# EUROPEANA SOUNDS

## Project Number: 620591

# MS7 End-user contributions defined

**Document Identifier:** EuropeanaSounds-MS7-End-user-contributions-defined-v1.0.docx

**Document link:** http://pro.europeana.eu/web/europeana-sounds/documents

**Date:** 29/10/2014

## Abstract

This Milestone provides the definition of the different types of end-users that are addressed and the different types of user contributions that the project expects to support. This Milestone describes the primary focus and how these types of end-users and user contributions are currently understood.

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

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

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

## I. COPYRIGHT NOTICE



## II. REVISIONS

| Version | Status | Author | Partner | Date | Changes |
|---|---|---|---|---|---|
| 0.1 | ToC | Maarten Brinkerink | NISV | 16/06/2014 | |
| 0.2 | Draft | Maarten Brinkerink | NISV | 07/07/2014 | First draft for review |
| 0.3 | Draft | Harry van Biessum | NISV | 08/07/2014 | Changes to 2.2 |
| 0.4 | Draft | Maarten Brinkerink | NISV | 25/07/2014 | Changes in response to review |
| 0.5 | Draft | Gregory Markus | NISV | 15/09/2014 | Additions to 3 |
| 0.6 | Draft | Lise den Brok | WAWWD | 19/09/2014 | Additions to 2 |
| 0.7 | Draft | Maarten Brinkerink | NISV | 15/10/2014 | Changes in response to feedback |
| 1.0 | Final | Maarten Brinkerink | NISV | 28/10/2014 | Final version |

## III. DELIVERY SLIP

| | Name | Partner/WP | Date |
|---|---|---|---|
| **Document Author** | Maarten Brinkerink<br>mbrinkerink@beeldengeluid.nl | NISV / WP2 | 28/10/2014 |
| **Reviewed by** | **Reviewers:**<br>David Haskiya<br>Breandán Knowlton<br>Valentine Charles | <br>EF / WP4<br>WAWWD /WP2<br>EF/WP4 | 15/09/2014 |
| **Approved by** | **Coordinator & PMB** | | 29/10/2014 |

## IV. DISTRIBUTION

| No. | Date | Comment | Partner / WP |
|---|---|---|---|
| 1 | 30/10/2014 | Submitted to the Europeana Commission | BL/WP7 |
| 2 | 29/10/2014 | Posted on Europeana Pro | BL/WP7 |
| 3 | 30/10/2014 | Distributed to the Project Consortium | BL/WP7 |

## V. APPLICATION AREA

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

## VI. DOCUMENT AMENDMENT PROCEDURE

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

## VII. TERMINOLOGY

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

Further terms are defined below as required:

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

## VIII. PROJECT SUMMARY

Europeana Sounds is Europeana's 'missing' fifth domain aggregator, joining APEX (Archives), EUscreen (television), the Europeana film Gateway (film) and TEL (libraries). It will increase the opportunities for access to and creative re-use of Europeana's audio and audio-related content and will build a sustainable best practice network of stakeholders in the content value chain to aggregate, enrich and share a critical mass of audio that meets the needs of public audiences, the creative industries (notably publishers) and researchers. The consortium of 24 partners will:

- Double the number of audio items accessible through Europeana to over 1 million and improve geographical and thematic coverage by aggregating items with widespread popular

appeal such as contemporary and classical music, traditional and folk music, the natural world, oral memory and languages and dialects.

- Add meaningful contextual knowledge and medium-specific metadata to 2 million items in Europeana's audio and audio-related collections, developing techniques for cross-media and cross-collection linking.
- Develop and validate audience specific sound channels and a distributed crowd-sourcing infrastructure for end-users that will improve Europeana's search facility, navigation and user experience. These can then be used for other communities and other media.
- Engage music publishers and rights holders in efforts to make more material accessible online through Europeana by resolving domain constraints and lack of access to commercially unviable (i.e. out-of-commerce) content.

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

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

## IX. STATEMENT OF ORIGINALITY

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

# X. EXECUTIVE SUMMARY: END-USER CONTRIBUTIONS DEFINED

This Milestone provides the definition of the different types of end-users that are addressed and the different types of user contributions that the project expects to support.

To establish maximum impact a differentiation between “experts” and “non-experts” (general public) for the crowdsourcing micro-tasks is proposed, in order to be able to distinguish between tasks that range from simple tasks for a broad audience, to tasks that require certain domain knowledge and target a more specific audience. These two categories have been further developer into culture snackers and culture vultures.

In relation to the end-user contributions to enrichment a distinction into three types of contributions can be made: User-generated metadata, subjective contributions and user profile related interactions.

The first category of end-user contributions consists of contributions that enrich the aggregated collections on the metadata level. The second category of end-user contributions refers to the social and personal interaction with the content on the object level. The last category of end-user contributions includes the - mostly implicit - relations and user interactions with the content that stem from the functionality offered by Europeana to the end-user through the “My Europeana” user profile.

# TABLE OF CONTENTS

# 1 INTRODUCTION

As stated in the Description of Work (DoW, Part B, page 9) [REF 1] end-user engagement lies at the very heart of Europeana Sounds. This Milestone provides the definition of the different types of end-users that are addressed and the different types of user contributions that the project expects to support. Please note that this provides a first categorisation and description of end-user types and user contribution types, which will be further refined and elaborated on over the course of the project (for instance within upcoming Milestones and Deliverables). This Milestone describes the primary focus and how these types of end-users and user contributions are currently understood.

# 2 END-USERS

Europeana Sounds has a significant challenge to face with regard to metadata quality. Data providers are using various standards of metadata. Some Europeana Sounds data providers will supply high quality descriptive metadata while descriptions from other data providers will be of poor quality. The DoW explains how the scale of the problem - "potentially more than 1.5m items of metadata describing just the audio content" - calls for innovative means to enrich existing metadata and generate new metadata, including "user participation" (DoW, Part B, page 3):

> *"User participation provides access to the vast knowledge that resides in each partner's community of interest and helps GLAMs to redefine their role as custodians of knowledge transfer in the digital domain. "*
>
> (DoW, Part B, page 3)

In order to engage end-users in the enrichment through crowdsourcing a "distributed infrastructure that enables crowdsourcing to happen in online environments with an existing crowd" will be set up under the supervision of Work Package 2 (WP2). D2.1 *Crowdsourcing infrastructure and exchange policy* and D2.2 *Functional design of semantic enrichment* will document the (functional) requirements for this in more detail. To establish maximum impact with this endeavour, the DoW explains that a differentiation between "experts" and "non-experts" (general public) will be made, in order to be able to design different "micro-tasks" that range from "simple tasks" for a broad audience, to tasks that require certain "domain knowledge" and target a more specific audience (DoW Part B, page 13).

## *2.1 General public*

The general appeal of sound and music to a wide audience is already alluded to in the DoW, as is the growing importance of the Internet as a distribution mechanism and point of access to sound and music (DoW Part B, page 19). It also mentions the "general public" as being the "Key Audience for Europeana, seeking to improve user engagement" (Part B, page 39).

As recently reported in Chenchen Shen's paper "Design for User Engagement on Europeana Channels" Google Analytics shows that the main visitor of the Europeana search portal "fall into the age groups of 25-34 (33.50%) and 18-24 (27.50%), followed by the range of 45-55 (12.50%)", while a survey shows that 48.06% of the respondents, indicated to be the most active users, are also 25-34 years old, followed by the 35-44 and 45-54 age segments[1]. The research also shows that there is a fairly high level of education among the current end-users with 64% having a university degree (Final report, page 20). It therefore seems justified to keep these characteristics in mind when talking about the general public in relation to Europeana and Europeana Sounds.

[1] Shen, C. "Design for User Engagement on Europeana Channels". Delft University and Europeana, Master of Science Graduation Project-Design for Interaction. 2014.

Another valuable insight from Shen is the distinction between "culture snackers" and "culture vultures", where the "culture snacker" is described as following (and most relevant to the general public category):

> *"They enjoy viewing culture-related content on the Internet, and share it with friends and followers, but the information on cultural heritage is just part of their daily information consumption. [...] For them, the overall pleasant experience sometimes overweighs the quality of content."*
>
> (Final report, page 20)

To make it a bit more concrete, culture snackers from now on are perceived as users with a casual interest in Europeana or the type of content it serves. In the context of Europeana Sounds special attention will go out to the culture snackers with a specific interest in audio and/or music.

The survey from Shen also shows some positive signs for the potential of enrichment through crowdsourcing by the general public, since it has indicated that 49.0% of the respondents are interested to "share my knowledge, expertise and insights", followed by "create or participate in a discussion topic I am interested in", "translate interesting content to my own language" and "recommend relative materials on some content", which also has more than 40% support rate (Final report, page 27). If designed cleverly, all of the above can be translated in enrichment micro-tasks.

## *2.2 Experts*

While it might seem tempting to define experts in relation to Europeana Sounds in a narrow way, the discussion above informs us that this is not the way to go. While the DoW explicitly mentions "Researchers" as a target group (Part B, page 39), and one might be tempted to map that one-on-one to the category of an expert end-user, the connection to academia is not the only prerequisite for being able to fulfill a specific micro-task that requires domain knowledge. And conversely, the Final report shows that 75.5% of the current users of the Europeana Portal want to use the resources of Europeana for their academic life (Final report, page 25). However, that doesn't mean that all these students, scholars and researchers using Europeana can always be motivated to participate in expert crowdsourcing enrichment tasks.

The concept of "culture vultures", as mentioned by Shen above, might be a more informative approach to distinguishing experts from the general public:

> *"They are the culture enthusiasts and professionals. They have a strong interest in cultural heritage and probably a good knowledge in a specific area(s). They are likely to work professionally with culture in one form or another, or to be a lifelong culture enthusiast, including researchers, students, professionals and interested laymen. While having a broad general interest a culture vulture has a special interest in, and knowledge of, one or a small number of specific topics, subjects, styles or genres."*

(Final report, page 20)

Again, to make this idea a bit more concrete, culture vultures from now on in this document are perceived as users who work professionally with the Europeana portal, or the type of content it serves. And again, special attention will be focused on the culture vultures with a specific interest in audio and/or music, in the context of Europeana Sounds.

Please note that this concept does not require an expert (or culture vulture) to be a 'professional'. A culture vulture can also be an enthusiast or amateur interested in and knowledgeable about a (few) specific 'niche' topic(s). For those familiar with, for instance, Internet fora, this is probably not an uncommon phenomenon. The report goes on to explain that culture vultures can come from any domain. This concept that culture vulture do not necessarily need to be professionals is also illustrated by Historypin[2]'s intended audience - traditional Irish and Scottish musicians. Based on preliminary observational research on the website The Session[3], it seems that traditional Irish music enthusiasts engage with traditional culture mostly in their spare time, but are very knowledgeable and keen to learn more. For them, the social aspect of teaching and learning about traditional culture is also very important. It's also important to realize that culture snackers and culture vultures are not static or mutually exclusive roles. Depending on contextm a single user is sometimes a snacker and sometimes a vulture.

Achieving successful enrichment through crowdsourcing by experts in the context of the Europeana Sounds project depends on identifying the relevant domains in relation to the collections that will be aggregated and on finding successful ways to engage with these experts. This could be achieved by providing them with micro-tasks that appeal to their specific interest and domain knowledge, for example. Within the SEALINC media project[4] this is referred to as "nichesourcing" (Oosterman et al, 2014).

The point of departure for identifying the areas of expertise are the underlying content categories mentioned in the DoW: Classical music, Traditional music and oral testimony, Popular and contemporary music, Natural sounds, Languages and oral testimony (DoW Part B, page 20-21). The DoW also already hints at the potential - and challenge - of the communities that the Data Providers already engage with:

> *"Another crucial impact factor is that user participation provides access to the vast knowledge that resides in each partner's community of interest and helps GLAMs to redefine their role as custodians of knowledge transfer in the digital domain."*
>
> (DoW Part B, page 13)

[2] http://www.historypin.org
[3] http://www.thesession.org
[4] http://www.commit-nl.nl/projects/socially-enriched-access-to-linked-cultural-media

In response to the 'Data Providers Survey on Enrichments through Crowdsourcing'[5] issued by WP2 in May 2014 (M4), the Data Providers made clear that they are indeed involved with numerous communities already. One of the results of the survey is that the Data Providers mainly focus on researchers (named 10 times explicitly) and students (named 7 times explicitly), who can be considered experts in the audience dichotomy this Work Package employs. Furthermore the following communities were named as currently engaging communities by the 14 Data Providers that responded: teachers (named 7 times), musicians (6 times), general public (4 times), artists (2 times), the library community (2 times), the media community (2 times) and other specific communities were also named (6 times).

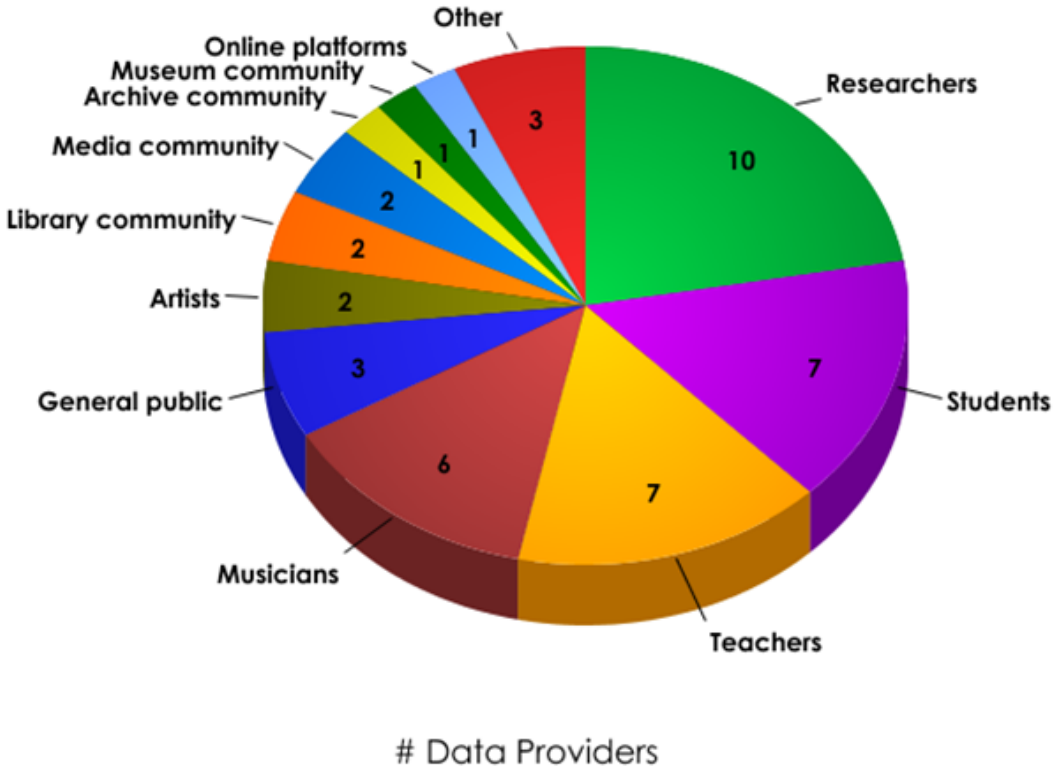


Figure 1: Communities that are currently engaging with Data Providers

5 https://basecamp.com/1936492/projects/5105397/uploads/17288804?enlarge=114233678#attachment_114233678

# 3 END-USER CONTRIBUTIONS

As the DoW states, the enrichment through crowdsourcing strategy is based on the central concept of breaking up activities into micro-tasks (Part B, page 10). This is part of a wider strategy for enrichment that also consists of (semi-)automatic semantic enrichment - or rather assisted (metadata) enrichments - and collaboration with Wikimedia chapters (Part B, page 80).

In relation to the end-user contributions to enrichment, a distinction into three types of contributions can be made: User-generated metadata, subjective contributions and user profile related interactions. For enrichment through crowdsourcing, the first category is deemed most relevant to meeting the goals of the project. However, the potential of the other two categories should be taken into account in further research and development, since they are the by-product of a larger user engagement strategy (within the project as a whole).

As a result of the bi-weekly teleconference between the technical partners involved in Europeana Sounds (the Technical Coordination Group), as organised by the Technical Coordinator of the project Johan Oomen, an alternative - more technologically driven - distinction between different types of enrichments was created:

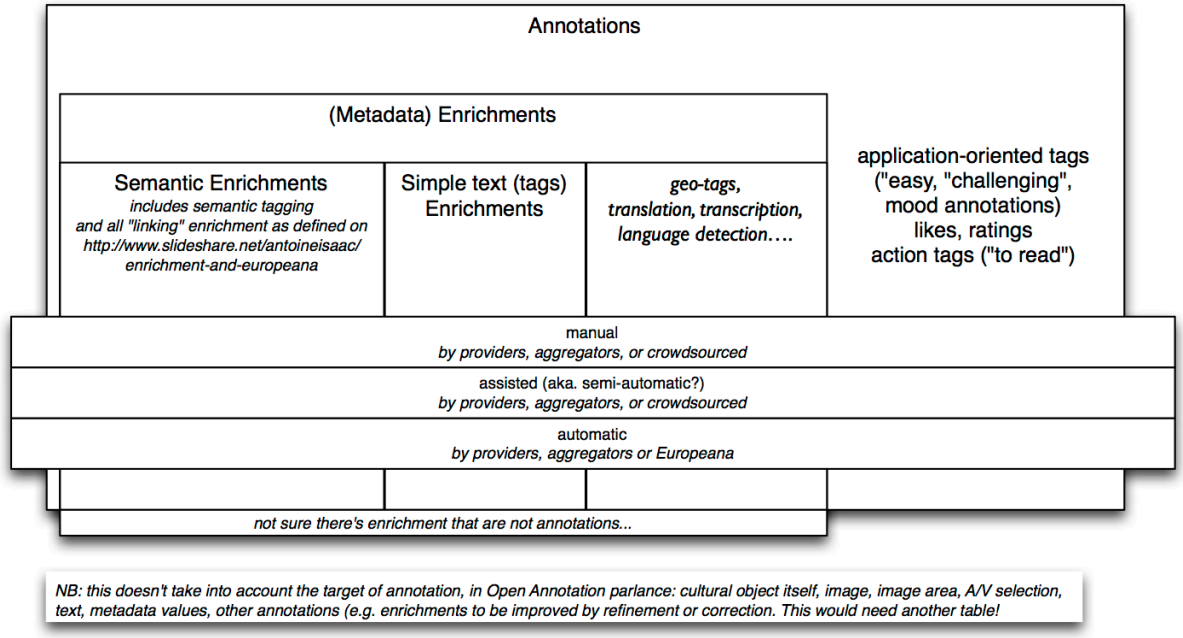


Figure 2: Different types of annotations within the technical infrastructure

The figure above shows how different types of annotations - including, but not restricted to, user contributions - sit within the technical infrastructure of the project. In relation to the types of user contributions described below, Section 3.1 corresponds with (metadata) enrichments, which can be either of a manual or assisted nature. Section 3.2 can correspond with either manual or assisted

(metadata) enrichment, but also just general annotations, which are application oriented. Section 3.3 exclusively corresponds to annotations that are application oriented.

## *3.1 User-generated metadata*

The first category of end-user contributions consists of contributions that enrich the aggregated collections on the metadata level. These correspond to the following types of crowdsourcing from the "Classification of Crowdsourcing Initiatives", as put forward in the article 'Crowdsourcing in the Cultural Heritage Domain' by Johan Oomen and Lora Aroyo (2011, page 140), correction and transcription task, classification & contextualisation (objective):

| Crowdsourcing type | Short definition |
|---|---|
| **Correction and Transcription Tasks** | Inviting users to correct and/or transcribe outputs of digitisation processes. |
| **Contextualisation** | Adding contextual knowledge to objects, e.g. by telling stories or writing articles/wiki pages with contextual data. |
| **Complementing Collection** | Active pursuit of additional objects to be included in a (Web)exhibit or collection. |
| **Classification** | Gathering descriptive metadata related to objects in a collection. Social tagging is a well-known example. |
| **Co-curation** | Using inspiration/expertise of non-professional curators to create (Web)exhibits. |
| **Crowdfunding** | Collective cooperation of people who pool their money and other resources together to support efforts initiated by others. |

Figure 3: Classification of crowdsourcing initiatives

### 3.1.1 Correction and transcription tasks

Transcription and Correction tasks "Invite users to correct and/or transcribe outputs of digitisation processes" (Oomen & Aroyo, 2011, page 140).

An inspiring transcription driven project currently underway is Transcribe Bentham[6]. Transcribe Bentham is a crowdsourcing undertaking coordinated by University College London that strives to transcribe the collected manuscripts of Jeremy Bentham. The project is making use of MediaWiki[7] as a crowdsourcing platform. So far the project has transcribed more than 10,000 of Bentham's manuscripts. The project has more than 16,000 registered end-users. Interesting spin-off results

[6] http://blogs.ucl.ac.uk/transcribe-bentham/
[7] https://www.mediawiki.org/wiki/MediaWiki

have resulted from this process, including the development of a cookbook with recipes[8] Bentham (presumably) wrote for the panopticon prison. St. John Restaurant in London (a Michelin starred establishment) even put one of the recipes on their menu[9].

The University of Maryland made it a course requirement for Introduction to Digital Humanities that students were to report on their experience using Transcribe Bentham[10]. The translation of metadata into other languages can also be seen as a form of transcription of existing metadata. In the Final report the user research has shown a great willingness of almost half of the respondents (44.9%) to translate metadata into their own language (Final report, page 37).

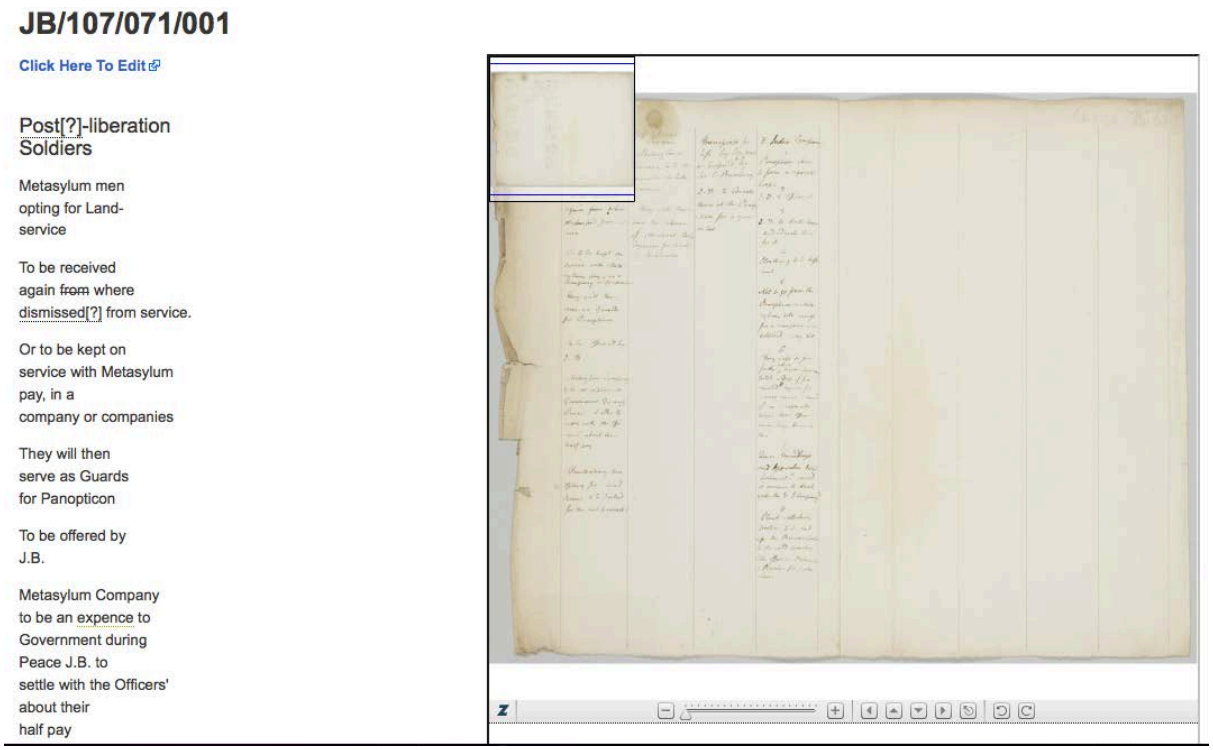


Figure 4: Screenshot of Transcribe Bentham

For the collections of Europeana Sounds, transcription and correction as forms of metadata enrichment can also be envisioned in many use cases. Classical music scores for example, are extremely difficult to transcribe digitally. This is due to numerous factors, including lack of consistency in style and handwriting quality. This makes a crowdsourced transcription endeavour within Europeana Sounds related to classical music scores significant and would no doubt produce academically and culturally significant results.

[8] http://blogs.ucl.ac.uk/transcribe-bentham/2013/03/15/bake-it-like-bentham/
[9] http://www.ucl.ac.uk/news/news-articles/0813/30082013-Bentham-recipe-St-John
[10] http://mith.umd.edu/engl668k/?s=bentham&submit=Search

Correction of metadata is most relevant in the context of the assisted (metadata) enrichments within Europeana Sounds, where validation of the output of automatic enrichment processes - either as a result of data enrichment as part of the aggregation by Data Providers or the 'standard' enrichment that is part of the ingestion by Europeana - is foreseen by the DoW in some cases (Part B, page 76). However, in principle - depending on the use case - correction can be applied to every available metadata field, although this sometimes calls for a verification or moderation process.

Correction can also pertain to the correction of transcriptions. This is common in transcription projects because users have direct access to the transcriptions and original texts. Such a structure would also be suitable within Europeana Sounds. For example, DIY History from the University of Iowa Library[11], even though similar to Transcribe Bentham, allows other users to review each other's transcriptions, almost in a group context. In that sense, end-users can collectively transcribe and correct transcriptions.

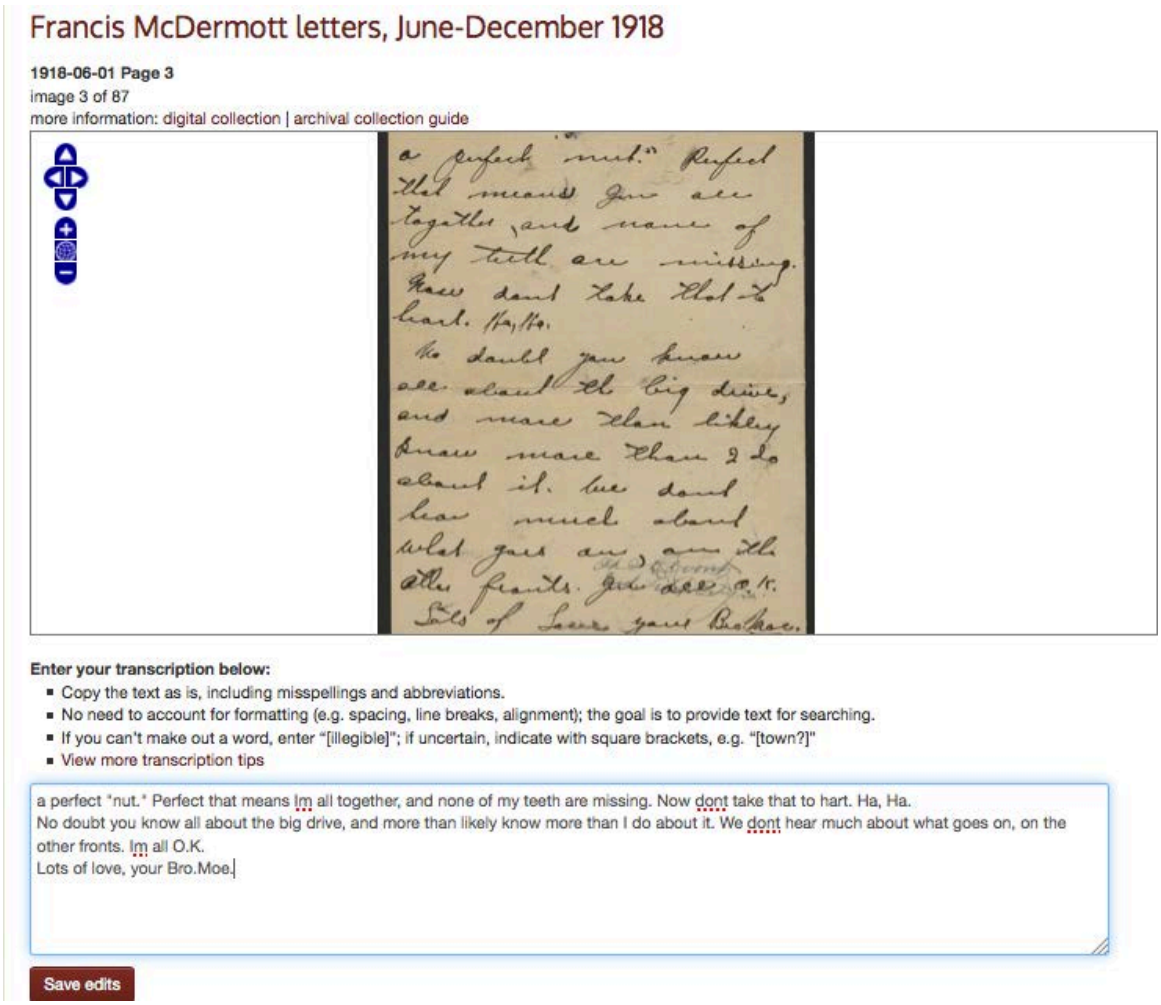


Figure 5: Transcription and Review page from DIY History

[11] http://diyhistory.lib.uiowa.edu/

### 3.1.2 Classification

Classification refers to gathering metadata related to objects in a collection. Social tagging is a well-known example (Oomen & Aroyo, 2011, page 140). Classifying, according to the Open Annotation Data Model, "represents the assignment of a classification type, typically from a controlled vocabulary, to the target resource(s). For example to classify an Image resource as a Portrait." (Van de Sompel et al, 2013)

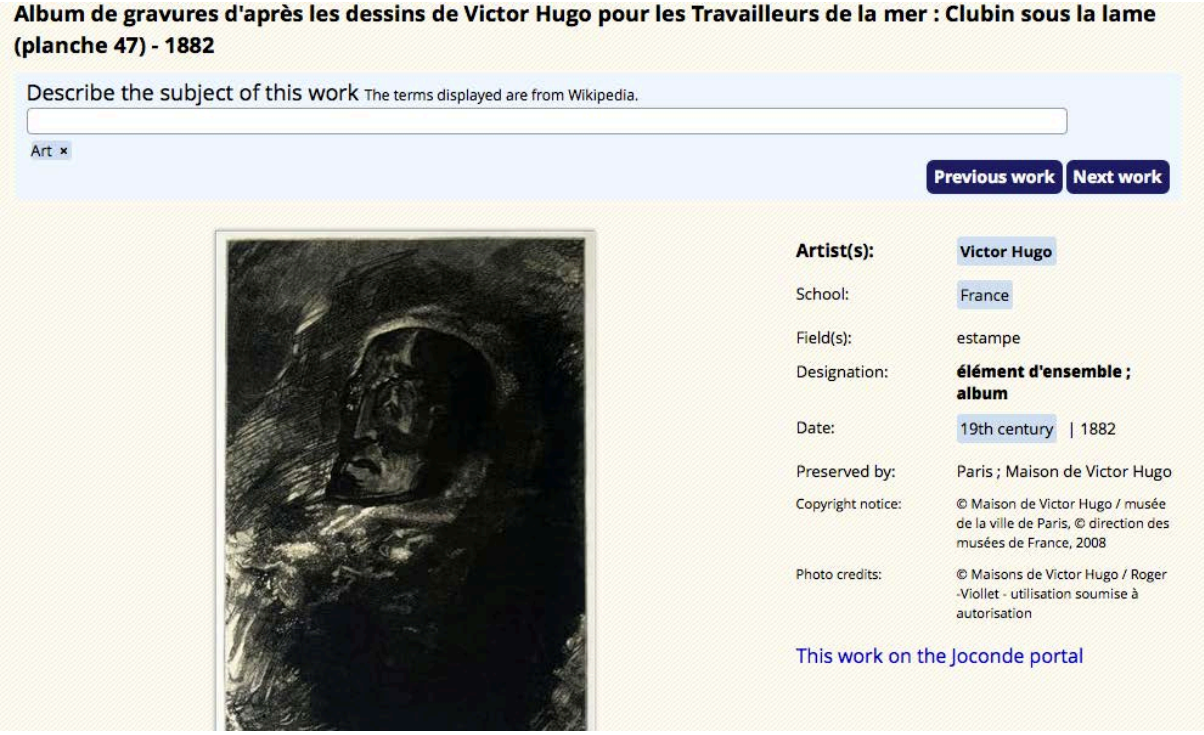


Figure 6: Classifying painting on Joconde as "art"

Classification - and the above mentioned example of social tagging - is probably the most well-known and most often applied type of crowdsourcing. Relevant examples in the context of Europeana Sounds would be adding a genre to a piece of music, or adding the name of a specific species of bird to a sound recording. This could either be facilitated in free form, or based on existing ontologies or other knowledge structures. D1.3 *Ontologies for Sound* [REF 2] of WP1 in Europeana Sounds has completed a new ontology for sounds that can and will be used for this purpose. As part of WP1 Task 1.3 the EDM model has also been extended to accommodate metadata from archival audio material better.

An inspiring classification project is the JocondeLab[12] proposed by the French Ministry for Culture and Communication. The project uses an implementation of assisted enrichment, where users' tags are aligned with DBpedia[13] concepts in real-time and the multi-lingual labels of DBpedia are indexed

[12] http://jocondelab.iri-research.org/jocondelab/
[13] http://wiki.dbpedia.org/

for improved multi-lingual retrieval. The platform is also available in 14 different languages, which allows for a larger and more diverse user reach. However, because DBpedia links to Wikipedia articles, this is also a form of contextualisation because of the deeper level of information gained through this connection.

Annotorious[14], developed by the Austrian Institute of Technology (AIT, also a partner in WP2) is another tool that allows social tagging. It also makes use of DBpedia to support the generation of structured data from crowdsourcing activities. There are several examples of institutions making use of this software. These include ARCHER (Archives of the American Numismatic Society)[15] as well as Europeana 1914-1918[16].

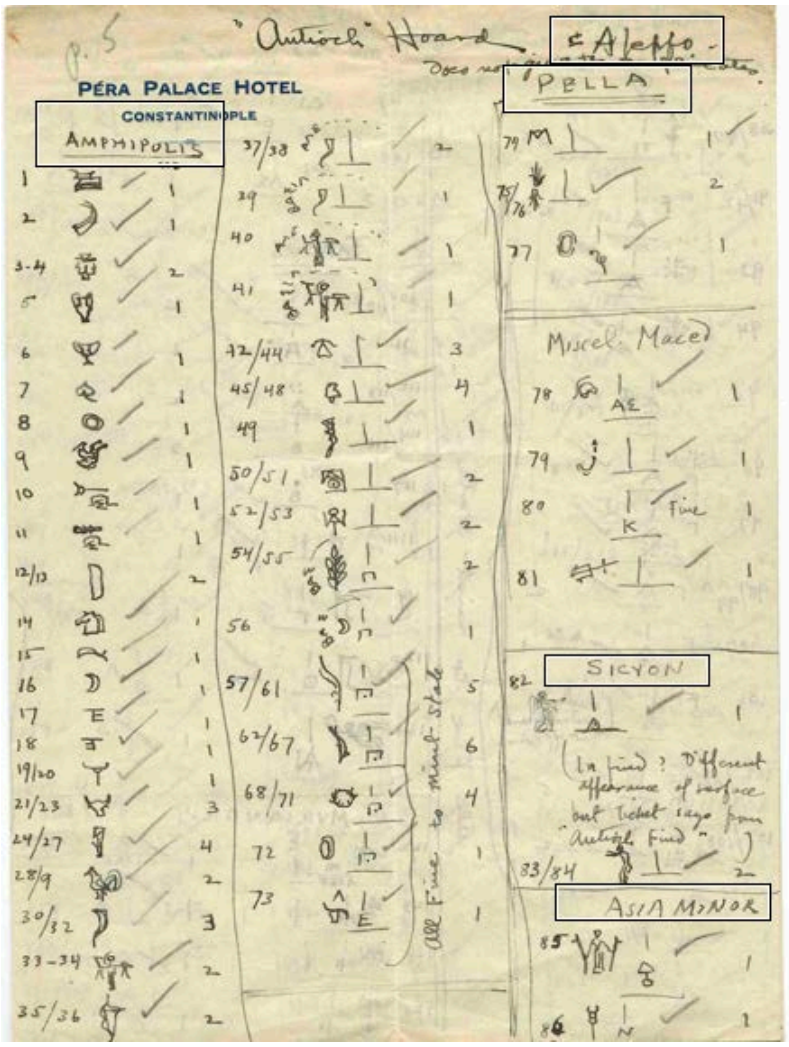


Figure 7: ARCHER project with tagged objects in blue boxes

[14] https://annotorious.github.io/
[15] http://numismatics.org/archives/ark:/53695/nnan187715
[16] http://www.europeana1914-1918.eu/

### 3.1.3 Contextualisation (objective)

Contextualisation refers to the addition of contextual knowledge to objects, written articles/wiki pages with contextual data, or enrichment based on specific knowledge about the object.

For instance, the "Sound Connections" pilot developed as part of Europeana Creative[17] makes use of the HistoryPin interface and encourages communities to enrich sound assets with other records from Europeana, Wikimedia Commons or other similar resources i.e. an audio file of a Robin singing could be contextualised with images of other Robins, videos, ornithological info etc.

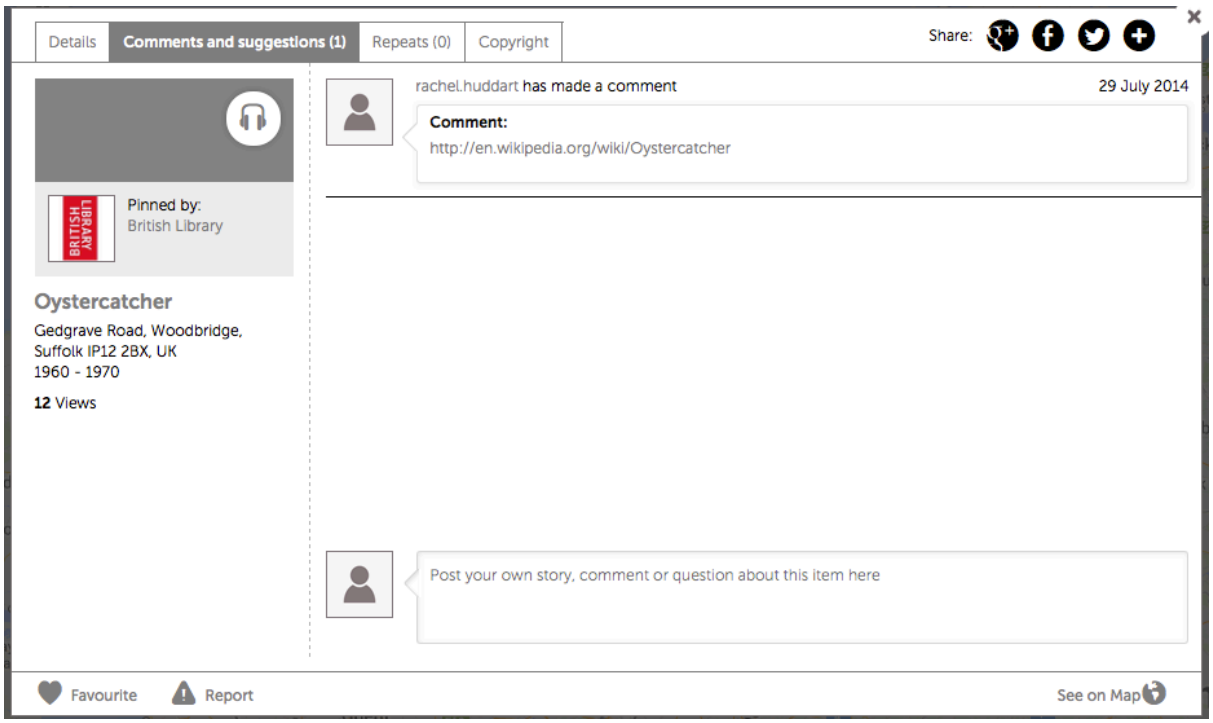


Figure 8: Screenshot of "Sound Connections" with contextualised bird recording

As part of Europeana Sounds HistoryPin will, together with data providers Tobar an Dualchais, Comhaltas Ceoltóirí Éireann and the Irish Traditional Music Archive, help traditional musicians to engage with culture more deeply, by connecting websites popular with traditional musicians to the collections of multiple archives. By linking archives and traditional musicians using digital tools, HistoryPin can reduce barriers to teaching and learning.

HistoryPin will create an online widget or tool that enhances parts of this process. In the case of either a recording or an ABC transcription, the widget might use the algorithm of the TunePal[18] service (invoked via an online Application Programming Interface, or API) in order to develop the best statistical match against a generally accepted corpus of traditional Irish and Scottish music. At

[17] http://pro.europeana.eu/web/europeana-creative
[18] http://tunepal.org/tunepal/index.php

that point, the widget can search the tune index at The Session and also search Europeana's metadata repository, which will include the collection catalogues of the various traditional Archives.

Contextualisation, in relation to user-generated metadata and the Europeana Sounds project, can refer to digital asset contextualisation via users providing information gained from other resources, either within Europeana or other (trusted) web resources. Conversely, it can also be the product of an object being reused in another relevant context, for instance within a Wikipedia article.

**Taxonomy** [edit]

The greater white-fronted goose is divided into five subspecies. The nominate subspecies *A. a. albifrons* breeds in the far north of Europe and Asia, and winters further south and west in Europe.

Three other restricted-range races occur in northern North America: *A. a. gambeli* in interior northwest Canada, and wintering on the coast of the Gulf of Mexico, slightly larger than the nominate form, Pacific white-fronted goose, *A. a. frontalis* and tule goose, *A. a. elgasi*, in southwest Alaska, largest and longest-billed of all, both wintering in California. All these races are similar in plumage, differing only in size.[7]

The very distinct Greenland white-fronted goose, *A. a. flavirostris*, breeding in western Greenland, is much darker overall, with only a very narrow white tip to the tail (broader on the other races), more black barring on its belly, and usually has an orange (not pink) bill. It winters in Ireland and western Scotland.

*Branta albifrons* (Scopoli, 1769)

Greater white-fronted geese, Texel Netherlands (2013)

Birds breeding in the far east of Siberia east to Arctic Canada, wintering in the United States and Japan, have been described as *A. a. frontalis* on the basis of their slightly larger size and a marginally longer bill. Another putative East Asian subspecies *albicans* has also been described. A 2012 study has found that *frontalis* and *albicans* do not merit subspecies status, the former being synonymised with *gambelli* and the latter with the nominate subspecies; this study found that these forms had been named on the wintering grounds from specimens whose breeding grounds were unknown.[8]

Ecological studies in 2002 suggest the Greenland birds should probably be considered a separate species from *A. albifrons*.[9] Of particular interest is its unusually long period of parental care and association, which may last several years and can include grandparenting, possibly uniquely among the Anseriformes.

Figure 9: Video uploaded to Wikimedia Commons by the Netherlands Institute for Sound and Vision (NISV) being used to contextualise Wikipedia Article

Contextualisation can also be the linking of media to a Wikipedia article. This is done regularly within the community of volunteers that maintain Wikpedia, but specific events are also held, which encourage such contextualisation. Such events are called edit-a-thons. Europeana Fashion has already held multiple events of this type and has produced a handbook[19] on how to organise these events. Europeana Sounds will, as written in the DoW, "Collaborate with Wikimedia chapters in Europe to add contextual knowledge on the Europeana Sounds collection". To accommodate this aim, a series of edit-a-thons tailored to sound and music specific topics, related to the collections aggregated by Europeana Sounds, will be developed together with Wikimedia chapters all over Europe.

## *3.2 Subjective contributions*

The second category of end-user contributions refers to the social and personal interaction with the content on the object level. This corresponds to the following types of crowdsourcing from the Classification of Crowdsourcing Initiatives (Oomen & Aroyo, 2011, page 140), contextualisation (subjective) & co-curation (for public display).

[19] http://blog.europeanafashion.eu/download/Fashion%20Edit-a-thon

### 3.2.1 Contextualisation (subjective)

Contextualisation refers to adding contextual knowledge to objects, e.g. by telling stories or writing articles/wiki pages with contextual data (Oomen & Aroyo, 2011, page 140).

For this type of end-user contributions contextualisation can refer to interactions with the object such as adding a (personal) comment or story to an object, or participating in a discussion or asking a question about it, for example in the form of a threaded conversation. This is part of the "social interaction" that appears in the suggested "Communicate" strand in the Final report (page 38).

For example, The Great War Archive Flickr Group[20] was a particularly active environment for subjective and social contextualisation. The Flickr page was filled with WWI photos gathered during Oxford University's RunCoCo crowd sourcing campaign[21]. On Flickr users could easily add comments and discuss the photographs. For example this photo[22] led to a discussion about who the unnamed "Arbeiter" in the photo was. Another user uploaded an additional photo[23] from the same event but taken from a different angle.

Another example of this contextualisation is Trove[24]. Trove is a platform developed by the National Library of Australia that allows users to have access to a large, rich collection of images, maps, music, video books and more. Users are able to add comments providing contextualisation. In a recent month (11/8/2014-11/9/2014) 1,857 new comments were added to items. Users are also able to do other, less subjective things such as adding tags as well as splitting and merging items. Creative Australia praises Trove, writing "Trove users have become a community, too. Its members interact with each other in many ways. They provide each other with hints and tips about using Trove, and comment on each other's work. They organise themselves into communities of interest completely independent of the National Library."[25]

[20] https://www.flickr.com/groups/greatwararchive/pool/
[21] http://projects.oucs.ox.ac.uk/runcoco/
[22] https://www.flickr.com/photos/48140075@N04/14881253299/in/pool-greatwararchive
[23] http://search.socialhistory.org/Record/845678
[24] http://trove.nla.gov.au/
[25] http://creativeaustralia.arts.gov.au/module/creative-australia-pathways/theme-connecting-to-national-life-for-a-social-and-economic-dividend/pathway-access-interpretation-and-innovation-of-national-collections/case-study-powered-by-trove-new-ways-for-australians-to-engage-with-culture-and-history/

| When | By | About | Comment |
|---|---|---|---|
| 2014-08-26 16:10 | expertopinion | Mark Jeffrey (Version) | This photograph of Mark Jeffrey was taken at the Hobart Gaol, Campbell St Hobart, by police photographer Thomas J. Nevin in the week of 17th April 1877, when Jeffrey was transferred from the Port Arthur prison to Hobart. Rating: 4 stars |
| 2014-08-26 13:59 | lynell.charlton | Photographic record, description and criminal history of Walter Joseph Cashman, 27 May 1920 (Version) | CASHMAN Walter Joseph aka Walter John (1870-1924) son of George Cashman (1825-1890) and Catherine Cashman nee Griffin (c1833-1919). |
| 2014-08-25 12:55 | Anonymous | For the love of Kate / Annette d Chand ; [editors: Peter W. Whitewood, Deb Gard] (Version) | This book can be purchased from the Author. website: www.thewillowaye.webs.com or phone 0415426225 for your copy . |
| 2014-08-22 20:59 | aleago2 | The persimmon tree and other stories / Marjorie Barnard | A woman who was ill and , maybe, secluded by the world, starts to see the things around her again, with a different "eye": the eye of the soul. She's like a child in front of nature and she reflects in herself about the seasons and their effects on her. The author chooses a very vivid and colourful tree: the persimmon. The lady in the story, who remains without a name or an age is looking at a woman living near her house. She's probably the same age and she also remains without a name or an age. It's unusual for the season to s... |
| 2014-08-22 13:41 | Anonymous | Architectural model, Sydney Opera House, Jorn Utzon's design for the major hall, wood / plastic, Fin.. (Version) | This was a fantastic project. I visited the site in Crows Nest often and supplied some of the acrylic needed for the model. Laurie Green |
| 2014-08-21 14:50 | Anonymous | Captain George White (Version) | George White was my great, great grandfather. His daughter Jessie was my great grandmother and her daughter, Vivian Irene, was my grandmother. I want to write a book about them, so any comments would be gratefully received. jandjtrefle@gmail.com |

Figure 10: Example of subjective contextualisation on Trove

### 3.2.2 Co-curation (for public display)

Co-curation refers to using the inspiration/expertise of non-professional curators to create (Web) exhibits (Oomen & Aroyo, 2011, page 140).

Co-curation is the rearrangement and compilation of objects, based on (informed or random) selection by a user. This software can be paired with individuals' collections and/or the Europeana API or open datasets such as the ones available on Europeana Labs. This gives individuals, smaller institutions or community organisations the capability to creatively and professionally present their stories and heritage.

While the development of fully fledged, new, co-curation applications is outside the technical scope of Europeana Sounds, there are plenty of examples of what is possible. Timeline.JS[26] allows users to make extremely user-friendly, visually-striking timelines. With StoryMap.JS[27] end-users can create their own interactive maps that tell a story, by using pinned data which can be elaborated upon easily.

[26] http://timeline.knightlab.com/
[27] http://storymap.knightlab.com/

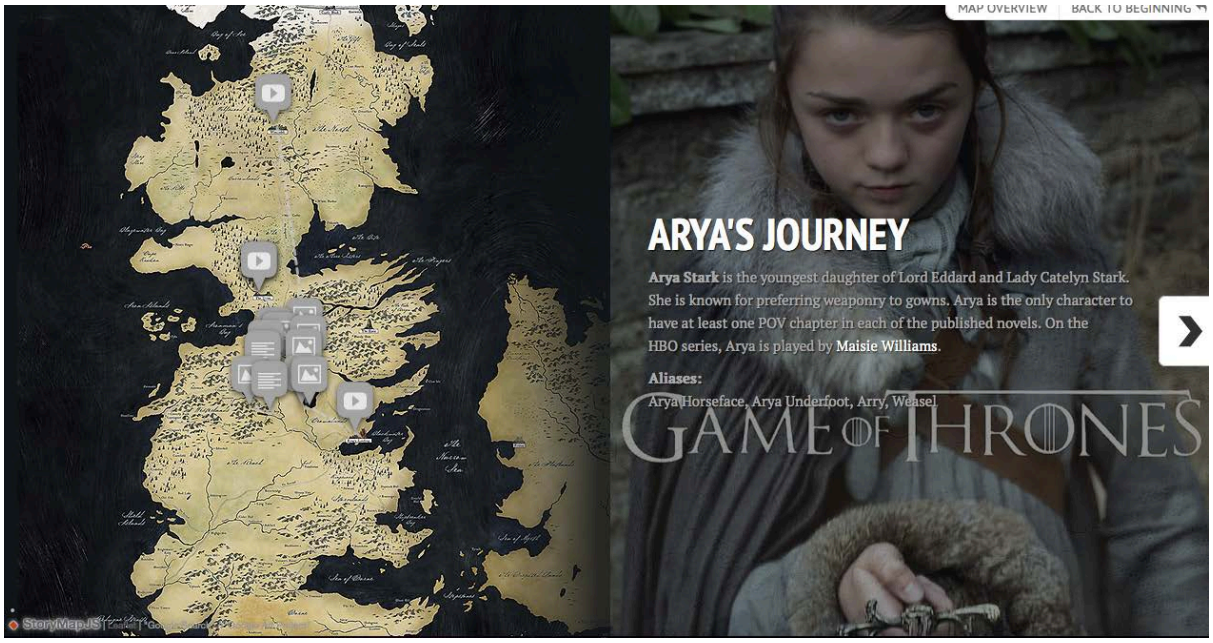


Figure 11: Storymap.JS platform being used by Game of Thrones

In the context of Europeana Sounds, one could also imagine creating a playlist that is shared online or a virtual exhibition documenting the life and travels of Mozart and the music he composed or performed in different cities.

## *3.3 Profile related interactions*

The last category of end-user contributions includes the - mostly implicit - relations and user interactions with the content that stem from the functionality offered by Europeana through the "My Europeana" user profile. This functionality may be extended as part of the work of Work Package 4 and may offer additional interactions with the content that can be used to extract information. Currently it allows users to save their favourite items, searches, tags and to translate search keywords into other languages. There is some overlap with the above-mentioned subjective contributions, although privacy considerations may be higher for this category.

Most user-generated contribution platforms require some form of registration. Predominantly this is done via a social media account (Facebook, Twitter, Google +). This is the case for example, HistoryPin[28]. Other projects are more formal and internal such as Wikipedia, which to some extent is its own social community of users. This is one additional aspect of profile-related interactions - the sense of community. Soundcloud[29] also displays this characteristic, although the time-stamped annotations added to most audio are entirely subjective and personal.

[28] http://www.historypin.com/user/
[29] http://www.soundcloud.com/

Waisda[30] is another profile-related crowdsourcing annotation program that uses gaming as a method to annotate television heritage. It actively seeks collaboration with communities connected to the content, and uses curated vocabularies as a means to integrate tags with professional annotations. Waisda awards different prizes to those participating.

Transcribe Bentham as mentioned previously also requires users to create a profile. It provides a forum for discussion, in addition to awarding transcribers with badges and praise, depending on the amount of transcriptions they complete.

Since these platforms already exist, it is within the scope of Europeana Sounds to promote and ensure wider usage of the Europeana API in profile-related platforms such as the ones mentioned above.

[30] http://www.waisda.nl/

# 4 SUMMARY

This Milestone provides the definition of the different types of end-users that are addressed and the different types of user contributions that the project expects to support.

To establish maximum impact a differentiation between “experts” and “non-experts” (general public) for the crowdsourcing micro-tasks is proposed, in order to be able to distinguish between tasks that range from simple tasks for a broad audience, to tasks that require certain domain knowledge and target a more specific audience. These two categories have been further developed into culture snackers and culture vultures.

To summarise, culture snackers are perceived as users with a casual interest in Europeana or the type of content it serves. Culture vultures are perceived as users who work professionally with the Europeana portal, or the type of content it serves. In the context of Europeana Sounds special attention will go out to the users - of both kinds - with a specific interest in audio and/or music. It’s important to realise that culture snackers and culture vultures are not static or mutually exclusive roles. Depending on context a single user is sometimes a snacker and sometimes a vulture.

In relation to the end-user contributions to enrichment a distinction into three types of contributions can be made: User-generated metadata, subjective contributions and user profile related interactions. For enrichment through crowdsourcing the first category is deemed most relevant to meeting the goals of the project, however the potential of the other two categories should be taken into account in further research and development, since they are the by-product of a larger user engagement strategy (within the project as a whole).

The first category of end-user contributions consists of contributions that enrich the aggregated collections on the metadata level. The second category of end-user contributions refers to the social and personal interaction with the content on the object level. The last category of end-user contributions includes the - mostly implicit - relations and user interactions with the content that stem from the functionality offered by Europeana to the end-user through the “My Europeana” user profile.

# 5 REFERENCES


| | |
|---|---|
| **Ref 1** | EC-GA including Annexe I ("Description of Work")<br>http://pro.europeana.eu/documents/2011409/8d0e9833-4608-494e-af77-681e68f8a8c8 |
| **Ref 2** | D1.3 Ontologies for Sound<br>http://pro.europeana.eu/documents/2011409/ed7f27af-65ff-4ea3-9d00-8345c31489cf |
| **Ref 3** | Design for User Engagement on Europeana Channels (C. Shen, 2014)<br>https://basecamp.com/1936492/projects/4984678/attachments/100527076/download |
| **Ref 4** | Crowd vs. Experts: Nichesourcing for Knowledge Intensive Tasks in Cultural Heritage (J. Oosterman et al, 2014)<br>http://dl.acm.org/citation.cfm?doid=2567948.2576960 |
| **Ref 5** | Answers Data Providers Survey on Enrichments through Crowdsourcing<br>https://basecamp.com/1936492/projects/5105397/uploads/17288804?enlarge=114233678#attachment_114233678 |
| **Ref 6** | Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges (J. Oomen and L. Aroyo, 2011)<br>http://www.iisi.de/fileadmin/IISI/upload/2011/p138_oomen.pdf |
| **Ref 7** | Open Annotation Data Model (H. Van de Sompel et al, 2013)<br>http://www.w3.org/ns/oa#d4e731 |