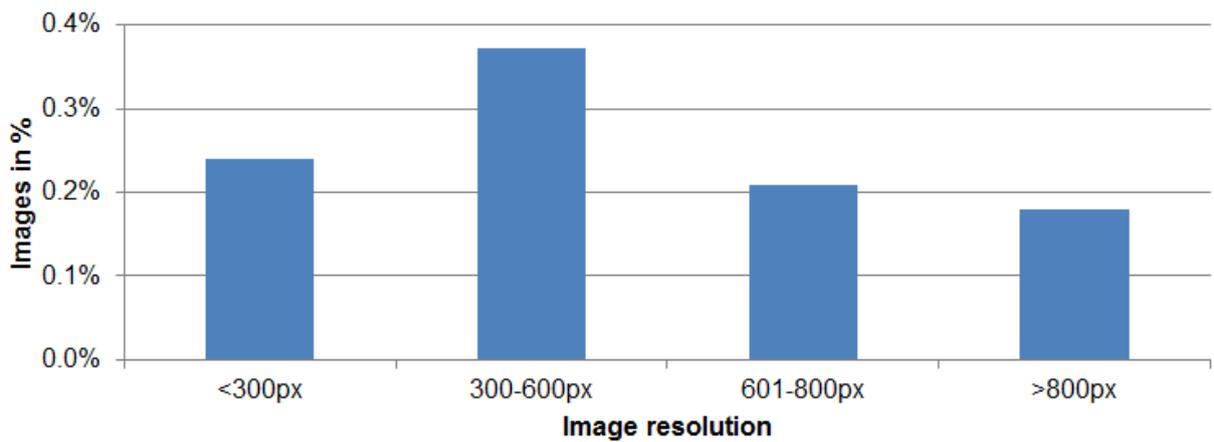
Fig. 3: Image resolution of 4.5 million images out of the 6.2 million images where resolution information was available
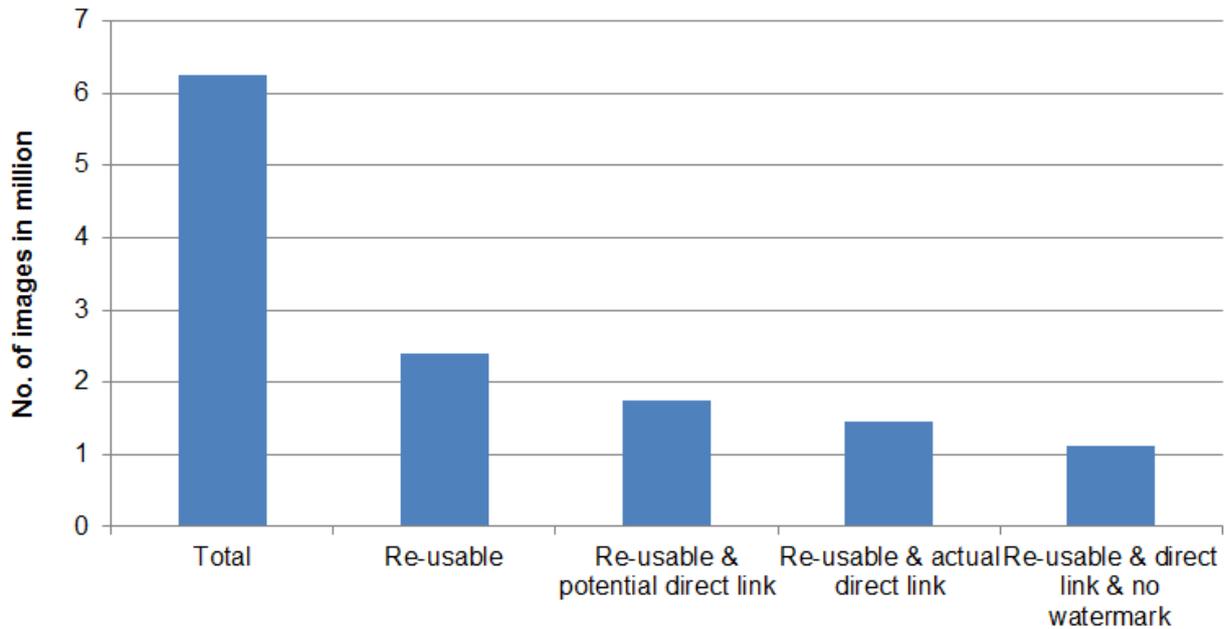
Fig. 4: Re-use criteria applied to 6.2 million images from the Europeana repository
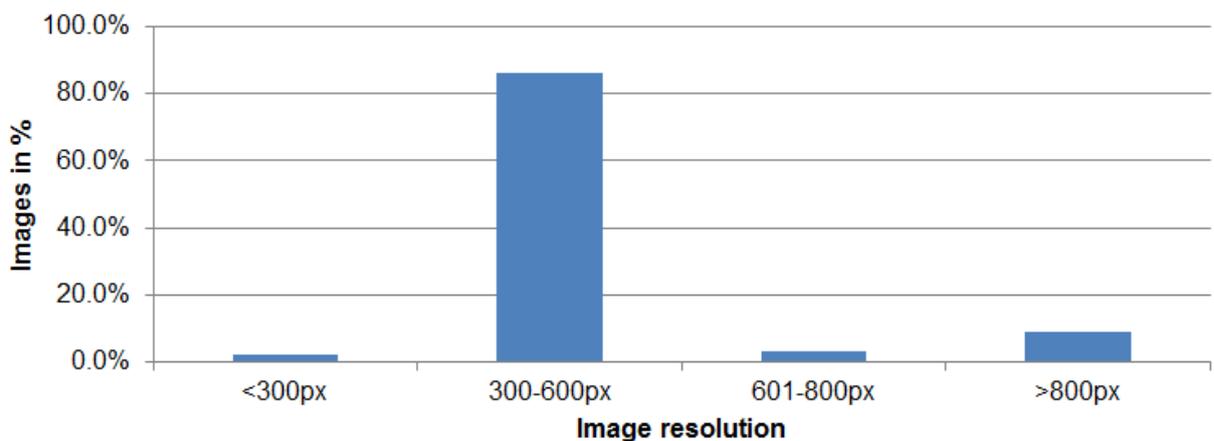
Fig. 5: Image resolution of all re-usable images with direct link and no watermark where resolution information was available
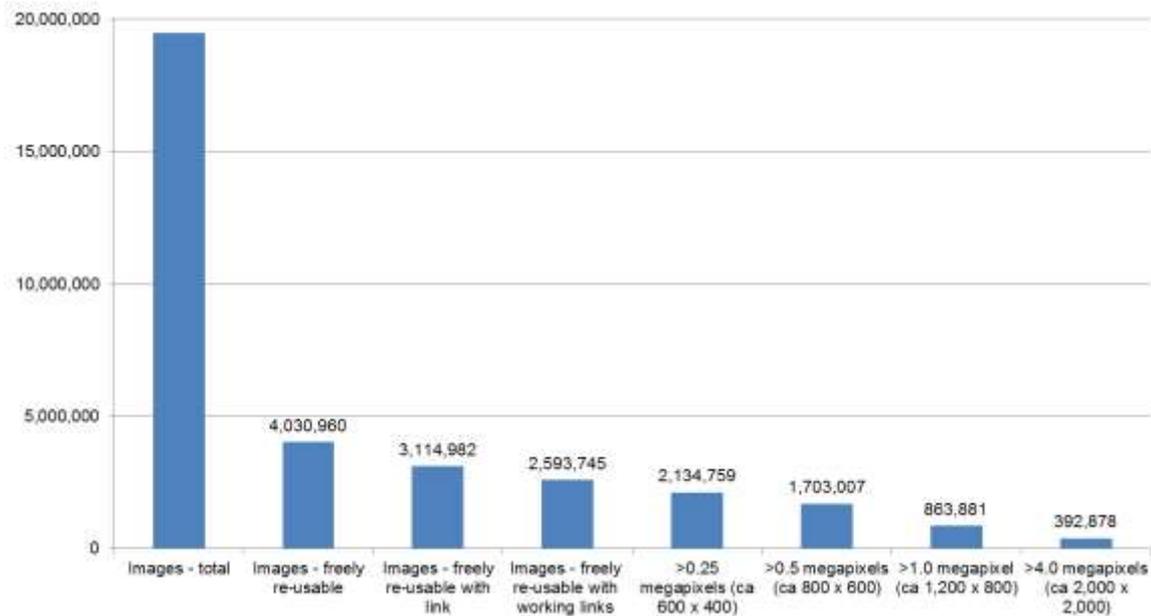
Fig. 6: Analysis of all images in Europeana based on API calls

2.2   Content Sourcing for the Five Themes of Europeana Creative

During the Europeana Creative project, five innovative Pilots are developed to showcase the potential to re-use Europe's digitised cultural heritage in innovative applications. The Pilot applications are games, learning apps or applications based on geo-referenced data. The five themes selected for these Pilot applications are History Education, Natural History Education, Tourism, Social Networks and Design. They are being developed through a process of co-creation involving education, tourism, social networks and design experts, software developers, business model specialists and other project partners.

The content sourcing for the five themes mainly followed the approach described further below in this document (section 4). The Europeana repository was the primary source for the content to be selected for the themes, starting from the full corpus of re-usable content that is described above. This was true for the Pilots and the Challenges. While My Europeana was used as the main tool to collaborate on content sourcing for the Pilots, Europeana Labs was used to make content available that is suitable for the Challenges (see below for a more detailed description of My Europeana and Europeana Labs). The requirements specified by the Pilot Leads were used to select the most suitable content out of the full corpus of re-usable content. While for most of the Pilots an itemised inventory was built in My Europeana, the corresponding datasets are offered on Europeana Labs for the Challenges.

2.3 Content Inventory for the History Education Theme

The Pilot was led by EUROCLIO, who developed Historiana as an online educational multimedia tool that offers students multi-perspective, cross-border and comparative historical sources (http://historiana.eu/). Within the History Education Pilot, a Historiana learning section was developed to stimulate the re-use of digitised heritage in history education. It offers history educators easy-to-find and free-to-use educational resources (sources, learning activities and tools) that are designed to stimulate historical thinking, multiperspectivity and active learning. In the Historiana Apps Environment, educators can create online learning activities using the tools, share these online learning activities with their fellow educators, and invite their students to do these learning activities. Educators can use pre-selected sources or upload their own sources.

The content for the History Education Pilot is focussed on World War 1. This decision was based on these three arguments:

- First World War is a major topic in most school history curricula in Europe and beyond.

- There have been three Europeana Projects that specifically focus on World War 1 (see below).

- The centenary of World War 1 has already and will continue to get a lot of press attention in most European Countries because of all the national and international commemoration ceremonies.

Sources for the History Education Pilot have been selected around nine content areas that highlight different aspects of the war. These are:

1. An interconnected world

2. The descent into war

3. The war begins

4. Key moments in the war

5. How the war was fought

6. Experiences and reporting the war

7. Who was to blame?

8. An uncertain peace

9. Remembering the war

As mentioned above, three projects were running in the last years dedicated to collect content related to World War 1. These projects are:

- [Europeana 1914-1918](#)[5]

- [Europeana Collections 1914-1918](#)[6]

- [European Film Gateway 1914](#)[7]

In particular some aggregators and data providers have very interesting World War 1 related content available via Europeana, which are:

- Imperial War Museum

- National Library of France

- Central Institute for the Union Catalogue of Italian Libraries (World War 1 collection in Europeana)

- [Austrian National Library](#)[8]

- [Great War Archive of Oxford](#)[9]

- [Berlin State Library](#)[10]

- EUscreenXL (http://www.euscreen.eu)

- Europeana Newspapers (http://www.europeana-newspapers.eu/)

Now that the Historiana database is improved and the Historiana Apps are functional, EUROCLIO will continue to add more of the content from the above collections, with a focus on including visual sources (example in Fig. 6 is in [Europeana](#)[11]).

---

[5] [http://www.europeana1914-1918.eu/en;](http://www.europeana1914-1918.eu/en) accessed November 28, 2014.

[6] [http://www.europeana-collections-1914-1918.eu/;](http://www.europeana-collections-1914-1918.eu/)accessed November 28, 2014.

[7] [http://www.europeanfilmgateway.eu/content/efg1914-project;](http://www.europeanfilmgateway.eu/content/efg1914-project) accessed November 28, 2014.

[8] [http://www.europeana.eu/portal/search.html?query=DATA_PROVIDER%3A%22Austrian+National +Library%22&rows=24&qf=PROVIDER%3A%22The+European+Library%22&qf=First+World+War](http://www.europeana.eu/portal/search.html?query=DATA_PROVIDER%3A%22Austrian+National+Library%22&rows=24&qf=PROVIDER%3A%22The+European+Library%22&qf=First+World+War); accessed November 28, 2014.

[9] [http://www.europeana.eu/portal/search.html?query=*%3A*&rows=24&qf=PROVIDER%3A%22The +Great+War+Archive%2C+University+of+Oxford%22&qf=First+World+WAr](http://www.europeana.eu/portal/search.html?query=*%3A*&rows=24&qf=PROVIDER%3A%22The+Great+War+Archive%2C+University+of+Oxford%22&qf=First+World+WAr); accessed November 28, 2014.

[10] [http://europeana.eu/portal/search.html?query=DATA_PROVIDER%3A%22Staatsbibliothek+zu+B erlin+-+Preu%C3%9Fischer+Kulturbesitz%22&rows=24&qf=First+World+War](http://europeana.eu/portal/search.html?query=DATA_PROVIDER%3A%22Staatsbibliothek+zu+Berlin+-+Preu%C3%9Fischer+Kulturbesitz%22&rows=24&qf=First+World+War); accessed November 28, 2014.

Fig. 7: Example of a visual source already selected for Historiana

At the moment EUROCLIO is in contact with Europeana Newspapers and EUscreen to explore opportunities to get direct access to their collections. In the Europeana CEF project, EUROCLIO hopes to work on the development of a Search and Select function using the API's of Europeana and Europeana related projects to help educators to find sources that they can use more easily.

2.4 Content Inventory for the Natural History Education Theme

2.4.1 Final Content Available for Natural History Education

The main aim of the Natural History Education Pilot is to demonstrate effective collaboration between memory institutions and creative industry partners, using digitised natural history content (mainly from natural history collections) published in Europeana. There are two Natural

---

[11] http://www.europeana.eu/portal/record/9200365/BibliographicResource_1000055502980.html?start=1&query=europeana_collectionName%3A9200365*&startPage=1&qf=hark+dogs&qt=false&rows=24; accessed November 28, 2014.

History Education Pilot applications which use different models, approaches and target different audiences. Both products are game-based and aim to introduce users to natural history content in an attractive, interactive and educational ways, whether that's at home with the family, or in schools and museums. One of the Pilot applications is a memory game called "Memory Match - Natural History Edition". The second Pilot application is a serious adventure game called "The Secret Legacy".

Most of the content selected for the Pilots was provided by OpenUp! project partners to Europeana[12]. OpenUp! currently provides more than 2 million metadata records to Europeana and is not only the largest data provider of natural history related content but the third largest data provider to Europeana over all. The majority of the data have a rights statement that allow re-use (1.5 million) thus making a lot of content available for experimentation.

According to the first content inventory from 2013, content for the Natural History Education Pilot can be related to the following topics: plants, zoology, entomology, palaeontology and geology. Figures 7 to 12 show screenshots of collections from OpenUp! partners available in Europeana and related to the five main natural history topics.

Fig. 8: Example for an OpenUp! collection from the Europeana portal suitable for the Natural History Education, topic "plants"

---

[12] http://www.europeana.eu/portal/search.html?query=*:*&qf=PROVIDER:OpenUp%21; accessed November 28, 2014.

Search ∨ | DATA_PROVIDER:"ETI BioInformatics, Leiden, the Netherlands" | **Q Search** | Help

Results per page: 24 ∨ | Results 88729 - 88752 of 93,286 | << < | 3698 | of 3887 > >>

Rhinolophus rouxii

Cynopterus sphinx

Eonycteris spelaea

Sphaerias blanfordi
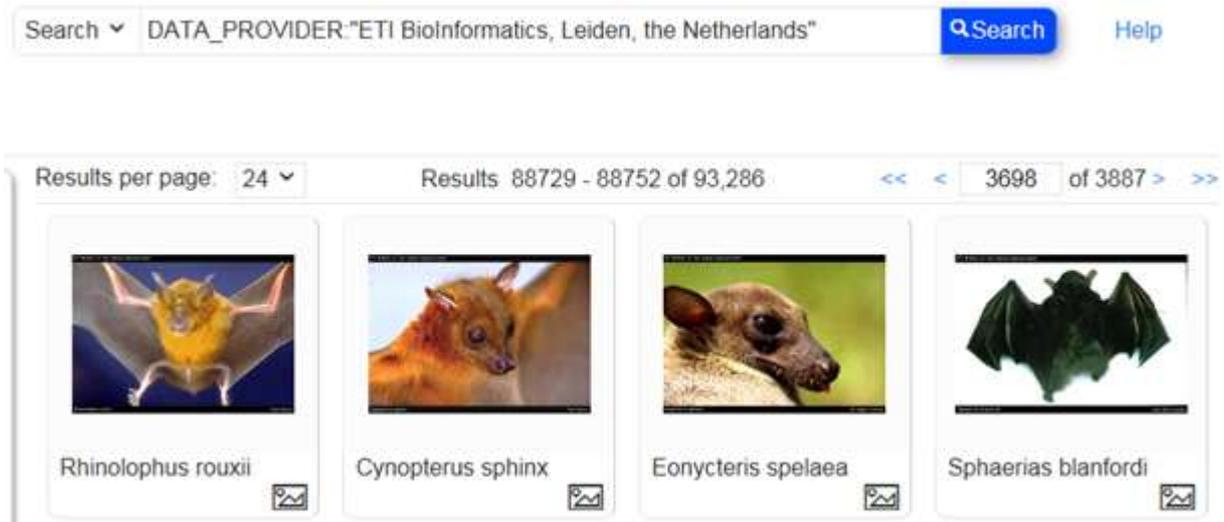
Fig. 9: Example for an OpenUp! collection from the Europeana portal suitable for the Natural History Education, topic "zoology"

Search ∨ | DATA_PROVIDER:"Natural History Museum of Denmark" | **Q Search** | Help

Results per page: 24 ∨ | Results 1177 - 1200 of 22,449 | << < | 50 | of 936 > >>

Eurysacca annulata Povolný 1986

Symmetrischema andinum Povoln...

Temnoaphelus hispidus Ferrer 1987

Saprosites carinatus Bordat 1992

Fig. 10: Example for an OpenUp! collection from the Europeana portal suitable for the Natural History Education, topic "entomology"
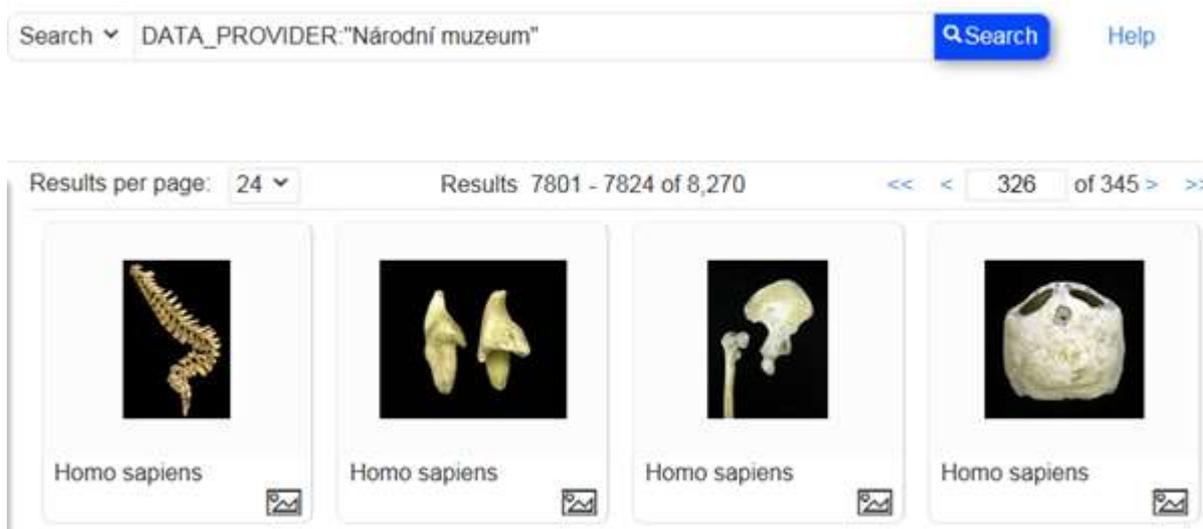
Fig. 11: Example for an OpenUp! collection from the Europeana portal suitable for the Natural History Education, topic "anthropology"
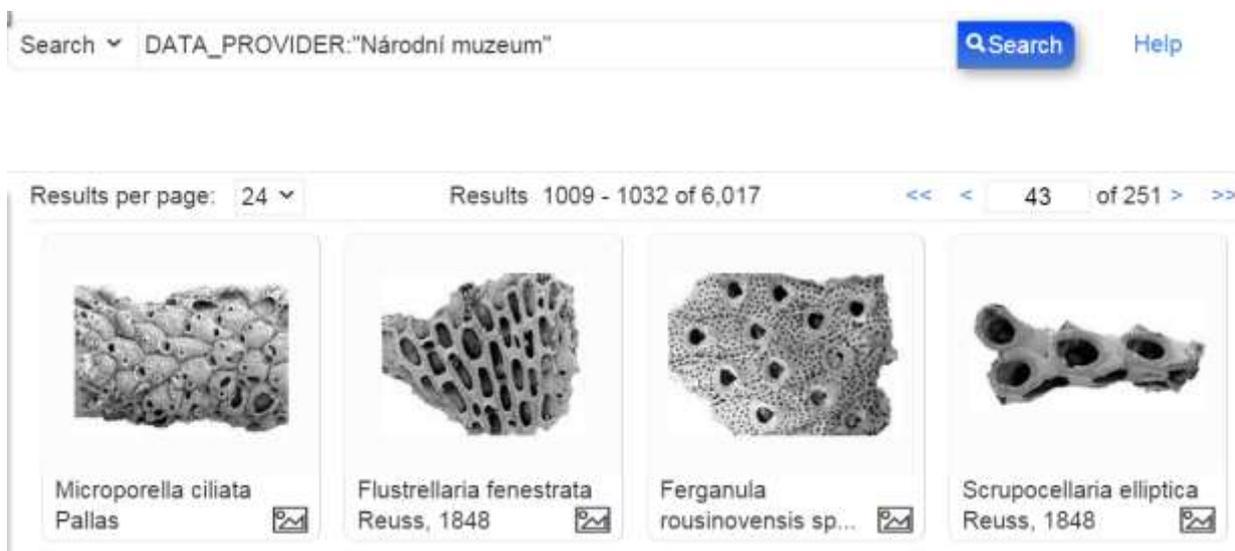
Fig. 12: Example for an OpenUp! collection from the Europeana portal suitable for the Natural History Education, topic "palaeontology"
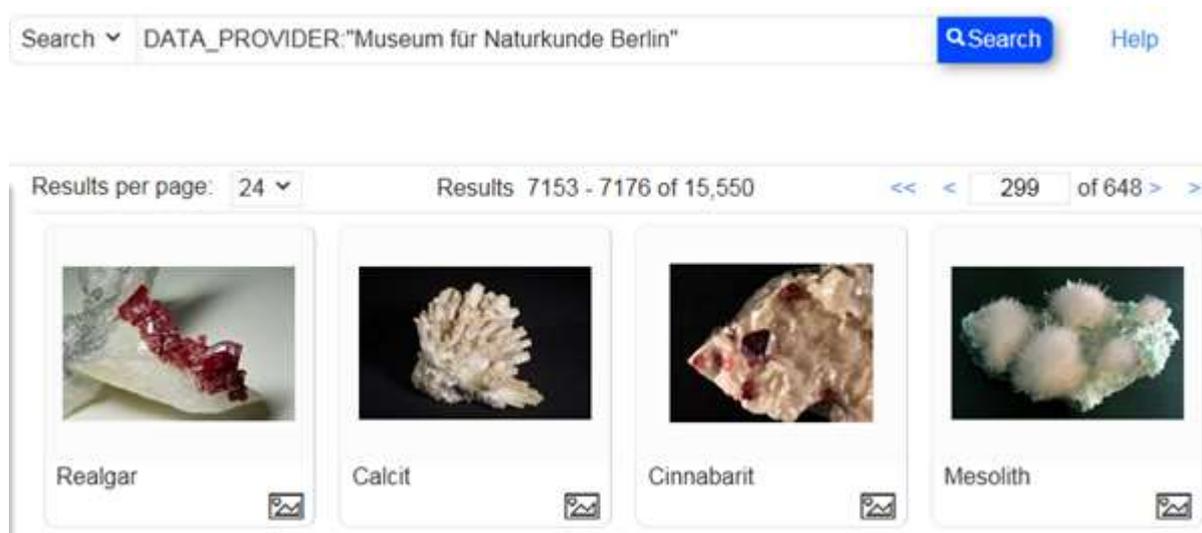
Fig. 13: Example for an OpenUp! collection from the Europeana portal suitable for the Natural History Education, topic "geology"

2.4.2 Final Content Used in the Memory Game

Memory Match – Natural History Edition is an educational game combining the classic memory game of finding pairs with the quiz game testing player's knowledge of natural history. The game is primarily aimed at school children and their teachers and parents. The Pilot app uses preselected content provided by the OpenUp! project to Europeana. The content was selected by the team at National Museum Prague. A dedicated My Europeana account was set up to store the queries and items used for the memory game card decks.

Three card decks with pre-selected content and related questions were prepared for the game already.

- Minerals: includes 42 minerals from NMP (17) and MfN (25) collections; sufficient to play the extreme version of the game with 42 pairs in the card deck.

- Insects: includes 32 insects from NMP (4), Oberösterreichisches Landesmuseum (3), University of Tartu - Natural History Museum (10), Royal Belgian Institute of Natural Sciences (9) and Zoological Research Museum Koenig (6) collections; sufficient to play the harder version of the game with 30 pairs in the card deck.

- Fossils: includes 32 fossils from NMP (29), Natural History Museum London (2) and University of Tartu - Museum of Geology (1) collections; sufficient to play the harder version of the game with 30 pairs in the card deck.

Content for a fourth set on animals is currently in preparation, as the license for the content needs to be negotiated with the data provider. If this negotiation is going to be successful, the card deck can be released for the game.

2.4.3 Final Content Used in the Adventure Game

The content used for the Natural History Education Pilot "The Secret Legacy" was highly diverse, covering the topics zoology, entomology, palaeontology and geology from eight different institutions. Also famous specimens of *Archaeopteryx lithographica* and *Brachiosaurus brancai* were included in the adventure game (Table 1).

Table 1: Final content from the Europeana portal used in the Natural History Education Pilot "The Secret Legacy"

MfN = Museum für Naturkunde Berlin, NMP = Národní muzeum Prague, TEL = European Library, NHM = Natural History Museum London, ONB = Austrian National Library.

Puzzle	Object	Provider	Rights Statement
Side Quest	*Luehdorfia japonica* (Leech, 1889)	MfN	CC-BY-SA
Side Quest	*Appias albina* (Boisduval, 1836)	MfN	CC-BY-SA
Side Quest	*Papilio dardanus* (Brown, 1776)	MfN	CC-BY-SA
Side Quest	*Zerynthia polyxena* (Denis & Schiffermüller, 1775)	MfN	CC-BY-SA
Side Quest	*Parnassius charltonius* (Gray, 1853)	MfN	CC-BY-SA
Side Quest	*Atrophaneura semperi* (Felder & Felder, 1861)	MfN	CC-BY-SA
Side Quest	*Graphium antheus* (Cramer, 1779)	MfN	CC-BY-SA
Side Quest	*Papilio nephelus* (Boisduval, 1836)	MfN	CC-BY-SA
Side Quest	*Belenois theuszi* (Dewitz, 1889)	MfN	CC-BY-SA
1	Map Humboldt	TEL	CC0

1	Vasa parrot	MfN	CC-BY-SA
2	*Hippocampus breviceps* (Peters, 1869)	MfN	CC-BY-SA
2	*Tinca tinca* (Linné, 1758)	MfN	CC-BY-SA
2	*Achatina murrea* (Reeve, 1849)	NHM	CC-BY
3	*Archaeopteryx lithographica* (H. von Meyer, 1861)	MfN	CC-BY-SA
4	*Rophalis relicta* (Hagen, 1856)	MfN	CC-BY-SA
4	*Mammuthus primigenius* (Blumenbach, 1799)	MfN	CC-BY-SA
4	*Schlueteria tetracheles* (Fritsch & Kafka, 1887)	NMP	CC-BY
4	Skull of *Homo sapiens*	NMP	CC-BY
4	*Mammites nodosoides* (Schlüter, 1871)	NMP	CC-BY
4	*Bubo lacteus* (Temminck, 1820)	MfN	CC-BY-SA
4	*Falco rusticolus* (Linné, 1758)	MfN	CC-BY-SA
4	*Brachiosaurus brancai* (Janensch, 1914)	MfN	CC-BY-SA
5	Calcite	MfN	CC-BY-SA
5	Hematite	MfN	CC-BY-SA
6	Picture Office Dr. Coleman	Rijksmuseum	PD
6	*Corophium volutator* (Pallas, 1766)	MfN	CC-BY-SA
6	*Leptocheirus pilosus* (Zaddach, 1884)	MfN	CC-BY-SA
6	*Aeginina longicornis* (Krøyer, 1843)	MfN	CC-BY-SA
6	*Podocerus brasiliensis* (Dana, 1853)	MfN	CC-BY-SA
7	Darwin: "On the Origin of Species"	Univ. of Edinburgh	CC-BY
7	Picture of Lamarck	ONB	PD

7	Picture of Linné	Upplandsmuseet	PD
8	*Araneus diadematus* (Clerck, 1757)	MfN	CC-BY-SA
8	*Lycosa accentuata* (Latreille, 1817)	MfN	CC-BY-SA

2.5 Content Inventory for the Tourism Theme

The Tourism Prototype is named 'VanGoYourself' (VGY) and is a genuinely different and surprisingly deep way for visitors to engage with cultural heritage, based on emotion, playfulness and curiosity. A web application in responsive design helps to recreate a painting with friends and family for virtually everyone with access to the Internet, then upload and share the photograph, twinned with the original, for others to enjoy (http://vangoyourself.com/).

The Pilot Lead (Plurio.net) has identified a set of criteria for content selection. In general, the service started with paintings with a great impact on the user. The criteria have been defined according to the overall demand driven approach of the Tourism Pilot and to provide content that is fit for touristic purposes. The criteria were shared with the Pilot team who applied them throughout the sourcing activities.

General criteria are:

- Great classics (impact = recognition of famous painter, famous painting, i.e. Last Supper from Leonardo Da Vinci);

- Famous scenes, people or "sujets" (impact = recognition of the scene, i.e. Last Supper, but from another painter than Leonardo; portraits of famous kings and queens, etc.);

- Lesser known images, painters or scenes / sujets with a surprising element that trigger an impulse to recreate it (because it's fun / weird / challenging).

Other significant aspects are:

- Different numbers of people, from portraits of 2-3 people to larger groups of people;

- People of different ages, sex, ethnicity;

- Great diversity in periods: antiquity, middle ages, renaissance, baroque, classical, romantic, symbolists, impressionists, expressionists, surrealists;

- Great diversity in provenance: from Italy to Scandinavia, from Spain to Greece, from Russia to the UK;

- Great diversity in subject: interior, exterior, day, night, dawn, allegories, mythology, Christian, pagan, European, extra-European.

Legal criteria, format, metadata:

- All the images of paintings must be licensed CC0 or CC-BY or be in the public domain.

- Short URL (to retrieve painting) or the image as file

- Name of painting (if known)

- Name of artist (if known)

- Period of painting or date (if available)

- Name of collection / museum where painting can be found

- Fairly good resolution, high resolution if possible

- No B/W images of colour paintings

In the beginning of the content sourcing process, some inventories were built in a dedicated My Europeana account and in Google docs (particularly on ONB content). The Pilot Lead also created a gallery on the Google Arts project[13] that gives an idea of the type and quality of content that is needed for touristic audiences and services.

Based on this input, the current inventory was build, which includes more than 60 paintings and is accessible at http://vangoyourself.com/category/paintings/. A lot of the content is in Europeana, while the content from some institutions is not yet accessible via Europeana. The content includes paintings from the following institutions, to give some examples of the diversity of the institutions that were considered: Amsterdam Museum, Herbert Gallery Coventry, National Gallery Denmark, Rijksmuseum, Royal Museums Belgium, Royal Pavilion Brighton, Saarlandmuseum, Simeonstift Trier, Smithsonian Institution, Van Gogh Museum, Villa Giulia, Villa Vauban Luxembourg.

Beyond the content that was selected for VGY to date, the Tourism Pilot team has identified new collections to be added to Europeana and existing collections in Europeana that need to open up in order to be re-used. Based on a priority list of the Pilot Lead, EF is working on improving the availability of content that is suitable for VGY.

In addition to the content for the Tourism Pilot, a number of datasets have been identified by EF being suitable for tourism in a more general sense. These datasets where added to Europeana Labs (http://labs.europeana.eu/data/), and tagged as either 'Tourism' or 'Geodata'. The data providers to Europeana also start to acknowledge them being featured on Labs, which hopefully helps to spread the message among data providers and then increase the number of open datasets suitable for Europeana Labs. Tourism was the theme that created the most interest among potential data providers to share their data more widely and also contribute to

---

[13] https://www.google.com/culturalinstitute/user-galleries?q=vangoyourself&hl=de&projectId=art-project; accessed November 28, 2014.

Europeana. It shows that the concept of the project actually works as it created a higher interest among cultural heritage institutions to open up parts of their collections for creative re-use. Some institutions in the list above are still lined up to be included in Europeana. This needs to be followed through more seriously to make this new data available in Europeana.

2.6   Content Inventory for the Social Networks Theme

The Social Networks Pilot creates simple and flexible tools and experiences that will allow learning as much as possible about the relationship between social networks and Europeana content. Communities of interest are invited to enrich specific sets of sounds in different ways.

There are two data providers from within the project that will deliver sounds for the Pilot, namely the Netherlands Institute for Sound and Vision (NISV) and the British Library (BL). Having actual data providers as part of the Pilot team helps a lot to source the necessary content as the Pilot team knows the requirement and knows the content they can make available to Europeana. In addition to these identified data providers, Europeana investigated if there were additional collections of sounds that would be suitable for the Pilot.

Europeana investigated all 486,547 sounds available at that point in time (January 2014) and checked if they were suitable for the Pilot. Out of this corpus, only 10% (45,878 items) of the objects were labelled with a Public Domain mark or a Creative Commons license. Only 5% of the total sounds data had both an open rights statement and a direct link (23,631 items). The data provided by the Pilot partners were however the most suitable for the Pilot.

The Netherlands Institute for Sound and Vision is using sounds gathered during the Geluid van Nederland (Sound of The Netherlands; http://www.geluidvannederland.nl/) project. These sounds come from both the institution's archive as well as user generated content (UGC). The content is all currently hosted on SoundCloud. NISV worked on mapping their sound data to the Europeana Data Model and delivered this data to Europeana via the Dutch National Aggregator (Digitale Collectie) in October 2013. The Sound of the Netherlands dataset can be accessed with the following search term on Europeana: europeana_collectionName:2021613*. The dataset is also featured on Europeana Labs[14].

All content available for the Pilot areaudio recordings, but the subjects vary widely from nature and city sounds to transportation and ambient sounds. NISV did an analysis of a portion of their sound collection to detect which are common subjects or themes. The sounds chosen for the Pilot are birds, airplanes and city sounds from Amsterdam and London. These were chosen due to the amount of high quality sounds with rich metadata available and the accessibility of relevant communities in the Netherlands and the UK.

---

[14] http://labs.europeana.eu/data/sounds-of-daily-life-and-birds-from-netherlands-institute-of-sound-and-vision/; accessed November 28, 2014.

Almost all the sounds (2,505 items) from NISV are available for re-use either under a CC BY, CC BY-SA, or CC BY-NC-ND license. Very few problems were encountered with the content. The problems that did occur were either the language since nearly all the descriptions and titles are in Dutch making it harder for non-Dutch speakers to read the metadata, lack of geotags, and improper dating. Improvement of geotags for the UGC records has been achieved with the help of Ontotext and although NISV does have a portion of their sound collection that does not include a date, the sounds that were selected for the prototype are sounds with existing dates.

For the Pilot, NISV provided audio content for all three subthemes. About 40-50 sounds for each subtheme were selected initially. The small amount of content was mainly because this was needed at the start of the Pilot to demonstrate that the sounds could be integrated via the Europeana API into the Historypin interface. NISV provided additional sounds for the aviation, birds and Amsterdam city soundscape themes during the process. For instance, the Amsterdam city sounds were increased to 360+ sounds.

The audio recordings supplied by the British Library came from a number of different sources. Around 1,000 recordings came from the UK Soundmap project, comprising crowd sourced recordings, which were uploaded onto Audioboo. The remainder came from various environmental and wildlife recordings, including the British Wildlife Recordings (640 recordings) and the A.R. Gregory recordings (around 700 recordings of birdsongs from Kenya). Where possible, the content is being made available under the CC BY license. For the UK Soundmap, Water and Weather recordings, the content is being made available under the CC BY-NC license. All the content needed to be mapped to Europeana Semantic Elements (ESE) format in order to be ingested by The European Library in November 2013. The European Library then exported the records to Europeana, in EDM format, in December 2013. The content for UK Soundmap is downloadable from the Audioboo platform. The remainder of the content can be either streamed (or, in future, downloaded) from the British Library "Sounds" website – http://sounds.bl.uk – but for the purpose of the Social Networks Pilot the content will be downloadable from SoundCloud platform. The British Library's content can be found on Europeana by searching for sounds and the British Library as a data provider. The British Library content will be used for two of the three themes: birdsongs and cityscapes (London).

Although the content from NISV and British Library was identified to be most suitable for the Pilot, some more content on the bird theme is available from other data providers in Europeana, one being also a partner in the Europeana Creative project (MfN). However, this additional content comes with restrictions, either with a restrictive rights statement, no direct link or a snippet of the original audio file. Therefore, these sounds are not considered for the prototype, but Europeana is using the prototype to negotiate data improvements with the data providers concerned.

2.7 Content Inventory for the Design Theme

The Design Pilot has three main strands of activities:

1. Prototypes that aim to stimulate and inspire professional designers, artists and practitioners – and creative industries in general – to creatively re-use Europeana content.

2. Digital Fabrication and hands on workshops, where practitioners and designers explore European digital cultural heritage and create together open cultural works.

3. Exhibition that is open for everyone to submit an artwork or design building upon the Europeana content.

The first set of requirements for content were specified in early September 2014 and focussed on two of the prototypes developed within the Design Theme:

- Culture Cam

- Teapot Machine

Since September 2014, content is sourced and prepared by Europeana, SAT and Aalto in very close collaboration. As the Pilot is still underway, this is not yet the final inventory for that theme.

2.7.1 Culture Cam

The basic idea is to create a tool that makes it easy and intuitive for designers and artists to browse the Europeana archive. The tool is a digital "live" similarity tool that helps recognise a colour, a shape or a pattern by using a web camera. You simply scan an object in front of your computer screen or tablet, the digital tool captures and analyses the given object and comes up with search results of similar colour, shape or pattern.

The content requirements for Culture Cam are listed below:

- Digital objects need to be labelled PDM in Europeana;

- Good resolution images (approx. 300x300px or higher);

- Direct links to the digital objects are not required;

- A selection of around 1,000 images;

- The images should be widely represented in colour, pattern and shape;

- Contemporary (design) content as part of the selection.

A My Europeana account was created for this prototype (access available on request). The Europeana content sourcing team selected suitable datasets and SAT handpicked suitable items from these sets to build the content base for the Culture Cam prototype. This was finished in early October 2014 with more than 1,000 images saved in My Europeana. The majority of the selected images are from the following data partners / institutions:

- Armémuseum (SOCH)

- Atria, Institute on Gender Equality and Women's History , Amsterdam

- Biblioteca Valenciana Digital

- Bibliothèque municipale de Lyon

- Bibliothèque nationale et universitaire de Strasbourg

- Bibliothèque nationale de France

- CIMEC - Institutul de Memorie Culturală, Bucharest, Romania

- Institut für Realienkunde (Kulturpool)

- Koninklijke Sphinx BV, Sociaal Historisch Centrum voor Limburg

- Music Library of Greece "Lilian Voudouri"

- National Library of the Netherlands

- Östasiatiska museet (SOCH)

- Rybinsk State Architectural, Historical and Art Museum Preserve

- Skoklosters slott

- The State Tretyakov Gallery

- Uměleckoprůmyslové museum, Praha (eSbirky)

- Universitat de Barcelona

- Ville de Bourg-en-Bresse

Going beyond the prototype phase requires much more content to be accessible via Culture Cam. Instead of using a pre-selected set of about 1,000 images, a much larger part of the Europeana repository should be available. This can be done be running the image similarity tool via the Europeana API to return previews of digital objects that are similar to the objects that were scanned via the webcam. There is no need to make a pre-selection, but the audience of the Culture Cam needs to be agreed on to define the types of content returned by the app. If creative industries are the main audience, it would make sense to limit the search results to re-usable items. If a broader audience is envisaged and Culture Cam is used as a different view on the Europeana repository for everyone to browse, it might not be necessary to limit the search results to re-usable items. This has to be further discussed with SAT and Aalto in order to define the content requirements for the Culture Cam app.

2.7.2   Teapot Machine

The Teapot Machine is a wall-mounted, mechatronic relief, which brings pictures of teapots and other old everyday objects from Europeana alive. The idea is to attach cut-outs of printed photos onto a three-dimensional structure made of board material, and to create an animated collage. The aim is to produce sequences of movements where the objects interact with each other in a playful way, giving the impression of a chain reaction. Sounds associated with the movements will also be produced in order to give a more lively expression. The plan is to make the machine interactive by using sensors that trigger the movement patterns as people approach the work.

The idea in this prototype would be to use mainly openly licensed (PDM, CC0, CC BY, CC BY-SA) and at least 1,000px wide pictures of everyday objects like teapots and cups, as well as patterns/textures from different Europeana data partners. The final set used in the prototype does not have to be large. A set of max. 20 objects and some pattern images to create the background would be enough.

One of the key issues in selecting the images for the prototype is their style. The shapes should be visually clear and interesting. All the images in the Teapot Machine should also work together aesthetically. They do not have to be similar, but somehow create a coherent atmosphere.

The list below gives some examples of datasets in Europeana that have suitable images:

- Koninklijke Sphinx BV[15]
- Rybinsk State Architectural, Historical and Art Museum Preserve[16]
- Rijksmuseum[17]
- Hallwylska museet[18]
- Amsterdam Museum[19]
- Museum Rotterdam[20]

---

[15] http://www.europeana.eu/portal/search.html?query=Koninklijke+Sphinx+BV&rows=96&start=1; accessed November 28, 2014.

[16] http://www.europeana.eu/portal/search.html?query=europeana_collectionName%3A2023831*&rows=24; accessed November 28, 2014.

[17] http://www.europeana.eu/portal/search.html?query=europeana_collectionName%3A90402*&rows=24; accessed November 28, 2014.

[18] http://europeana.eu/portal/search.html?query=*%3A*&rows=96&qf=DATA_PROVIDER%3A%22Hallwylska+museet%22&qf=porslin+OR+keramik; accessed November 28, 2014.

[19] http://www.europeana.eu/portal/search.html?query=DATA_PROVIDER%3A%22Amsterdam+Museum%22&rows=24&qf=PROVIDER%3A%22Digitale+Collectie%22&qf=theepot; accessed November 28, 2014.

Further looking into content that is suitable for the Pilot Lead to a refinement of the inventory. These 12 items from Sociaal Historisch Centrum voor Limburg / Koninklijke Sphinx BV, Het Nationaal Glasmuseum and Rijksdienst voor het Cultureel Erfgoed are most suitable for the prototype and negotiations are underway to make the necessary image quality available for the Teapot Machine:

- Cup and plate with blue dots, Decor: Chabl. Decor 23 blauw

- Cup and plate with yellow ornaments, Decor zonder naam

- Teapot and cup with yellow flowers, Decor: 200 a

- Cup and plate with red sawtooth ornament, Decor: Chabl. Decor Lotus

- Cup with red dots, Decor: Chabl. Decor 23 Corail

- Sketch of glasses, Onbekend

- House, Maagdenhuis

- Architectural elements, Grote of N.H. St.Nicolaaskerk

- Architectural element, Maagdenhuis

- House, Paleis Kon. Moeder

- Architectural elements, Diversen

- Church tower, N.H.Kerk

---

[20] http://www.europeana.eu/portal/search.html?query=DATA_PROVIDER%3A%22Museum+Rotterdam%22&rows=24&qf=PROVIDER%3A%22Digitale+Collectie%22&qf=theepot; accessed November 28, 2014.

3. Lessons Learned and Recommendations

Europeana Creative was the first large project where several creative industry partners were involved to experiment with content made accessible through Europeana. The experiences that data providers, creative industry partners and Europeana made during the content sourcing process are summarised below. Recommendations are developed as well, which will be followed up in the following chapter where the content sourcing process is described in more detail.

3.1.1 Content Requirements

Well defined requirements for the content to be re-used within the Pilot applications were essential for EF to support the content sourcing process from the start of the work on the Pilots. It is equally essential to involve the EF content sourcing as early as possible even in the design of the Pilot concept. Otherwise, Pilot Leads and EF staff enter the content sourcing process with different background information and expectations. Without a common understanding of the goals of the Pilot, it is very difficult to efficiently source for suitable content. A very close collaboration of Europeana and the creative industry partner during the content sourcing process also helps to speak the same language when specifying the content needs for the Pilot.

It became obvious that the earlier the content requirements are available the more likely it is to have the content available on time. This is particularly true if new content is required that still needs to be ingested into Europeana. This has to do with the set-up of the ingestion and publication process of data in Europeana (including how the aggregator model works). Adding new data to Europeana simply takes several weeks or even months. Also the change and update of collections in Europeana to meet the content requirement needs a considerable amount of time. This process needs to be transparent for creative industry partners in order to manage expectations and plan accordingly.

Content requirements need not only be available to EF to help the Pilot Leads to source suitable content. They also need to be available and communicated to the data providers that are most suitable for providing this content to Europeana and the aggregators responsible for processing these data for Europeana. This is particularly important when content is prepared for Europeana specifically for the purpose of being used in a certain project. It simply takes time (up to several months) to prepare the content, rethink strategies and policies for re-use or set up a technical infrastructure to submit the data according to the requirements. Examples from within the project are the Austrian National Library and the British Library. Their progress with providing data to Europeana following the requirements for re-use is described in a following chapter, where they also describe the challenges they were facing.

3.1.2   Search and Browse

The Europeana portal search interface does not work intuitive enough to make it easy for creative industry partners to find what they are looking for. Finding the right content is complicated by an insufficient metadata quality in many cases. At least it is not always clear how to search and in what metadata fields to search for what information to surface the required content. Therefore, the help of the EF content sourcing team was necessary, at least to get the Pilots started. EF is happy to also support creative industry partners in the future to source suitable content in Europeana.

Browsing was found to be a very important way to get an understanding of the composition of a dataset. The availability of thumbnails is the key for making browsing efficient. EF needs to make sure to have previews for all datasets in the repository.

The 'Can I use it?' filter was acknowledged by the Pilot Leads being the most valuable filter to narrow down for usable content. However, filtering for technical quality is a very necessary option to make the search for suitable content more efficient. For images, a quality filter should also consider aspects like JPG-compression artefacts. The media file checker that is currently under development will make filtering for technical quality criteria much easier.

Filtering for file type or genre was mentioned as another important amendment that would make search and retrieval of suitable content easier. Filtering for genre is difficult to realise now, while filtering for type is to some extent possible even now. If a direct link to a digital object is given to Europeana with the file extension in the URL, this information can be surfaced with a specific query. EF staff can help in designing such a query to return only content available in specific file types.

Seeing the difficulties of creative industry partners to find sufficient usable content when searching Europeana directly, offering pre-defined and well curated collections of usable content would help to make re-use easier. Europeana Labs is already and will continuously be used to promote and describe openly licensed and high quality datasets from Europeana.

3.1.3   Available Content

Most Pilot Leads concluded that the majority of content in Europeana is not fit for purpose and not available for re-use. Not everything in Europeana is available under open licenses and with direct links. Also the technical quality of the content is not high enough to support re-use. With regards to images, the resolution is too low, watermarks are quite common and sometimes even coloured paintings are transformed into B/W derivatives. The outcomes of the Re-use Task Force, which started in October 2014, will help to make more content in Europeana fit for purpose.

High metadata quality is not only important to find the content in Europeana. It is also important when working with the material in a re-use scenario. A good and meaningful title and the correct creator are necessary to cite the content properly. The information about the content holder is important to attribute correctly (see also next chapter). It is currently investigated how we can best advise data providers to improve the metadata quality at source.

Open licenses and correct rights information is very critical for creative industry partners to know what they can possibly do with the content. It is sometimes difficult enough to find openly licensed content in Europeana. However, this information is not always reliable. Contradictory rights information is not the exception, where the rights statement on the side of the data provider is different from what it says in Europeana. In some cases, unclear rights information are also provided with the metadata, where the information in dc:rights is not a clear indicator for the possible use. Creative industry partners tend to be very careful and do not use any content in case they have doubts that this is possible without infringing copyright. Improving the rights statement (incl. re-negotiate with data providers) and their consistency in Europeana will be a real added value for creative industries. Europeana also needs to give more clear guidance on how to clear rights, why to open up content for re-use and what the benefits of the public domain are (see below chapter on value for partners). In the meantime, creative industry partners are most welcome to contact Europeana if they have questions about rights statements being unclear or contradictory.

For more experiences of creative industry partners with regards to the availability of content for re-use, their lessons learned and recommendations, the blogpost of Melissa Terras[21] is recommended.

3.1.4   Manage and Collaborate

My Europeana was used a lot by the Pilot Leads to collaborate during the content sourcing process. My Europeana is your personal space on Europeana. With My Europeana[22], you can save your favourite items and searches, and add tags. However, it is not powerful enough to support the process most efficiently. Managing a large number of saved items is quite cumbersome. More sharing, filtering and sorting options would be acknowledged as well as the possibility to easily export some metadata (e.g. data provider information).

My Europeana was also used further on in the process, using the My Europeana API to for example make tagged items accessible to a Pilot application. Project partner Semantika was experimenting with objects tagged in My Europeana to create new card decks for the memory card game (Natural History Education theme). However, this was a cumbersome process as Semantika had authentication issues with My Europeana. Semantika had to build a dedicated web service to overcome this issue and make the API connection work.

Attribution is good practice for all data, even if it is labelled public domain. Pilot Leads experienced that it is not always easy to attribute properly. One reason for this problem is the large number of items from different data providers that creative industry partners are working with, making it difficult to always use the correct information. This is particularly true if creative

---

[21] http://melissaterras.blogspot.co.uk/2014/10/reuse-of-digitised-content-1-so-you.html; accessed November 28, 2014.

[22] http://europeana.eu/portal/usingeuropeana_myeuropeana.html; accessed November 28, 2014.

industry partners are not very technical to query Europeana (portal or API) properly. Also the way that the content holder is presented in the metadata is not always clear enough to attribute correctly.

Related to the attribution of the data provider is the suggestion to find ways to link back derivative works made during the creative re-use process to their originals in Europeana based on the attribution information. This would then close the loop also for the institutions and help them to understand for what the content was used for (see also next chapter).

3.1.5   Value for Partners

For many of our data partners it is not clear why re-use and opening up for re-use is beneficial for them, what it actually means to them and how creative industries and cultural heritage organisations can work together more efficiently. The Europeana Strategy 2015-2020[23] highlights a few aspects why it is important for data partners and how it creates value for them. However, more concrete examples and case studies are necessary to elaborate on the opportunities for data partners and help them achieving their goals. The Content Re-use Task Force kicked off recently to work with data partners and creative industries on a vision, business models and case studies. This Task Force will also give clear guidance on what re-use scenarios can be realised with what data quality.

---

[23] http://strategy2020.europeana.eu/; accessed November 28, 2014.

4. Process for Content Sourcing and Selection

Based on the lessons learned during the content sourcing for the five Europeana Creative themes, the process was refined in a way to be most efficient and universally applicable to related and comparable cases now and in the future.

Europeana is following a staggered approach in order to source suitable content for creative re-use:

1. Starting point for the content sourcing are requirements from the creative industry partners that want to use content for experimentation.

2. Content available in Europeana that meets the content requirements for the re-use scenario needs to be sourced.

3. Content available in Europeana which is in principle suitable for the re-use scenario but not available under conditions that allow re-use needs to be sourced.

4. Existing data providers need to be identified who can provide new content to Europeana to make it available for experimentation.

5. New potential partners (i.e., Europeana Network members but also any other institution holding relevant digitised cultural heritage objects) need to be identified in order to provide content for re-use to Europeana.

One reason for choosing this approach is the time that is required to publish data in Europeana. If suitable data are available in Europeana, they are available for re-use immediately. If data need to be updated in Europeana to be suitable for re-use, this takes at least one month (Europeana is working with a monthly publication cycle). If data from a new provider need to be made available in Europeana, it can take up to several months. A new data provider first has to find the best aggregation channel to contribute. The institution then has to comply with the formal requirements (e.g. set force by the Data Exchange Agreement) and prepares the data. After completion of these steps, the data will be submitted to Europeana via the appropriate channels. All data need to be validated, checked, processed and published before they are available for re-use. Seeing these constraints, it is much more efficient to not start with new data providers when sourcing content for re-use.

In the Description of Work of Europeana Creative, a storage space was envisaged to be available for new data providers to make content available for experimentation. However, this option was never used during the project. EF always followed the existing aggregation model relying on the content to be available on the side of the data provider or aggregator. This was also working because the majority of the new content that was made available for experimentation came in through existing and well established aggregation routes. In these cases in particular, an additional storage layer would have also made the entire process more complex.

The entire content sourcing process is described below in more detail after the tools involved in the content sourcing process are introduced.

4.1 Collaborative Workspace for Content Selection

My Europeana[24] was identified as the collaborative workspace for content selection. For every theme EF created a new account to tag and save items and save search queries. However, the thematic My Europeana accounts are not public and the credentials to them are not publicly available (but available on request). My Europeana is for the content sourcing team to collaborate more easily also with the creative industry partners.

4.2 Europeana Labs

Europeana Labs is the playground for remixing and using our cultural and scientific heritage, a place for inspiration, innovation and sharing. It is also our public interface to re-usable datasets in Europeana: http://labs.europeana.eu/data/. Here Europeana features datasets that fulfil the following criteria: freely re-usable (PDM, CC0, CC BY, CC BY-SA), direct links to digital objects (images must be at least 800px wide). Any subsets of a dataset are eligible where one of the Europeana portal facets can be applied to narrow down the dataset. If a dataset contains both freely re-usable and in copyright content, the 'Can I use it?' filter can be applied to identify the re-usable part of the dataset for Europeana Labs. Using keywords or more specific queries is not eligible to identify a subset for Europeana Labs. It is also not possible to build thematic collections out of several datasets that fit a specific keyword. The reason for these restrictions is the plan to automatise the updates of Europeana Labs as much as possible, which is only possible if the data in Europeana Labs represent as much as possible real units in the Europeana repository.

Several staff members of EF are currently involved in updating the Europeana Labs dataset pages on a monthly basis. The updates of Europeana Labs are synchronised with the monthly publication cycle of data to the Europeana repository. Once every month new datasets will be published and existing datasets will be updated as much as possible. This also results in more open datasets with high quality objects that can be promoted on Europeana Labs.

In order to facilitate the collection of new datasets in a standardised way, EF set up a dedicated Google sheet that is amended every month with new data (access available on request). The following fields need to be filled in order to publish a dataset on Europeana Labs: Title, ImageURL, Description, Data provider, Link to data provider home page, Contact information, Copyright, Copyright URL, Dataset, Link to portal, Link to console, Categories. With the following guide anybody interested in preparing a dataset for Europeana Labs should be able to do so:

---

[24] http://www.europeana.eu/portal/myeuropeana#login; accessed November 28, 2014.

- Title: Every dataset that is featured on Europeana Labs need to have a title. The title has to be short and appealing to promote the dataset.

- ImageURL: Every dataset is illustrated with one picture to further promote the dataset. This image is selected out of the complete dataset. It should be in landscape format to allow a consistent presentation of datasets in Europeana Labs. For Europeana Labs it needs to be scaled down to 800 pixels width. Every picture is renamed to have the dataset ID in the file name plus a one word description of the set (e.g. data provider name). The images are then uploaded to the appropriate Github folder. As this is a manual process, all images that need to uploaded need to be available to the EF content sourcing team in order to batch upload them and adapt the ImageURL to reflect the Github folder where the images are uploaded.

- Description: The description of the dataset should have some more details of the dataset content, like theme, object types and rough numbers of objects. The number of objects should be given as a 'more than' number to also be sure that this information is not outdated when a dataset is updated in Europeana. It is recommended to study descriptions available on Europeana Labs to get an understanding how much detail is necessary.

- Data provider: The data provider information corresponds to the information about the data provider in Europeana (edm:dataprovider).

- Link to data provider home page: This is a public website where more information about the data provider is available. It can also be the main site of the repository of that data provider.

- Contact information: An email address of the data provider. By default, this is the info@ email address given on the website of the data provider. It can be more personalised depending on the collections that are made available, e.g. the email address of a person responsible for the particular dataset featured on Europeana Labs.

- Copyright: This is the rights statement for the digital objects forming the dataset, as displayed on the Europeana Portal and stored in edm:rights of the metadata for the objects. If it is a mixed dataset where several open licenses are used, all rights statements need to be given here. It is not necessary to split the dataset into subsets in that case.

- Copyright URL: This is the URL where more information about the rights statement is available. It corresponds to the URL for the rights statement that is used in Europeana.

- Dataset: This is the dataset ID, important to unambiguously identify the dataset in Europeana. In case this number is not known, it is always visible in the object page URL in Europeana. This URL starts with this string 'http://www.europeana.eu/portal/record/' followed by the ID of the dataset this object belongs to.

- Link to portal: To generate the link to the Europeana Portal, the query for the dataset should always be based on the dataset ID, which is less ambiguous than the name of the

data provider. Thus, the basic query to find a specific set always starts with 'europeana_collectionName:datasetID*'. It is important to remember to only use facets to narrow down a search, not keywords or anything else.

- Link to console: The link to the console is created by EF based on the portal URL.

- Categories: The categories for each set are chosen to reflect the thematic focus. In the context of Europeana Creative, the five themes of the project are used as much as possible to indicate the re-use potential of a dataset for one of the five themes. The categories (tags) currently used in Europeana Labs should be used as an inspiration what categories are suitable.

As said above, this process is a recurring task for EF every month to incorporate new datasets that were added to the Europeana repository. As an additional task, existing datasets in Europeana Labs need to be checked and updated manually in Github directly, to reflect any dataset changes like changed URL, thematic extensions or different object numbers.

4.3 Content Requirements

Building the content inventory for creative re-use was based on a demand driven approach during the Europeana Creative project: creative industry partners had very specific content requirements in order to realise their re-use scenario. As elaborated in the lessons learned chapter above, this has disadvantages as the amount of suitable content can be very low. Two solutions to this problem are proposed to ensure a successful collaboration with creative industries:

1. Ensure that content requirements are developed in close collaboration between EF and the creative industry partner. The content sourcing itself should run as an iterative process to develop a common understanding of what is needed and what is available, which also helps to manage expectations on both sides.

2. Choose a supply driven approach when developing creative re-use scenarios. A lot of re-usable content is accessible in Europeana. If this can be packaged and offered to creative industry partners, some of them may develop re-use scenarios based on the content that is actually available. This approach also helps to acknowledge the previous work done by data providers to open up their data and to invest in high data quality. The datasets available in Europeana Labs should support this approach.

As the content requirements are the starting point for any content sourcing, the timing is very critical in order to make content available in time. If no suitable content is available in Europeana, it can take up to several months to make content available (see below sections for more details on that). Therefore, content requirements need to be communicated to EF as early as possible in order to achieve the goals of the creative industry partners.

4.4 Metadata Available in Europeana

Europeana is the starting point for all sourcing of content for creative re-use. Almost 33 million digital objects are accessible via Europeana (as of October 2014) covering a huge diversity of topics and subjects across many cultural heritage domains. The process of evaluating metadata also includes the rich media content that has been collected by Europeana through user-contributed content campaigns and similar mechanisms (e.g., Europeana 1914–1918[25] campaign, Europeana 1989[26] campaign). Of these metadata records, it needs to be evaluated which records refer to media files that correspond to the Content Re-use Framework. These media files are suitable for immediate use by creative industries.

Based on the content requirements of the creative industry partner, EF is making a first inventory of content using My Europeana to save queries and items. The content sourcing is first based on content that corresponds to the Content Re-use Framework. Depending on the content requirements and the availability of suitable content in Europeana, the search is also expanded to content that does not correspond to the Content Re-use Framework. If the creative industry partners prefer to work with such content, EF will evaluate the content and approaches the responsible data provider or aggregator to negotiate how the conditions of the Content Re-use Framework can be met in order to provide access to re-usable media files.

A decision tree was created to facilitate the content sourcing process and to make the necessary negotiation steps transparent (Fig. 13). All negotiations with data providers take a considerable amount of time, as learned from previous experiences. Therefore, the recommendation is to pay attention to the re-use potential of the content and try to restrict the search for content to material corresponding to the Content Re-use Framework.

Given that the investigation and negotiation with a data provider was successful, the metadata records need to be updated in Europeana to reflect the changes and allow the API to retrieve the records and the underlying content. The metadata will be re-harvested and processed by EF, following the regular (monthly) publication cycle of the Europeana ingestion specialists, to be updated in the Europeana database.

When communicating with a data provider and planning the harvesting, it is important to have the Europeana aggregation model in mind. Europeana is communicating initially with aggregators who have to communicate with their data providers. Also metadata records that need to be updated in Europeana first need to be updated on the aggregator side before they can be updated in Europeana. This again adds to the time needed to update metadata records in Europeana.

---

[25] http://www.europeana1914-1918.eu; accessed November 28, 2014.

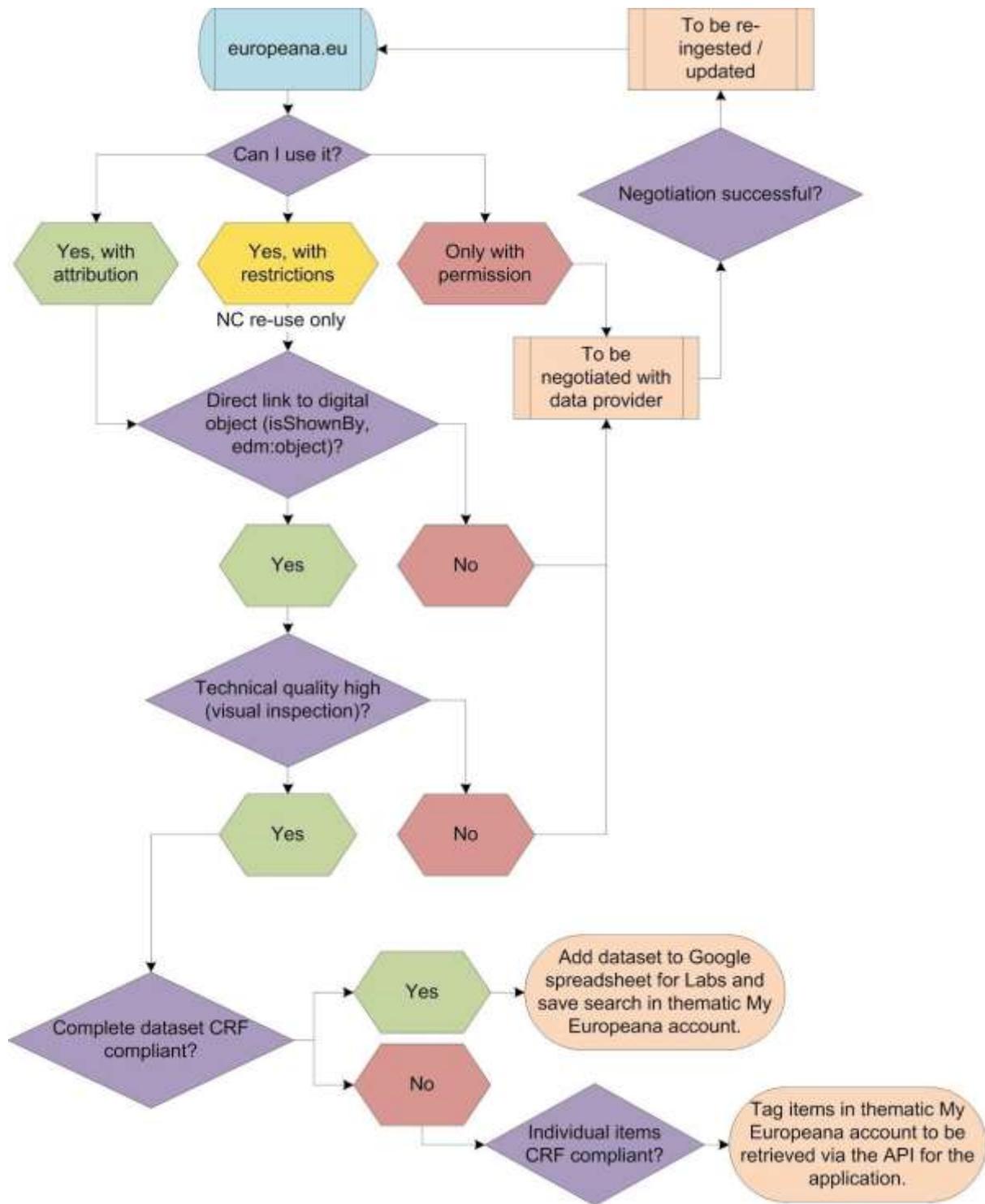[26] http://www.europeana1989.eu; accessed November 28, 2014.

Fig. 14: Decision tree for content sourcing

4.5   New Content from within the Consortium

If no content is available in Europeana that is suitable for the re-use scenario, new content that can be offered by existing data providers needs to be identified. The reasons for first approaching existing data providers and not reach out to the wider Europeana Network are (1) EF has an established relationship with the data provider, (2) EF knows the content type and data quality potentially available from existing data providers, (3) existing data providers know the formal and technical process to enable a shorter turnaround time when submitting data to Europeana. This approach was tested during the Europeana Creative project as several consortium partners are established data providers: Austrian National Library (ONB), the British Library (BL), the Museum für Naturkunde (MfN), the National Museum, Prague (NMP) and the Netherlands Institute for Sound and Vision (NISV). Each of these institutions have pledged to provide content suitable for experimentation and their progress in providing new data during the project time is described in one of the following chapters.

4.6   New Content from outside the Consortium

If no content is available in Europeana or from existing data providers that is suitable for the envisaged re-use scenario, new content needs to be sourced from external parties. Europeana does not launch an open call to the Europeana Network or wider ecosystem yet in order to achieve this. Rather EF (in collaboration with the creative industry partner) approaches potential data providers individually and negotiates the content to be re-used in a one-to-one manner. In some cases creative industry partners may already have the necessary contacts to potential new data providers and have negotiated the conditions with them for making content available to Europeana. EF is particularly relying on the subject expertise of the creative industry partner and their existing relationships with potential data providers to efficiently source suitable content.

Approaching providers individually is recommended as the full implications of the Content Re-use Framework are changing the ambitions of Europeana, currently being an aggregator of metadata and not an aggregator of content. Therefore, the re-use of actual content is a sensitive topic that Europeana needs to communicate carefully within the Europeana Network in order not to jeopardise existing relationships with data providers and aggregators. Attractive case studies demonstrating the successful re-use of content (e.g., Europeana Creative Pilots) are considered an important incentive for data providers to not only open up their metadata but also open up their content. A Europeana Network Task Force was launched in October 2014 to work on a strategy to increase the amount of re-usable content in Europeana and develop case studies to incentivise data providers to improve their data for re-use.

Once a potential data provider is identified and the general requirements for the re-use scenario are discussed and agreed, the data provider will enter the established provide-data workflow of Europeana to formalise the relationship with the data provider or refer the data provider to the appropriate aggregator or project to deliver data to Europeana. Eventually, the metadata are harvested and ingested by Europeana to make them available via the Europeana channels,

including the API. Again, the time constraints outlined above need to be considered, too, so considerable time is needed to make new content from new data providers available in Europeana.

5. Partner Progress on Preparing and Providing Data

Europeana Foundation is leading and coordinating Task 1.4 of Europeana Creative, which is the sourcing and preparation of content for the Pilot themes. A number of Europeana Creative partners are involved in this task to support the process and provide content, including the Austrian National Library (ONB), the British Library (BL), the National Museum Prague (NMP), the Museum für Naturkunde (MfN) and the Netherlands Institute for Sound and Vision (NISV).

5.1 Progress of the Austrian National Library

ONB revised all its provided collections to Europeana and checked it against the requirements of the Content Re-use Framework. After a process of evaluation regarding re-usable content in the perspective of rights and regarding technical possibility to provide the selected material, ONB made its decision on what kind of content to provide and how to do so.

In September 2014 the implemented changes in the providing system at ONB were finalised and the selected collection could be re-delivered to Europeana. The following collections have been re-ingested by The European Library:

- [WWI photographs from Austria-Hungary](#)[27]: More than 37,000 photographs illustrating the Great War at the eastern and south eastern front, the back country and further areas of the former Monarchy.

- [WWI text posters from Austria-Hungary](#)[28]: A collection of about 6,500 text posters from all parts of the Austrian-Hungarian monarchy including the important Emperor manifestos, announcements managing mobilisation, administration and shortages with information on dates and places.

- [Portrait Collection of the Austrian National Library](#)[29]: A selection of more than 150,000 portraits of the portraits collection of the Austrian National Library including European and

---

[27] http://www.europeana.eu/portal/search.html?query=europeana_collectionName%3A92070*&rows =24&qf=RIGHTS%3Ahttp%3A%2F%2Fcreativecommons.org%2Fpublicdomain%2Fmark%2F1.0%2 F*; accessed November 28, 2014.

[28] http://www.europeana.eu/portal/search.html?query=europeana_collectionName%3A9200290_Ag_ EU_TEL_a0564b_EUcollection14- 18&rows=96&qf=RIGHTS%3Ahttp%3A%2F%2Fcreativecommons.org%2Fpublicdomain%2Fmark% 2F1.0%2F*; accessed November 28, 2014.

[29] http://www.europeana.eu/portal/search.html?query=europeana_collectionName%3A92062*&rows =24&qf=RIGHTS%3Ahttp%3A%2F%2Fcreativecommons.org%2Fpublicdomain%2Fmark%2F1.0%2 F*; accessed November 28, 2014.

non-European regent dynasties, divided into categories like popes, emperors, kings, nobles, Austrian dukes and other kings.

- "Fideikommiß"-Vues of the Austrian National Library[30]: A selection of more than 3,500 topographical recordings, images of cities, villages, buildings, architecture, landscapes and pictures of natural events from the areas of the former Austrian crown-lands, but also from modern day Germany, Italy, France, England, Spain, Switzerland and The Netherlands.

The ingestion process and the consequent delivery to Europeana went very well and until the end of 2014 all the provided collections will be available via Europeana Labs.

Further to the above collections, ONB (with the input of SAT) selected 202 images of book illustrations from the 15th and 17th century based on the criteria of shape and colour: Fighting knights from the "Book on Swordsmanship and Wrestling", colourful images of parrots, amazing studies of plants and flowers from the "Florilegium of Prince Eugene of Savoy", fish and other animals from the "Marine Fauna of the Adriatic Sea" and further nature studies from Dutch, German and Italian artists. This content is currently being processed by The European Library and will be available in Europeana in a few weeks for the Design Theme.

5.2 Progress of the British Library

The British Library has provided 3,000 sounds to Europeana, mainly from wildlife and sound collections:

- British Wildlife Recordings (640 tracks)

- A.R. Gregory (869 tracks). Birdlife recordings from Kenya in the 1960s and 1970s.

- Listen to Nature (318 tracks)

- Sounds and Nature (61 tracks)

- Weather (111 tracks)

- Water (54 tracks)

- UK Sound Map (1001 tracks).

These were crowd sourced recordings made from 2010-2012 and accessible on the Audioboo platform.

---

[30] http://www.europeana.eu/portal/search.html?query=europeana_collectionName%3A92070*&rows=24&qf=RIGHTS%3Ahttp%3A%2F%2Fcreativecommons.org%2Fpublicdomain%2Fmark%2F1.0%2F*; accessed November 28, 2014.

As mentioned in the "Content Inventory for Social Networks Pilot" above, 40 of these recordings are on the SoundCloud platform for the purpose of the Social Networks Pilot. Aside from the UK Sound Map tracks on CC-BY-NC, all tracks on a CC-BY license will be updated so that they can be downloaded directly from Europeana in December 2014.

5.2.1   British Library Flickr Photostream

The British Library agreed to submit further content to the project and the 1,000,000 images exported to the BL's flickr photostream were identified as re-usable, as they were high resolution images and in the public domain. The metadata of the images was sparse: the book's title and author from which each image came does not say much about the image. Fortunately, flickr users were tagging these images - so far, around 100,000 tags have been added to 40,000 images - and these tags give the images much more meaning. For example, an illustration of a species of bird would at least be tagged with "wildlife" and, perhaps, also the name of the species.

After some difficulties in exporting to The Europeana Library (TEL) in EDM format, it was decided to export in ESE format as with the 3,000 audio tracks. TEL would then convert to EDM and export the images to Europeana. A test run of 744 images have now been successfully uploaded to TEL (Fig. 14)[31].

The first tag of each image (portrait, tattoo, Blacksmith, etc.) and the book title are used as the title of the Europeana record. The description of the image record details the book from which it was scanned and how tags were user generated. Having successfully uploaded the test run to TEL, all tagged images will be exported to TEL in November and subsequently to Europeana in December 2014.

5.2.2   Catalogue of Illuminated Manuscripts

Another collection that will be submitted is the Catalogue of Illuminated Manuscripts - around 37,000 images from the BL's medieval manuscripts collections. These are high resolution images which are in the public domain. Unlike the BL's flickr images, the metadata is extremely dense and therefore will be of educational potential as well as re-use potential on Europeana. However, exporting this collection to Europeana will not be as straightforward as the BL flickr images, as some of the relevant information pertains to the images themselves whereas other information pertains to the books from which they originated.

---

[31] http://www.tel.ulcc.ac.uk/acceptance/search?query=a1225&locale=fr&offset=600; accessed November 28, 2014.
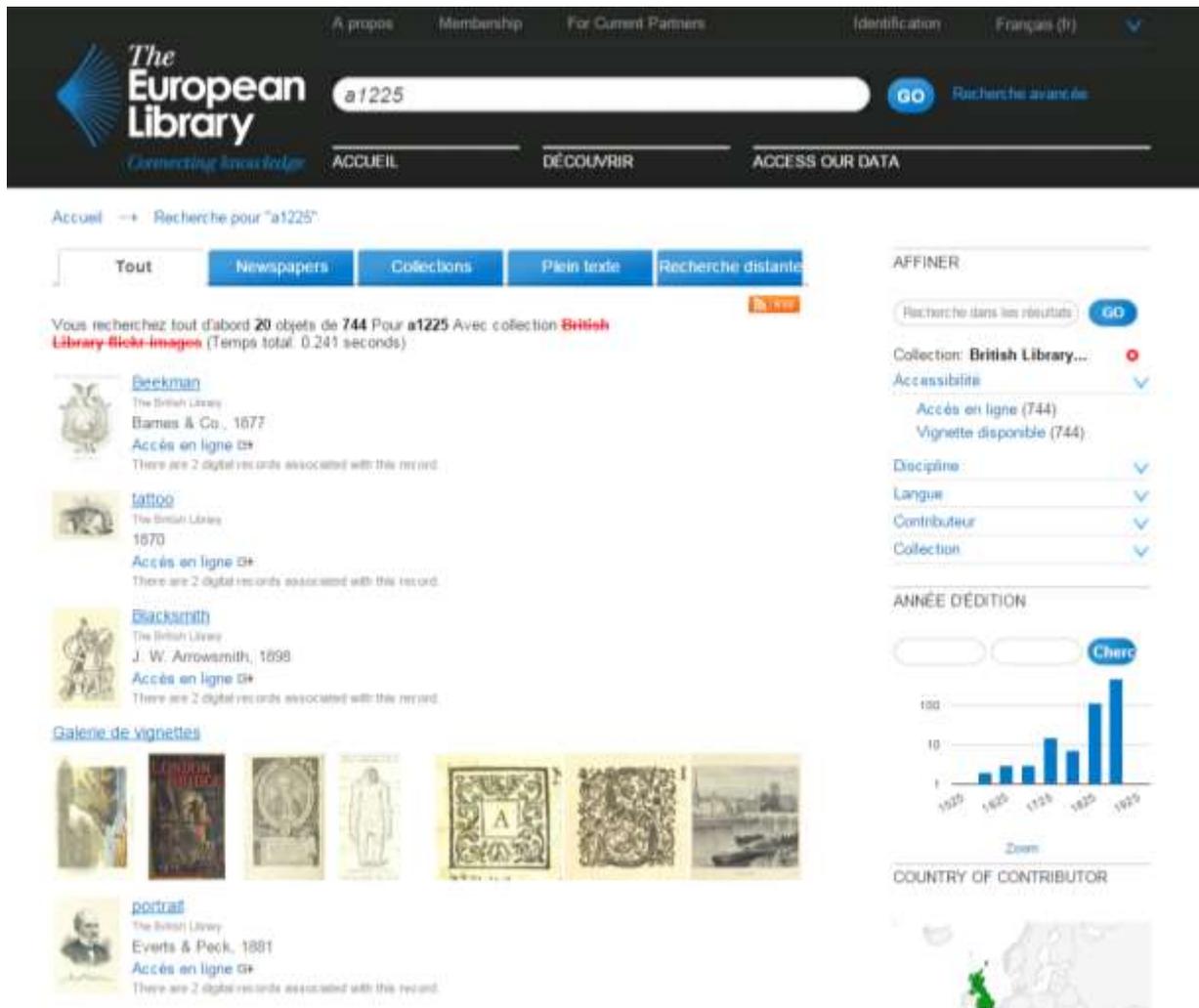
Fig. 15: Test ingest of 744 items of the BL flickr stream into The European Library

5.3 Progress of the Museum für Naturkunde

5.3.1 Content Provision MfN

MfN was partner and data provider in the OpenUp! project (Opening up natural history content for Europeana: http://open-up.eu/) which was successfully completed in February 2014. The developed and well-established infrastructures are still used by European natural history institutions involved in the project. The first harvest from MfN data after the project time took place in June 2014. With each ingest, the number of multimedia objects as well the diversity constantly increases and 32,500 objects from MfN will be available after the next Europeana publication in December 2014 (see Table 2; due to a problem with the MfN dataset, it was

disabled in October 2014 and will be re-ingested in January 2015 publication). Content from MfN is diverse and covers different topics as geology (Fig. 17), zoology (Figs. 15, 16 and 20), palaeontology (Fig. 19) and entomology (Figs. 16 and 20) as well as zoological sounds (Fig. 15). Various multimedia objects provided by MfN were used in the Natural History Education Pilot (see Table 1), but the content was not digitised or mobilised explicitly for the Europeana Creative project.

Table 2: MfN Content in Europeana in January 2015.

Collection	Topic	Type	No of records
Animal sound archive	Zoology, Entomology	Sound	11,300
Global Butterfly Information System (GLOBIS)	Entomology	Image	2,362
Mineral Collection	Geology	Image	7,200
Ehrenberg Drawings	Zoology	Image	3,555
Palaeontological Collections	Plant, Zoology, Entomology, Palaeontology	Image	4,991
Aves (recent)	Zoology	Image	464
Verhoeff Myriapoda Collection	Zoology	Image	834
Diptera Collection	Entomology	Image	346
Orthoptera Collection Dorsa	Entomology	Image	3,276
Drawers*	Entomology	Image	3,575

* collection not yet on Europeana, but content already visible via:
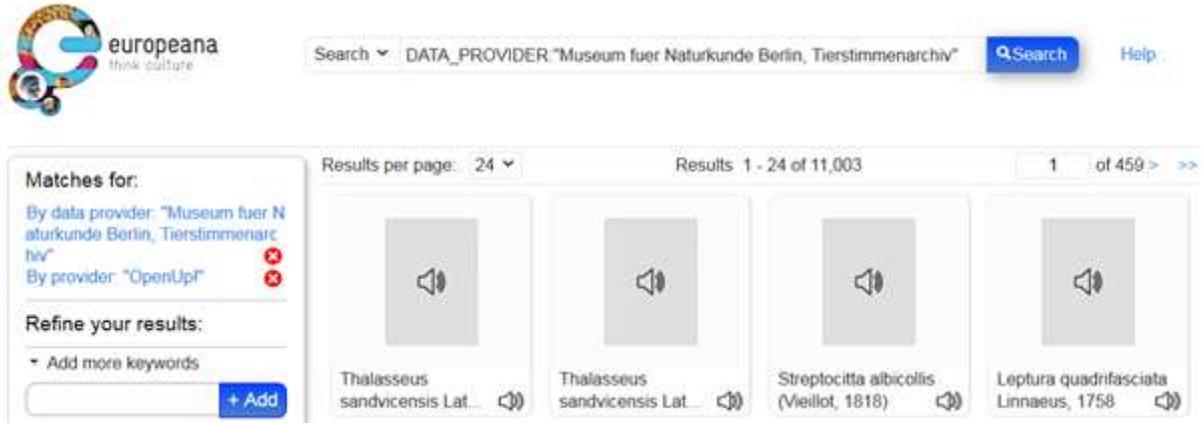http://eos.naturkundemuseum-berlin.de/collections

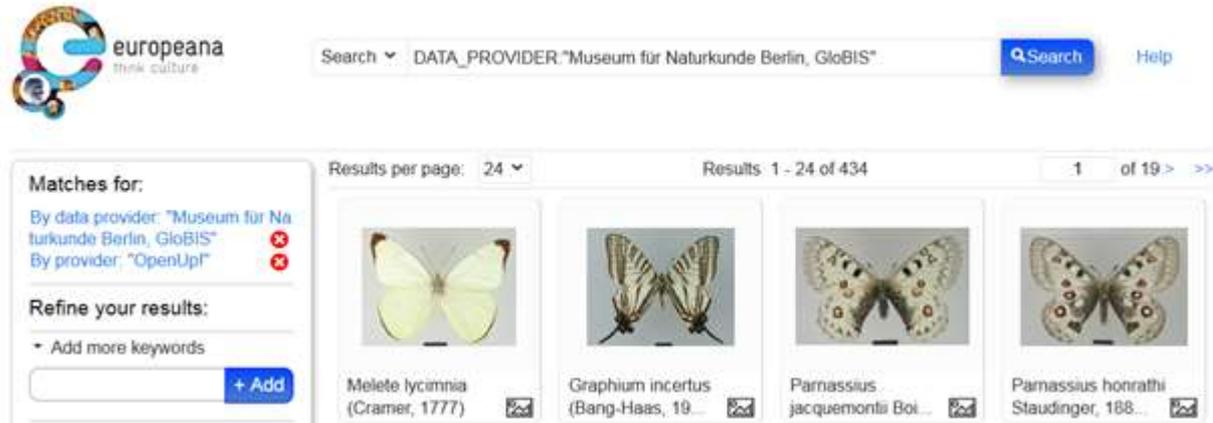**Fig. 16: Animal Sound Archive from MfN in the Europeana portal**

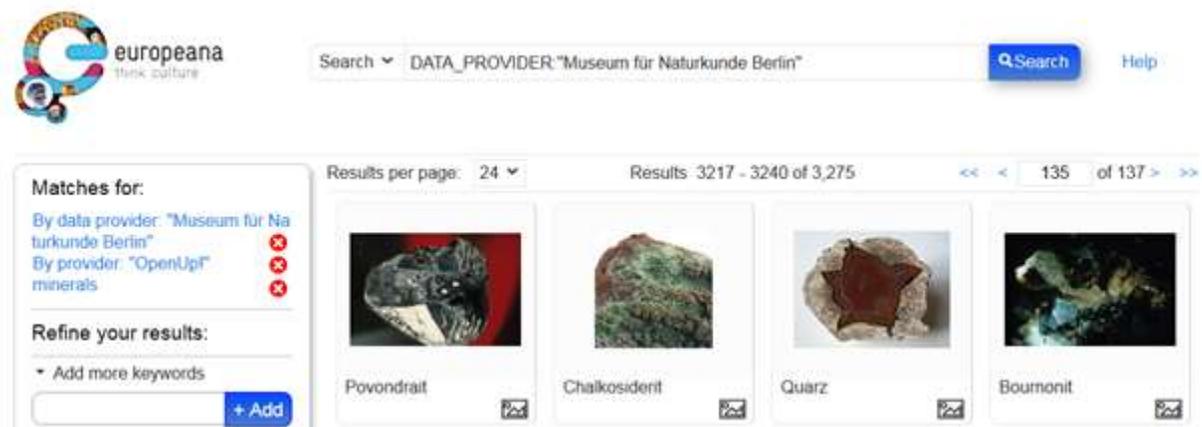**Fig. 17: Butterflies (GloBIS) from MfN in the Europeana portal**

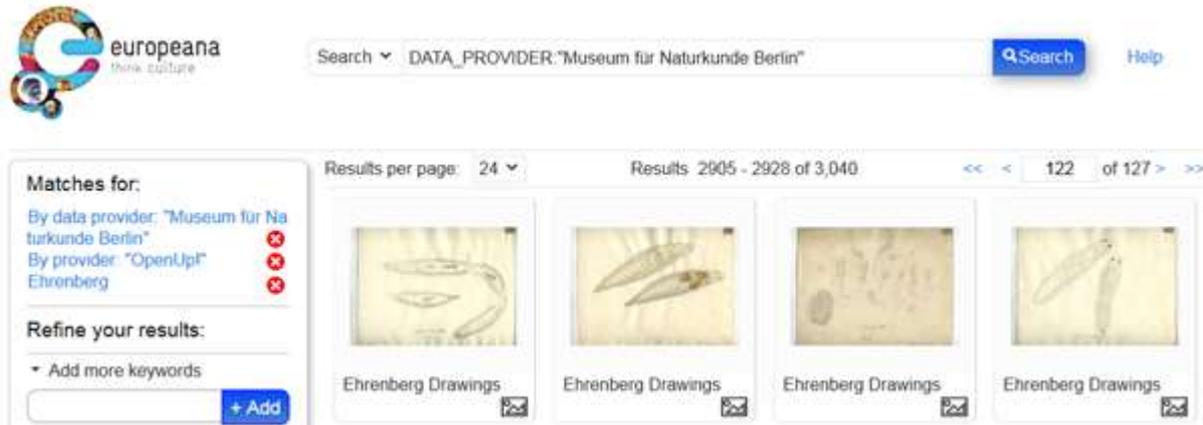**Fig. 18: Minerals from MfN in the Europeana portal**

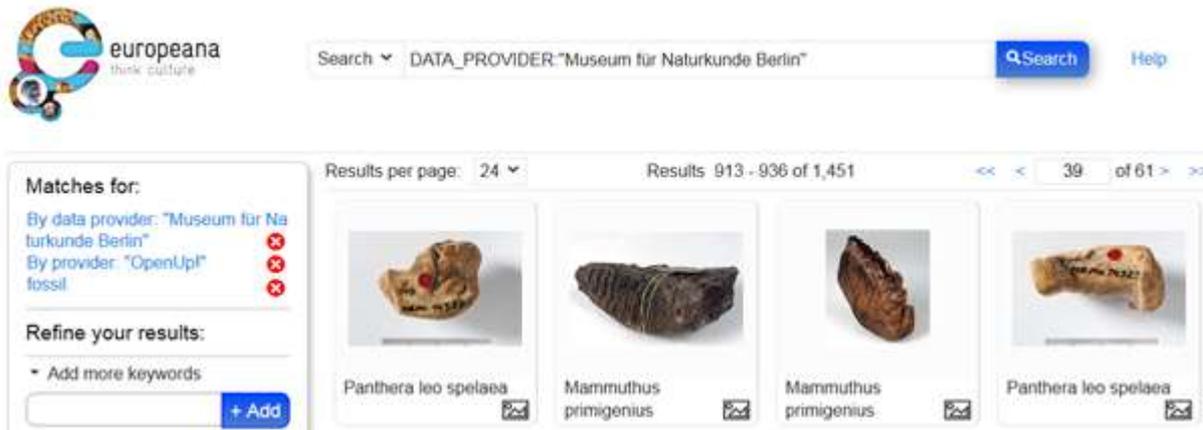**Fig. 19: Historical scientific drawings (Ehrenberg) from MfN in the Europeana portal**

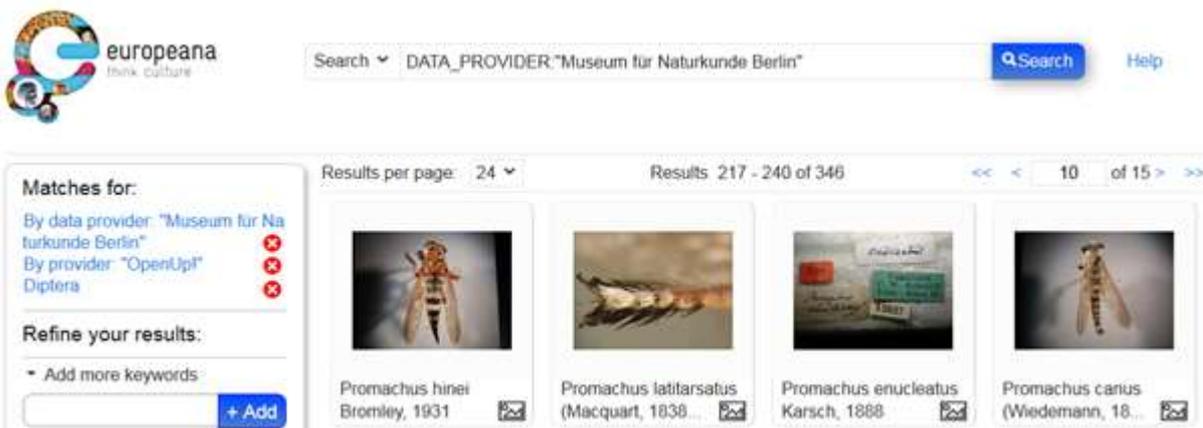**Fig. 20: Fossils from MfN in the Europeana portal**

Fig. 21: Flies (Diptera) from MfN in the Europeana portal

5.3.2 Licensing and Content Policy at MfN

A general content policy for multimedia objects at MfN is still in progress. The preparation takes longer than expected, as the legal status of some objects still poses questions. This particularly concerns the German copyright law for images created during the digitisation of collection items. Digitisation staff needs to be asked to abandon possible rights, e.g. the right to be named in order to make this material available under CC0. An agreement between the museum and digitisation staff was drafted to solve this problem and is now under review. This agreement will be part of a collection of agreements and contracts dealing with intellectual property rights at museums and archives in Germany. The collection is currently compiled by the service point digitisation Berlin (digiS)[32] in collaboration with a law office specialised in legal questions of the digital world and should later on be freely available for German GLAM institutions.

5.4 Progress of the National Museum Prague

National Museum Prague was like MfN a partner in the OpenUp! project (Opening up natural history content for Europeana: http://open-up.eu/) and within this project provided 8,270 metadata records to Europeana, which are linked to more than 14,000 objects. For many collection objects, digital representations of the labels were added as well, to have the complete set of information available in the collections of the museum. Working with the Europeana Data Model (EDM) allowed NMP to add several web resources to one metadata record, to make it clear what information belongs to which object. All natural history content is available under a CC-BY license, so it is suitable for re-use. Although the number of objects is not very large, it is very diverse and includes paleontological, botanical, entomological, zoological, mineralogical, mycological and anthropological collections (see Figs. 21-25).

The Natural History Museum is one section within the National Museum. Other sections include the History Museum, Czech Museum of Music, Naprstek ethnographic museum and a large museum library. Other museums under the National Museum are also making their collections online accessible via the platform eSbírky (www.eSbirky.cz), which works as a digital cultural heritage aggregator and also provides content to Europeana. This content is in most cases also licensed for re-use and some content is already foreseen to be re-used by the Design Pilot of Europeana Creative. The majority of the eSbírky is provided to Europeana via the AthenaPlus project. For the next harvest, which will be on November 25, 2014, eSbírky will provide 40,308 records. These records will link to 65,340 images, 1,628 audio files and 11 videos. The work on the eSbírky platform is coordinated by the National Museum, but contains content from about 40 memory institutions in the Czech Republic. This number will significantly grow in the near future after the launch of the new portal with new functionalities. The eSbírky platform is supported by the Czech Ministry of Culture, which is preparing to establish eSbírky as the national aggregator for digital cultural heritage and treasure in the Czech Republic.

---

[32] http://www.servicestelle-digitalisierung.de/confluence/pages/viewpage.action?pageId=917513; accessed November 28, 2014.
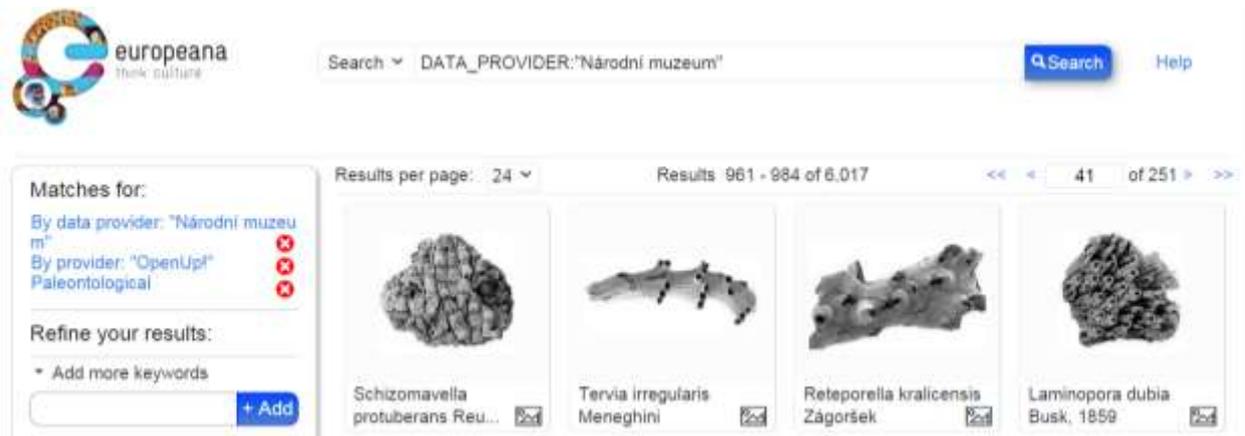
Fig. 22: Example from the NMP paleontological collections in Europeana
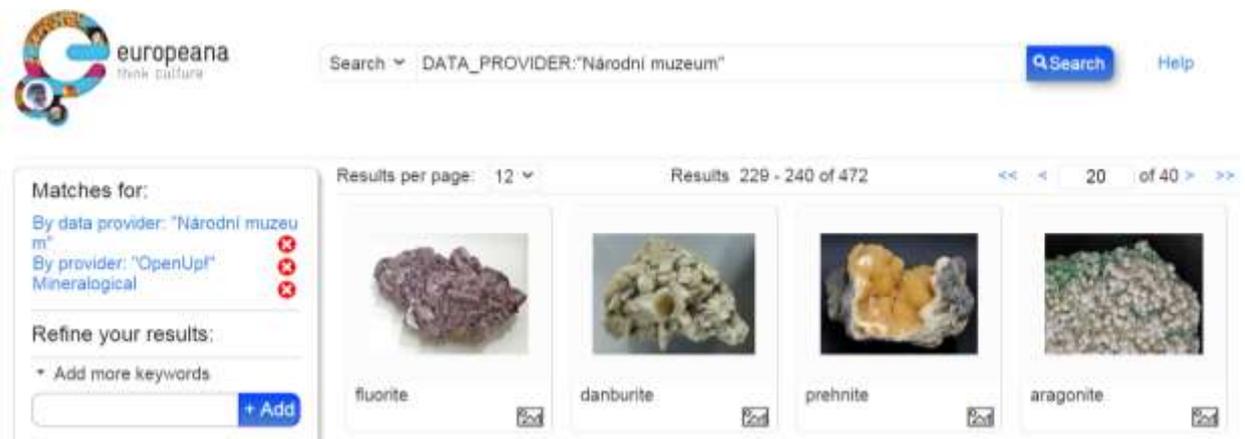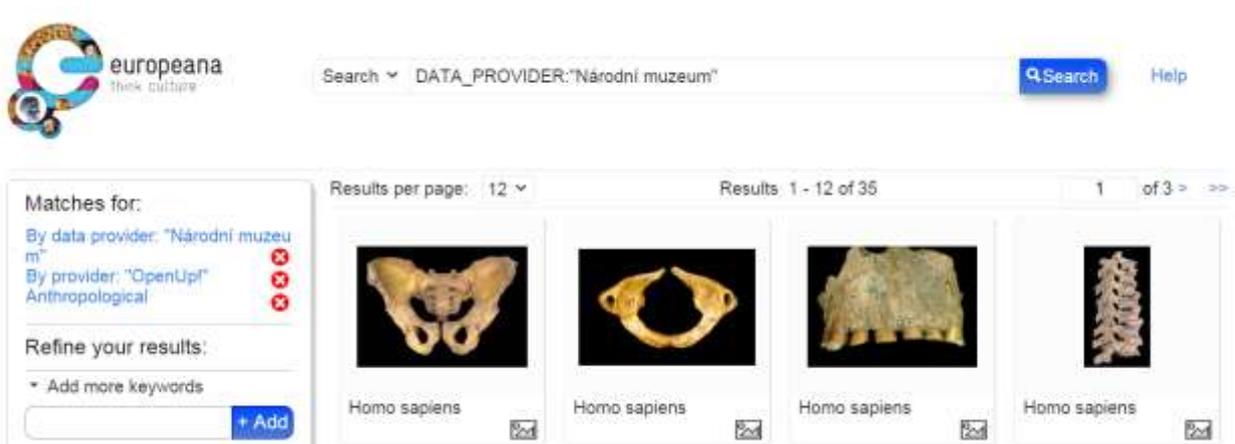

Fig. 23: Example from the NMP mineralogical collections in Europeana

Fig. 24: Example from the NMP entomological collections in Europeana

Fig. 25: Example from the NMP botanical collections in Europeana

Fig. 26: Example from NMP anthropological collections in Europeana

5.5 Progress of the Netherlands Institute for Sound and Vision

NISV has made available two datasets to Europeana and Europeana Labs as part of the activities in Europeana Creative. Both collections have been made available via the National Dutch aggregator (Digitale Collectie) to Europeana and both datasets are available under an open license (Creative Commons) so they are available for re-use.

The collections are:

- Open Images (Open Beelden) - www.openbeelden.nl: This dataset includes nearly 3,000 videos available via the Open Images platform of the Netherlands Institute for Sound and Vision. The collection contains newsreels from the Polygoon collection and several other

films on the Netherlands in the twentieth century. It contains historical news events, with topics such as art, culture, transportation, health and healthcare, nature, politics, sports, pets and fashion. Link to the dataset on Europeana[33] and Europeana Labs[34].

- Sound of the Netherlands (Geluid van Nederland) - www.geluidvannederland.nl: This dataset includes about 2,500 sounds from the Netherlands Institute for Sound and Vision of daily life available via the Sound of the Netherlands portal. The set includes sounds of daily life, birds, city sounds, etc. and they all have geo-location. Link to the dataset on Europeana[35] and Europeana Labs[36].

---

[33] http://europeana.eu/portal/search.html?query=europeana_collectionName%3A2021601*&rows=24&qf=REUSABILITY%3Aopen; accessed November 28, 2014.

[34] http://labs.europeana.eu/data/newsreels-from-sound-and-vision-on-the-netherlands-in-the-20th-century/; accessed November 28, 2014.

[35] http://www.europeana.eu/portal/search.html?query=europeana_collectionName%3A2021613*&rows=24; accessed November 28, 2014.

[36] http://labs.europeana.eu/data/sounds-of-daily-life-and-birds-from-netherlands-institute-of-sound-and-vision/; accessed November 28, 2014.

6. Conclusions and Outlook

The content sourcing for creative re-use and selection of suitable content via Europeana still involves a lot of manual work. This is not scalable for future re-use scenarios, where creative industry partners want to quickly access content to work with. The full implementation of the Content Re-use Framework as a filter in Europeana is currently underway to surface the most suitable content for creative re-use.

High data quality is a very important prerequisite to make re-use scenarios based on digital cultural heritage content work. The creative industry can help significantly to incentivise the opening up of content that is fit for purpose. Case studies are identified as a major component to show the potential of creative re-use and promote the benefits for cultural heritage institutions. Collaborative projects between cultural heritage organisations and creative industry partners can further help to make these case studies work for both sides.

The Europeana Network Task Force on content re-use is working on a more strategic approach to increase the amount of content in Europeana that is fit for purpose. It will work on the licensing and the technical standards and their value to data partners. The outcome should be a framework, a statement of shared understanding that declares how this is seen working in practice as the Europeana Network, how quality is defined and what instruments and mechanisms are needed to distinguish between service levels for data partners.