



Project Acronym: EAwareness
Grant Agreement number: 297261
Project Title: Europeana Awareness

D2.7 Report on Infrastructure and Tools for supporting User Contributed Content

Revision	Final
Date of submission	27 June 2014
Author(s)	Gregory Markus (NISV), Johan Oomen (NISV), David Haskiya (Europeana), Vassilis Tsouvaras (NTUA)
Dissemination Level	Public

Project co-funded by the European Commission within the ICT Policy Support Programme

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision No.	Date	Author	Organisation	Description
0.1	June 13 2014	Johan Oomen	NISV	Outline
0.2	June 16 2014	Gregory Markus	NISV	Chapter 2 & 3 additions
0.3	June 16 2014	David Haskiya	Europeana	Chapter 1 additions
0.4	June 17 2014	Vassilis Tsouvaras	NTUA	Chapter 1 additions
0.5	June 17 2014	Gregory Markus	NISV	Chapters 2&3
0.6	June 18 2014	Gregory Markus	NISV	Minor changes
0.7	June 25 2014	Gregory Markus	NISV	Incorporation of changes suggested by Wiebe de Jager (Europeana) and Natasa Sofou (NTUA)
0.8	June 26 2014	Gregory Markus	NISV	Finalisation and structuring

Table of Contents

Executive Summary	4
Introduction	5
Chapter 1: User Generated Content within the Europeana Platform	7
1.1 Overview of architecture supporting UGC	7
1.1.1 User Generated Content and Metadata and the Europeana software architecture	7
1.2 Overview of the Platforms Used	10
Chapter 2: UGC and Crowdsourcing within the Europeana Network.....	14
2.1 Categorising Crowdsourcing User Generated Content Software	14
2.1.1 Correction and Transcription Tasks	15
2.1.2 Contextualisation	16
2.1.3 Classification	17
2.1.4 Complementing Collections	20
2.1.5 Co-Curation	21
2.2 UGC Taskforce (More info in Annex 2)	23
Chapter 3. Measuring Impact of UGC in Europeana	24
3.1 Europeana Creative.....	24
3.2 Europeana 1914-1918.....	29
3.4 GLAMwiki Toolset.....	31
Conclusion and Future work.....	33
Annex 1: Tools for User Generated Content	34
Annex 2: Strategies for user generated content and crowdsourcing in museums and cultural heritage.....	48
A tale of three terms: User Generated Content, Crowd-Sourcing and Engagement: from prescription to process	49
Museums Visitors' Generated Content.....	51
Results from the Europeana Taskforce	52
User contributions in the Europeana 1914-18 and Europeana 1989 project.....	53
Conclusion	54
Annex 3: Wiki Loves Public Art 2013.....	55

Executive Summary

During its beginning years Europeana successfully aggregated millions of metadata records for digital cultural heritage assets across Europe. The extensive Europeana Network continued to grow and the technical infrastructure of the portal was increasingly bettered. The Europeana Awareness project sought to engage and Cultivate new ways for users to participate in their cultural heritage by: enhancing the users' experience; extending Europeana's social media programmes and use of web 2.0 tools and brokering a new relationship between curators, content and users.

This deliverable provides an update about the infrastructure and tools for supporting user contributed content in Europeana. This deliverable is a follow up to D2.1, delivered in 30 March 2012. The deliverable covers how Europeana is garnering user generated metadata through various ways. This is achieved multiple ways including the development of software and platforms that catalyze facilitation. Additionally, the report covers user generated content and crowdsourcing endeavours being undertaken within the Europeana Network and which open source tools are available to help further this task. This deliverable showcases Europeana's efforts to succeed in achieving the plan stated above. Also it addresses comments made by the second year review, specifically on the position of Wiki Loves Art (Annex 3).

Introduction

Crowdsourcing user generated content (UGC) allows cultural heritage to become a more participatory experience; no longer leaving curation solely to the curators. Making digitised content openly available online for crowdsourced enrichment allows for peer reviewing communities like Wikipedia offers and allows for a larger network from which ulterior information can be gained. Furthermore, crowdsourcing UGC makes cultural heritage more inclusive, allowing individuals to include their history to the larger collections.

This necessity for such inclusion can be discerned after reading this passage from scholars Jordan and Weedon, who write of cultural policy:

“Whose culture shall be the official one and whose shall be subordinated? What culture shall be regarded as worthy of display and which shall be hidden? Whose history shall be remembered and whose forgotten? What images of social life shall be projected and which shall be marginalized? What voices shall be heard and which silenced? Who is representing whom and on what basis?”¹

By utilising crowdsourcing and welcoming UGC, memory institutions like museums and archives can avoid having to ask the questions that Jordan and Weedon pose. Europeana works towards ensuring that this inclusivity is continually improved and made more visible to promote awareness. However, for UGC and crowdsourcing to be possible, open source software designed for such tasks must be available. Aware of this, Europeana has taken numerous actions to increase the findability, usability, and awareness of UGC software. These tools can be utilised by cultural heritage institutions, community organisation; any body looking to manage digital assets in an inclusive way.

Context of this deliverable

This deliverable is a follow up to D2.1, delivered in 30 March 2012 . It outlines work done within Task 2.1 *Operationalise tools to enable end user contributions to Europeana content.*

Specifically, it provides information and context divided in two chapters:

¹ Jordan, G., Weedon, C., 1995, *Cultural Politics: Class, gender, race and the postmodern world*. Blackwell, Oxford. p. 5

Chapter 1. User Generated Content within the Europeana platform

- overview of the UGC architecture
- overview of the platforms used
- position of the Digital Storytelling platform

Chapter 2. Describe UGC activity in the context of the Europeana Network

- Tools available through EuropeanaTech
- Work of the UGC taskforce

Annex 1. contains an overview of UGC tools available through the FLOSS inventory

Annex 2. contains the work of the UGC Task Force

Annex 3. addresses comments made by the second year review on the position of Wiki Loves Art.

Please note here that D2.5 Updated Report on the content gathering campaigns (delivered in January 2014) provides (partly quantified) information on the execution of the UGC campaigns 1914-1918, 1989 and the Wiki Loves Public Art project. It is relevant to see the impact of the tools through its practical usage.

Chapter 1: User Generated Content within the Europeana Platform

The Europeana Awareness proposal was written in 2011. In the past years, the strategy of Europeana has shifted, with more emphasis placed on Europeana acting as platform enabling various stakeholders to build tools and applications on.

1.1 Overview of architecture supporting UGC

Europeana Awareness has built technology that will play an important role in this transformation. The Digital Storytelling Platform is a major component of the 'future Europeana'. Its place will be different from the plans defined in 2011. This paragraph outlines how the Europeana platform will be enabled to manage user generated metadata and content. It also provides insights into the practical development efforts taking place in the final year of the Europeana Awareness project.

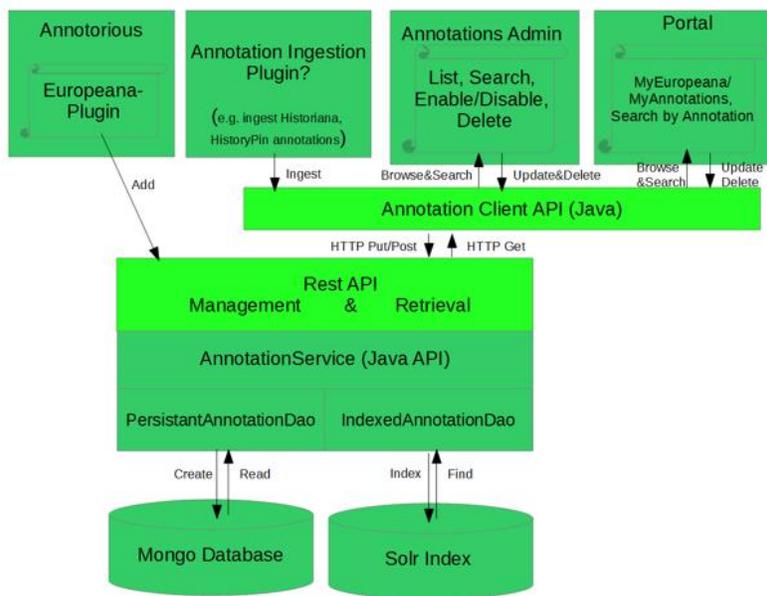
1.1.1 User Generated Content and Metadata and the Europeana software architecture

The Europeana architecture is designed to the storage of annotations, indexing, user tags, Image annotations, digital storytelling platform and user created sets. These components are built within the context of Europeana Awareness. If developed in related projects (notably EuropeanaCreative), this is mentioned explicitly in the subsection that follow below.

1.1.1.1 Annotations storage, indexing and API component

To extend the Europeana platform Europeana and NTUA are currently developing a component that is able to store and index (for search and retrieval) user annotations. As a true platform component it will have a public API allowing non-Europeana clients to annotate objects and web resources in Europeana.

The nascent [architecture](#) can be extended to support almost any kind of user annotations.



Within the scope and duration of Europeana Awareness (until the finalisation of the project early 2015) the focus is on tags, image annotations and sets as these are useful across multiple use-cases and domains/themes.

The user created metadata platform is developed in collaboration with the EuropeanaCreative project. The annotation backend service including the storage service, the rest api and annotorius plugin are under development in EuropeanaCreative. The portal extension, including the extension in MyEuropeana, MyAnnotation and the annotations admin services are under development in the EAwareness project. In particular, NTUA is responsible for the development of the user annotation extension in MyEuropeana using and extending the backend of the Digital Story player. Europeana is responsible for the frontend design including the creation of functional and design wireframes and their implementation.

Other types of annotations are likely to be developed during 2015-16. This will also be done in the scope of projects such as EuropeanaSpace. The backend service of the digital story player will also be used and extended in the framework the Europeana Space project. The backend service currently provides search and retrieval of metadata from Europeana. It is likely to extend its functionality to search and retrieve metadata from Digital Public Library of America, Youtube, Freebase and other external knowledge sources.

Note that Europeana will not necessarily implement the features in our own portal. Rather we will implement them in our thematic sites e.g. Europeana 1914-1918 which supports tagging and image annotation. Within Europeana Sounds we are currently working on refining, designing and then implementing what we call Channels. The Channels concept is intended to allow us to, together with partners in the Network, to in a scalable fashion develop more thematic sites like Europeana 1914-1918 and Europeana 1989 where user contributed content and user annotations play a larger role.

1.1.1.2 User tags

This feature allows users to add tags to Europeana objects. The tags will be aligned with semantic resources (so called “semantic tags”) and the resources’ multi-lingual labels indexed in order to improve discoverability across languages.

1.1.1.3 Image annotations

This feature allows user to add comments to specific areas of an image. Semantic tags will be extracted from the comments in order to improve discoverability across languages.

1.1.1.4 The Digital Storytelling Platform

After evaluating the (Digital Storytelling Platform) DSP-prototype and re-evaluating it in the light of Europeana’s platform strategy we have decided to refine the prototype and specialize it into a service (backend, API and client) that allows users, in Europeana or in a 3rd party application, to create and publish sets of content. The state of affairs after the second project year has been listed in D6.3: Report on Year Two of Europeana Awareness:

The DSP was evaluated by a representative group of potential users. The test results showed the strength of the back-end (server-side) platform as a way to store and connect user-uploaded stories, media and contextual connective narrative and the need to improve the front-end interface for better usability. As envisioned in the Description of Work, the DSP will be amended following this feedback, most notably:

- The DSP back-end is improved and will be part of Europeana Labs (once finished) and the Europeana GitHub's contrib branch as a prototype for others to use in their own campaigns. The project partners (NTUA, Europeana) are investigating using the platform

within Europeana Creative to deploy parts of the back-end for the development of user generated galleries as part of a task on user annotations.

- The DSP front-end will need to be improved in order to be more user-friendly. This work is currently being scoped and the public release is planned for early 2014, in line with the Europeana Communication strategy of 2014. The DSP will be featured on Labs in September 2014². This marks the conclusion of a milestone MS11 Final versions of the integrated toolset.

1.1.1.5 User created sets

User created sets are small collections of existing objects “curated” by Europeana users and which can be given a name/title, a description, and tags by the user. They can take the form of e.g. book lists (“Charles Dickens books translated into French”), playlists of music (“Traditional Irish Wedding Music”), mini art galleries (“My Rococo Gallery”) or small themed collections across media types (“Mozart’s music, sheet music and books about him”) - the possibilities are many and limited by the imagination of our users and how user-friendly we will be able to make the feature. Two benchmarks in this area are [DigitalNZ and their sets](#) and [Rijksmuseum and their Rijksstudio feature](#).

Published sets will be searchable via the Europeana API and via the Europeana Portal and sets can also be created via the API. This means that services and sites using the Europeana API will also be enable to create and re-publish sets. This development will take place the coming period, and will end be finished in September 2014 (*see Section 3.2 below*).

1.2 Overview of the Platforms Used

1.2.1 RunCoCo and Europeana 1914-1918

The RunCoCo (**Run Community Collections**) framework is extended and used to support the Europeana 1914-1918 campaigns. This development is done by the Europeana office as part of Europeana Awareness and in collaboration with the other WWI-themed Europeana projects.

A major piece of development already accomplished here is that RunCoCo "out of the box" supports data exchange with the core Europeana repository via technical interfaces (the Search

² <http://labs.europeana.eu/>

API for retrieval, OAI-PMH for harvesting UGC). The addition of these data exchange features together with the open source nature of RunCoCo opens up for any organisation to set up a community collection project and when doing so be able to both access existing Europeana data and provide their own community created content to Europeana. As such this development contributes to Europeana as a platform.

In terms of individual features to mention in the context of user created content is that we've recently added the ability for users of Europeana 1914-1918 to import photographs from their Flickr accounts.

1.2.2 Historypin and Europeana 1989

The Europeana 1989 site on Historypin was developed and provided the first simple semi-manual functions which facilitated the exchange of story objects and user annotations data between Europeana and Historypin. This data exchange will be further refined, automatised, in Europeana Creative and in Europeana v3.0 and will be documented as a first case of "roundtripping" data, data that happens to be user created stories and annotations, between a non-Europeana platform and the Europeana platform.

The process, code and data exchange formats will be documented and published on Europeana Labs and Europeana Pro so that other platforms than Historypin can emulate it in setting up a flow of user created content and annotations to and from Europeana. As such this development contributes to Europeana as a platform.

1.2.3 Wikimedia Commons

During 2014 Europeana experimented with exporting metadata from Wikimedia Commons and importing it to Europeana. We will primarily work with importing the resulting media from the Europeana sponsored competitions Europeana Loves Art Nouveau, Europeana Loves Public Art and Europeana loves GLAMs which were all part of the larger Wiki Loves Monuments competition.

If successful the solution can be extended to import any Wikimedia Commons metadata to Europeana.

1.2.4 Pinterest

In May 2012 Europeana began using the social media platform Pinterest as a new way to engage the public and see how they interact with cultural heritage content. The aim was to connect Europeana partner institutions with end-users via emerging social media platforms and vice-versa; connect end-users with trustworthy and credible sources of cultural heritage content.

Europeana ran a campaign from May 2012-June 2012 to raise the awareness of Pinterest for partner institutions. The results were significant, showing a higher visit rate and overall improvement of Europeana web traffic. From this study Europeana concretely suggested that memory institutions should begin working more with social media, especially Pinterest. Pinterest’s design allows most of the hard work, sharing, curating, interacting and engagement to be done by the users themselves. This is done so by incorporating a “pin it” feature on said institutions’ portals which enables the end-users thusly.

Two years after the campaign the Europeana Pinterest statistics from the last 5 months (Jan 2014-May 2014) remain positive.

Pins	Pinners	Repins	Repinners	Impressions	Reach	Clicks	Visitors
1094	358	6336	5539	1129992	467633	3855	1181

The “repin” feature is probably the most important feature when it comes to crowdsourcing because those who “repin” are doing the sharing and awareness work for Europeana on Pinterest. When an item is repinned it appears on the repinner’s Pinterest feed. So even though 358 originally pinned content, 5,539 users shared the pinned item. This allowed these pinned items to appear in 1,129,992 different search results, main feeds and board feeds and reach an estimated 467,633 different Pinterest users.

1.2.5 89 Voices

89 Voices³ is an oral history and social engagement project envisioned by Neil Bates and Michelle van Duijn of the Europeana Foundation. Built on the back of Europeana’s 1989 collection days in Poland, Hungary, Czech Republic, Germany and the Baltic States, 89 voices will connect with European history through the participants who attend these events via their objects and unique stories.

³ <http://89voices.eu/about>

The project uses Tumblr and Soundcloud, and will continue throughout 2014 once 89 voices have been recorded to mark the 25th anniversary of the start of the revolutions of Eastern Europe.

1.2.6 Other applications for User Contributed Content

Europeana has multiple partners who themselves have applications for user contributed content. We harvest such collections just like any other source and ensure the individual objects are marked as user contributed. One such application is Culture Shock ([example contribution](#)), another is the UK Sound Map ([example contribution](#)).

Chapter 2: UGC and Crowdsourcing within the Europeana Network

One of Europeana's strongest aspects is its network. The Europeana Network contains 2,500 cultural heritage institutions, archives, museums, galleries, universities, technical developers and creative enterprises. Focuses on the technical developer aspect of UGC and Crowdsourcing, Europeana has made use of the network to uncover and aggregate some of the best open source software, develop its own, and create a new platform where developers can easily browse, discover and share cultural heritage related software.

The following section will provide examples of tools which fit the first 5 categories that are available in the FLOSS Inventory⁴, Europeana Pro and Europeana Labs. For more examples refer to Annex 1.

2.1 Categorising Crowdsourcing User Generated Content Software

Europeana has made an initiative to develop and aggregate tools that enable crowdsourcing UGC. These initiatives can be found on Europeana Labs⁵ (beta site for developers) Europeana Pro⁶ (the technical and professional Europeana portal) and the FLOSS (Free, Libre, Open Source Software) Inventory⁷, an aggregated list of said tools developed as part of Europeana v 2.0.

Europeana Awareness deliverable 2.1 "D2.1 User requirements and IPR implications for User Contributed Content in Europeana" presented a categorisation of cultural heritage software that allows for UGC and crowdsourcing. These tools fall into 5 different categories as can be seen in the table below.

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

⁴ At the moment the FLOSS Inventory is being prepped for a migration to Europeana Labs to provide users with a better browsing experience.

⁵ <http://labs.europeana.eu/>

⁶ <http://pro.europeana.eu/>

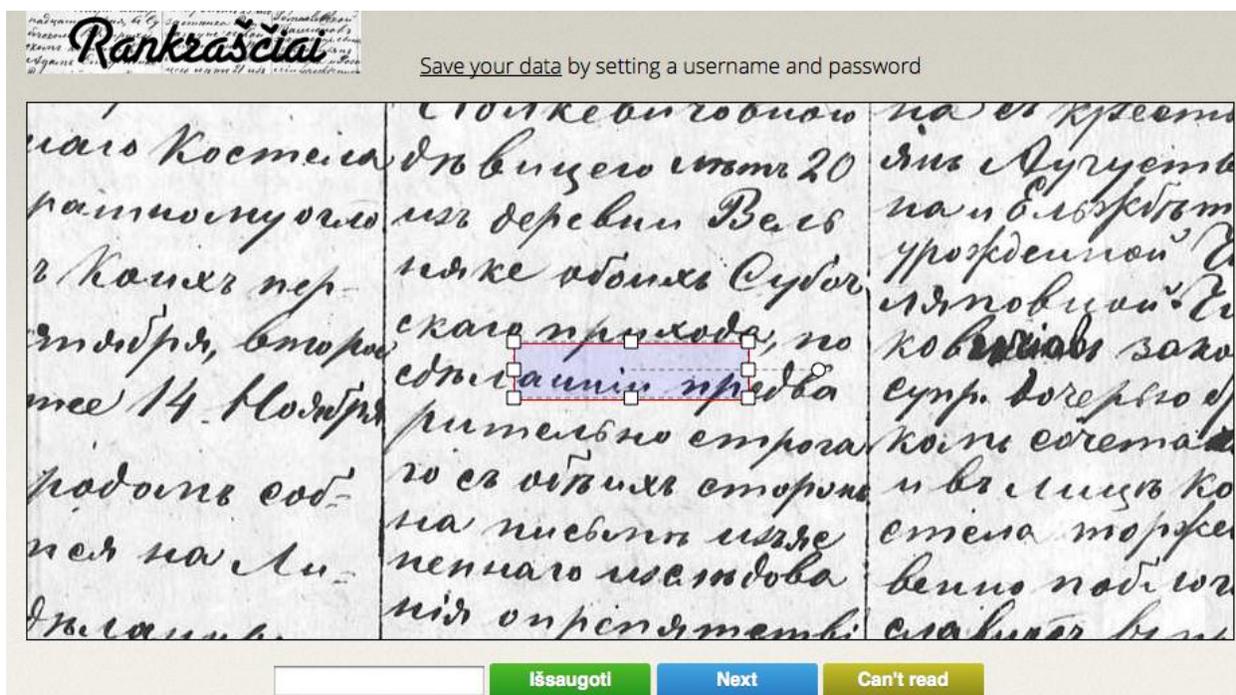
⁷ <http://tinyurl.com/bn6uty>

	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.

Table 1- categorisation of UGC/crowdsourcing methods

2.1.1 Correction and Transcription Tasks

The first example of a correction and transcription task tool available is Crowdhwr⁸. Crowdhwr won the 2012 Hack4LT⁹ hackathon, “Digital Humanities” Category. It can be found on Europeana Labs. The prototype allows users to mark words in a digitized manuscript taken from Europeana and prepare it for analysis. It provides a tool to convert image to text in order to perform automatic search of a manuscript’s content.



⁸ <http://labs.europeana.eu/apps/crowdhwr/>

⁹ <http://pro.europeana.eu/hack4lt>

Figure: Crowdhwr

The software is available under a GNU GPL license which means anyone can make use of the software, adapt it for their own needs and begin transcribing handwritten texts. Due to the readability problems that these old manuscripts present, such a tool is valuable for academics and researchers.

Another valuable tool findable on the FLOSS Inventory for transcription is the Tesseract OCR. Tesseract OCR is an Optical Character Recognition software. The product was developed in the 1980s by Hewlett Packard until being taken over by Google in 2006. The tool can be used to recognize text and process it digitally. This means that books no longer need to be transcribed by hand. Instead they can be scanned and the images can be processed by Tesseract OCR which will provide plain text version of the work. Tesseract OCR is capable of recognizing 40 different languages. The software is all open source and therefore can be used by any person or institution wishing to transcribe texts quickly. The benefit then of this software is that as more institutions digitise books or texts, the public is capable of transcribing them.

2.1.2 Contextualisation

Contextualisation refers to Adding contextual knowledge to objects, e.g. by telling stories or writing articles/wiki pages with contextual data.

FromThePage¹⁰, available on the FLOSS Inventory, is another tool that can be used for transcription, but it has multiple features including contextualisation. “FromThePage... allows volunteers to transcribe handwritten documents online. It's easy to index and annotate subjects within a text using a simple, wiki-like mark-up [internal-linking and external hyperlinks]. Users can discuss difficult writing or obscure words within a page to refine their transcription. The resulting text is hosted on the web, making documents easy to read and search”.¹¹

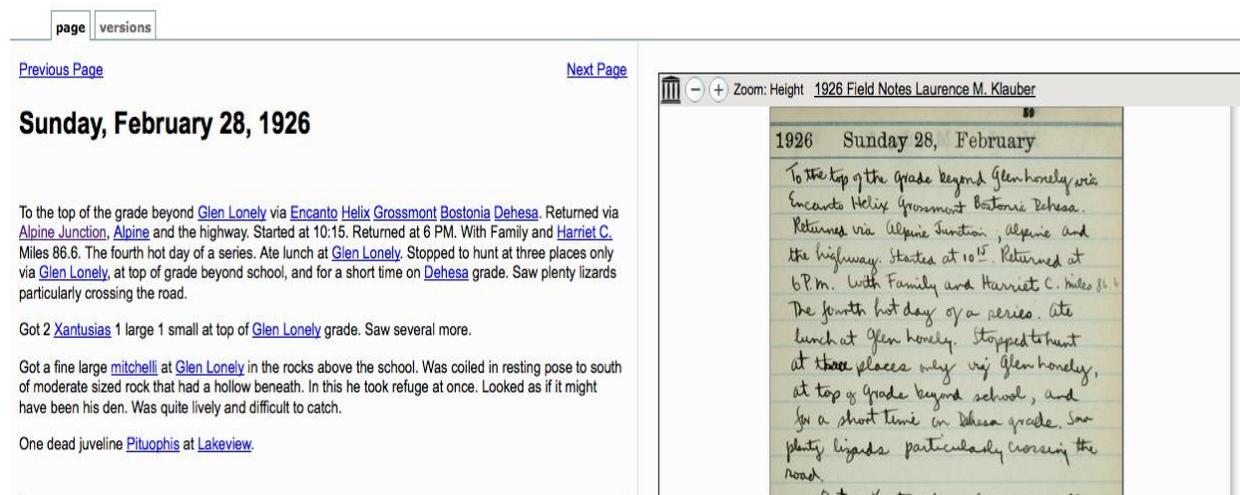
FromThePage is already being utilized for several projects. The San Diego Museum of Natural History is transcribing the [field notes of herpetologist Lawrence Klauber](#), Southwestern

¹⁰ http://beta.fromthepage.com/?ol=_hd_logo

¹¹ *ibid*

University is transcribing the [Mexican War Diary of Zenas Matthew](#) and Gavin Robinson is transcribing the his great-grandfather's [letters from a WWI prisoner of war camp](#).

Field Notes of Laurence M. Klauber — 1926 Field Notes Laurence M. Klauber



The screenshot displays a web page for a diary entry. At the top, there are tabs for 'page' and 'versions'. Below this, there are links for 'Previous Page' and 'Next Page'. The main heading is 'Sunday, February 28, 1926'. The text of the diary entry is transcribed and includes several hyperlinks to other pages, such as 'Glen Lonely', 'Encanto Helix Grossmont Bostonia Dehesa', 'Alpine Junction', 'Alpine', 'Dehesa', 'Alpine Junction', 'Alpine', and 'Lakeview'. To the right of the text is a zoomed-in image of the original handwritten diary page, showing the date '1926 Sunday 28, February' and the beginning of the entry: 'To the top of the grade beyond Glen Lonely via Encanto Helix Grossmont Bostonia Dehesa. Returns via Alpine Junction, Alpine and the highway. Started at 10:15. Returned at 6 P.M. With Family and Harriet C. Miles 86.6 Miles. The fourth hot day of a series. Ate lunch at Glen Lonely. Stopped to hunt at three places only viz Glen Lonely, at top of grade beyond school, and for a short time on Dehesa grade. Saw plenty lizards particularly crossing the road. Not a Xantusia. ...'.

Figure 1: FromThePage transcription and annotations for Laurence M. Klauber's diary

As can be seen in the screenshot above of FromThePage the interface is simple. You can see the transcribed text including the links to separate internal pages which act like the text's own encyclopedia. Clicking on one of these links takes the visitor to a similar looking page which contains more contextualisation like character biographies, place descriptions, images, as well as links to other pages where certain specific terms appear again.

FromThePage also has a comment feature that provides the annotation process with a social aspect. Much like Wikipedia, it allows for revisions, tags, etc. to be discussed adding yet another layer of crowdsourcing.

2.1.3 Classification

Classification tools help users gather descriptive metadata related to objects in a collection. Europeana has a classification prototype called The EuropeanaConnect Annotator Suite. This tool allows end-users to add metadata to images, maps, audio and video. The UGC this procures can be simple tags and annotations, stories, and external links, all of which provide both end-users and the collection owners with a richer digital asset than originally available.

A unique feature of the Annotator Suite is the Linked Open Data it utilises. This allows for information to be easily linked, discovered and shared from other resources. The Annotator Suite uses Dbpedia to accomplish this.

DBpedia¹² is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web. The Dbpedia database can be accessed through DBpediaSpotlight which is easily implemented. It allows users to actively search within the database for relevant secondary content to be used for enriching other content or webpages.

For example, upon entering the query “Alps”, one is given a link to a list of information including a summary of the Alps from Wikipedia, relevant images, information about other mountain ranges, as well as specific mountains in the Alps and “Alps” in various languages. Implementing Dbpedia in an online collection of texts allows for almost immediate link and enrichment and since Dbpedia is directly linked to Wikipedia, as Wikipedia is updated by its community so to will the respective text.



Confidence: 0.5 Language: English

n-best candidates **SELECT TYPES...** **ANNOTATE**

Alps

¹² <https://github.com/dbpedia-spotlight/dbpedia-spotlight/wiki>

Figure 2: Example of Dbpedia Spotlight

About: Alps
 An Entity of Type : [YagoPermanentlyLocatedEntity](#), from Named Graph : <http://dbpedia.org>, within Data Space : [dbpedia.org](#)

The Alps, one of the great mountain range systems of Europe, stretch approximately 1,200 kilometres across eight Alpine countries from Austria and Slovenia in the east, Switzerland, Liechtenstein, Germany, France to the west and Italy and Monaco to the south.

Property	Value
dbpedia-owl:abstract	<ul style="list-style-type: none"> ▪ The Alps, one of the great mountain range systems of Europe, stretch approximately 1,200 kilometres across eight Alpine countries from Austria and Slovenia in the east, Switzerland, Liechtenstein, Germany, France to the west and Italy and Monaco to the south. The Alps were formed over hundreds of millions of years as the African and Eurasian tectonic plates collided; the extreme shortening caused by the event resulted in marine sedimentary rocks rising by thrusting and folding into high mountain peaks such as Mont Blanc and the Matterhorn. Mont Blanc spans the French–Italian border, and at 4,810.45 m is the highest mountain in the Alps. The Alpine region area contains many peaks higher than 4,000 m, known as the “four-thousanders”. The altitude and size of the range affects the climate in Europe; in the mountains precipitation levels vary greatly and climatic conditions consist of distinct zones. Wildlife such as ibex live in the higher peaks to elevations of 3,400 m, and plants such as Edelweiss grow in rocky areas in lower elevations as well as in higher elevations. Evidence of human habitation in the Alps goes back to the Paleolithic era. A mummified man, determined to be 5,000 years old, was discovered on a glacier at the Austrian–Italian border in 1991. By the 6th century BC, the Celtic La Tène culture was well established. Hannibal may have crossed the Alps with a herd of elephants, and the Romans had settlements in the region. In 1800 Napoleon crossed one of the mountain passes with an army of 40,000. The 18th and 19th centuries saw an influx of naturalists, writers, and artists, in particular the Romantics, followed by the golden age of alpinism as mountaineers began to ascend the peaks. In World War II the Third Reich invaded the Alpine countries, with the exception of Switzerland and Liechtenstein; Adolf Hitler kept a base of operation in the Bavarian Alps throughout the war. The Alpine region has a strong cultural identity. The traditional culture of farming, cheesemaking, and woodworking still exists in Alpine villages, although the tourist industry began to grow early in the 20th century and expanded greatly after World War II to become the dominant industry by the end of the century. The Winter Olympic Games have been hosted in the Swiss, French, Italian, Austrian and German Alps. At present the region is home to 14 million people and has 120 million annual visitors.
dbpedia-owl:thumbnail	<ul style="list-style-type: none"> ▪ http://upload.wikimedia.org/wikipedia/commons/thumb/8/82/Mont_Blanc_oct_2004.JPG/200px-Mont_Blanc_oct_2004.JPG
dbpedia-owl:wikiPageExternalLink	<ul style="list-style-type: none"> ▪ http://www.alpine-space.eu ▪ http://modis.gsfc.nasa.gov/gallery/individual.php?db_date=2005-09-17 ▪ http://pages.unibas.ch/earth/tecto/research/Schmid_et_al_2004_Ecl.pdf ▪ http://www.alpconv.org/en/publications/alpine/Documents/Vademecum_web.pdf ▪ http://www.unige.ch/climate/Publications/Beniston/ESP2011.pdf
dbpedia-owl:wikiPageID	▪ 981 (xsd:integer)
dbpedia-owl:wikiPageInLinkCount	▪ 4047 (xsd:integer)
dbpedia-owl:wikiPageOutLinkCount	▪ 460 (xsd:integer)
dbpedia-owl:wikiPageRevisionID	▪ 548196151 (xsd:integer)
dbpprop:country	<ul style="list-style-type: none"> ▪ Austria ▪ France ▪ Italy ▪ Liechtenstein ▪ Monaco

Figure 3: Results from Dbpedia for the term “Alps”

2.1.4 Complementing Collections

Complementing collections allows for institutions to grow their collections by crowdsourcing relevant media via interactive platforms.

An example of software that enables institutions to do this easily is the open source software made available by the RunCoCo team at Oxford University. The CoCoCo software can be used for community collections online or any activity where contributors fill in a form and (optionally) attach images, sound files, text document or other files to their submission. The look and feel of the interface and the content of the pages can be changed to fit different projects and purposes¹³. The software only provides backend support for the time being, meaning that institutions will have to develop further if they want to surface the aggregated UGC.

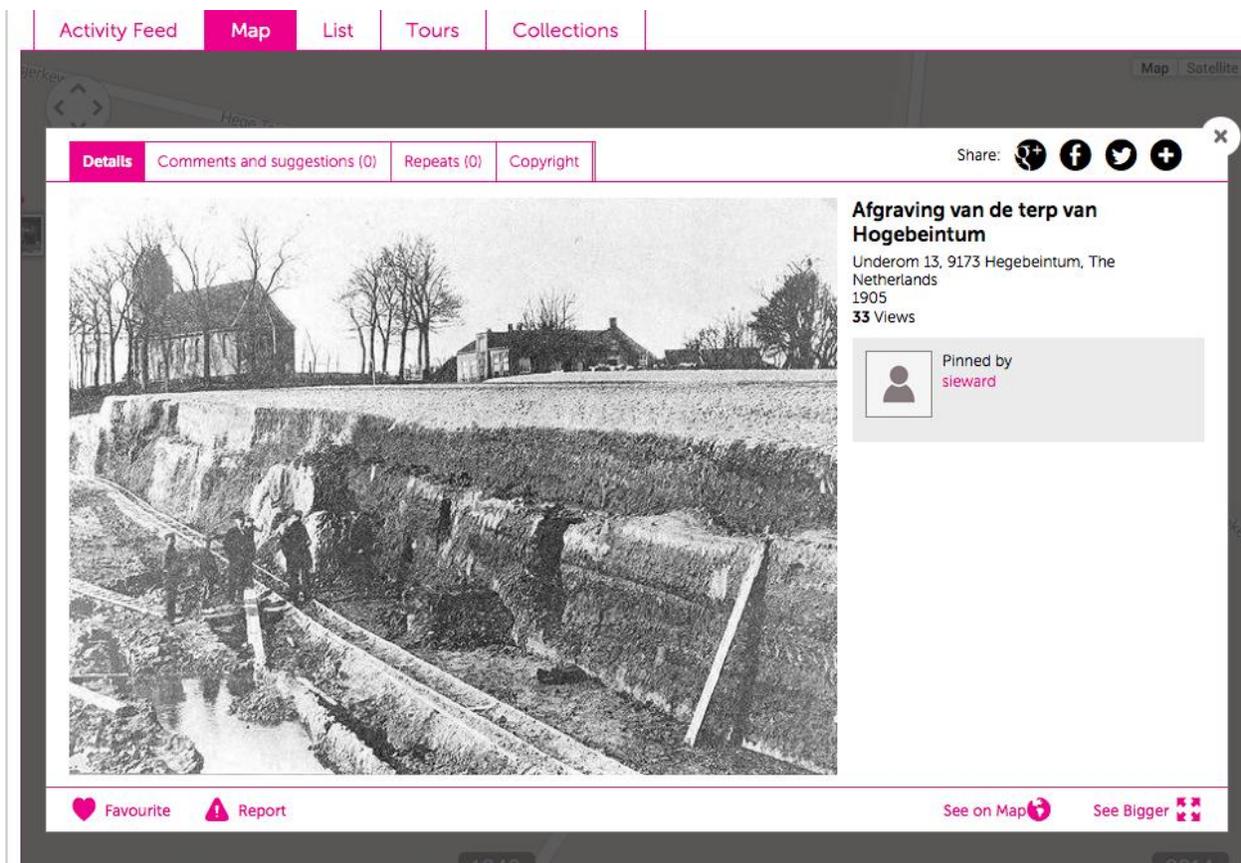


Figure 4: Example of UGC item pinned on Historypin

¹³ http://projects.oucs.ox.ac.uk/runcoco/downloads/CoCoCo/CoCoCo_interface.pdf

An example of a masterfully done front end for complementing collections would be Historypin's. Historypin allows users to "pin" content to their interactive map. The content is then made available on the portal. This allows for institutions to not only gain content and metadata but also public appeal due to the interface's sleek, modern design and layout. For the time being this is not open source. However, with the amount of interactive map open source software and collection management systems available on the FLOSS Inventory, it is within the reach of any institution with a skilled technical team to easily develop something similar with very little added software costs.

2.1.5 Co-Curation

Co-curation software is the most prevalent software available in the FLOSS Inventory after contextualisation and classification tools. Co-curation software allows individuals to create their own exhibitions, apps, timelines, and maps. This software can be paired with individuals' collections and content retrieval software like DBpedia, the Europeana API, open datasets like 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.

For example, Timeline JS is an open source tool that allows one to easily build visually striking interactive timelines which can be embedded on a website. The timelines are made by uploading a Google SpreadSheet. There is even a simple template, making the work that much easier.

Start Date [1]	End Date [2]	Headline [3]	Text [2]	Media [4]	Media Credit [2]	Media Caption [2]	Media Thumbnail [5]
10/1/2011 0:00:00		Google Spreadsheet Example	This is a test of using google spreadsheets as a source for the timeline tool Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi commodo, ipsum sed pharetra gravida, orci magna rhoncus neque, id pulvinar odio lorem non turpis. Nullam sit amet enim.	http://www.flickr.com/photos/zachwise/6115056146/ /n/photostream	Zach Wise/verite.co	Chicago by zach.wise	
11/12/2012 0:00:00	1/12/2012 0:00:00	A Flickr Example	Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi commodo, ipsum sed pharetra gravida, orci magna rhoncus neque, id pulvinar odio lorem non turpis. Nullam sit amet enim.	http://www.flickr.com/photos/zachwise/6887123779/ /n/photostream	Zach Wise/verite.co	Enjoying the snow while sheltered at the train station	
11/1/2011 0:00:00	11/1/2011 0:00:00	Another Flickr Example	Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi commodo, ipsum sed pharetra gravida, orci magna rhoncus neque, id pulvinar odio lorem non turpis. Nullam sit amet enim.	http://www.flickr.com/photos/zachwise/6894476779/ /n/photostream	Zach Wise/verite.co	Chicago to NYC	
11/16/2011 0:00:00	11/17/2011 0:00:00	Vimeo — Jay Z - "Hello Brooklyn"	Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi commodo, ipsum sed pharetra gravida, orci magna rhoncus neque, id pulvinar odio lorem non turpis. Nullam sit amet enim.	https://vimeo.com/10089801	Greg Solenström	Jay Z - "Hello Brooklyn"	

Figure 5: Timeline JS Upload Template

Timeline JS has been utilised by various organisations to create their own timelines including Time Magazine, LeMonde.fr, Al Jazeera, and VH1. While these use cases are not UGC or crowdsourced, Timeline JS's ease of use and adaptability make it possible for institutions to offer up data and media pertaining to certain historical events and allow users to generate their own personal timelines.

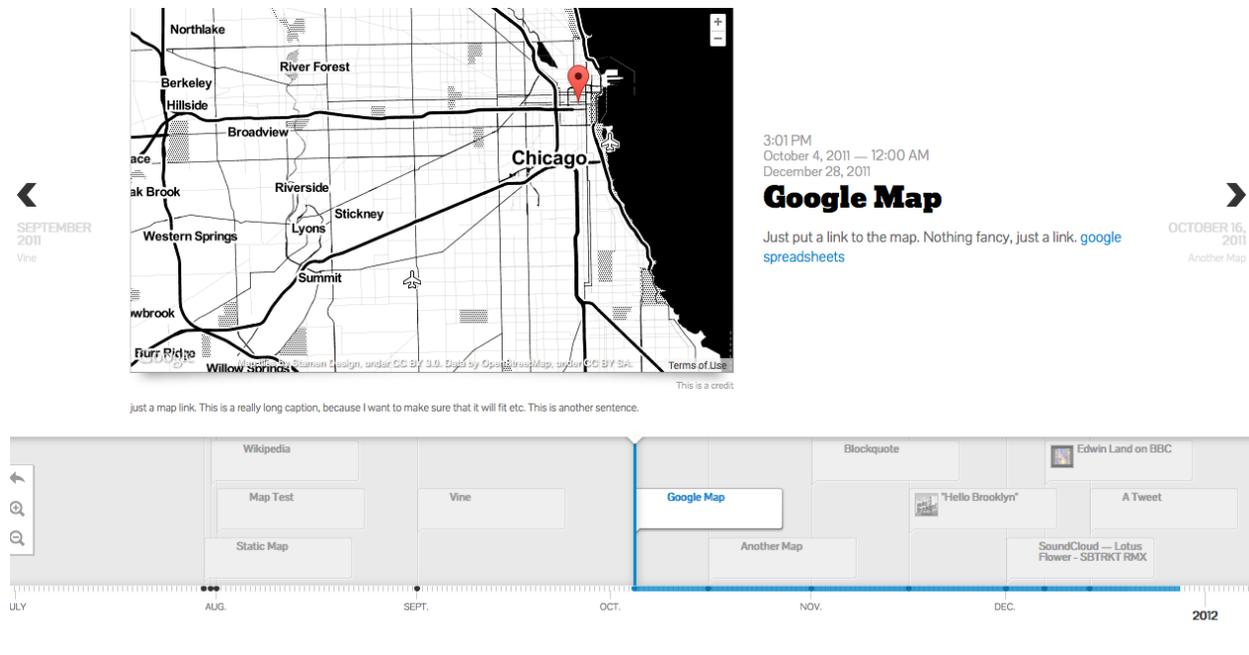


Figure 6: Timeline JS example timeline

Another possibility that Timeline JS offers when it comes to UGC and crowdsourcing media is an institution creating an open Google SpreadSheet that users can freely contribute content to. It can for example, be thematically based within a city neighborhood between 1900-2000. Users are able to upload their own images and videos along with text or stories alongside the institution's collection. The end result is an interactive timeline spanning a century of a neighborhood's history.

All of this can be done with very minimal technical work by the organising institution making it a viable option for those who wish to venture into crowdsourcing and UGC for the first time.

2.2 UGC Taskforce ([More info in Annex 2](#))

The goal of the Europeana UGC Taskforce is to improve the identification and benchmarking of services and best- practices as building blocks for the creation of a Europeana UGC ecosystem. To be a point of contact for all projects that want to apply UGC approaches within the Europeana network and to encourage cross-fertilization of ideas between projects and identify duplication of effort. And lastly, to make policy recommendations for the Europeana Network regarding the role UGC can play in the context of Europeana.

The above will be accomplished in three phases. Phase 1 will be identifying actors within and beyond the network and providing them with a very short questionnaire. The taskforce needs to find out who is using UGC, why, how and with how much success? Phase 2 will strive to uncover the trials and tribulations of UGC via a longer and more detailed questionnaire. Phase 2 will seek to understand UGC within Europeana. This will be accomplished through a detailed review of technology, organisation and impact.

Chapter 3. Measuring Impact of UGC in Europeana

Being able to quantifiably show the results of crowdsourcing UGC in Europeana related projects provides statistical support for further projects and policy recommendations. This section presents two pilot prototype currently under development as part of Europeana Creative which will facilitate UGC as well as statistical reports from Europeana 1914-1918 and Europeana 1989, two of Europeana's most successful UGC and crowdsourcing initiatives. Lastly, there is a discussion of the GLAMwiki Toolset which will make it easier for institutions or individuals with large collection to batch upload content to Wikimedia Commons as well as Europeana.

3.1 Europeana Creative

One way to encourage the public to contribute UGC is to do so in a creative way. Europeana Creative aims to promote creative re-use of digitised cultural heritage content through the development of 5 different pilots. These pilots are then used as a source of inspiration for other designers and developers who pitch their concepts and ideas during one of the Challenge Events. The best pitches are given incubation and developmental support from the the Europeana Creative consortium to help get them off the ground.

Two of these pilots are already showcasing how crowdsourcing and UGC can be optimised via creative applications. the Social Networks Pilot and Tourism Pilot.

3.1.1 Social Networks "Sound Connections"

The Social Networks Pilot, herein is referred to as "Sound Connections", makes use of the Netherlands Institute for Sound and Vision and British Library's field recording collections and the Historypin interface from We Are What We Do to engage communities and encourage them to enrich sounds with content from various open sources.

historypin

Can you help us enrich these field recordings of bird sounds?



The British Library and Netherlands Institute for Sound and Vision have many recordings of bird sounds from across the UK and Netherlands. We need your help to add contextual information to them. These could be photographs, videos, other sound recordings, articles, or other pieces of relevant information.

Use the map below to explore the sounds and click "comment" to add any contextual links to them. Here are some sources you might like to explore:

Figure 7: Sound Connections homepage

Sound Connections uses the Historypin interactive map from project partner We Are What We Do. The field recordings are pinned on the map and users can explore the sounds and can upload relevant content which corresponds to the sound. Three themes are being explored in this pilot, birds, aviation and city sounds (London and Amsterdam).

The pilot coordinators will approach members of these specific communities and encourage them to search for relevant content. The content can come from various sources like Europeana (whose API will be implemented), Wikimedia Commons, Flickr etc. However, to promote the generation of new UGC, community members will also be able to upload their own media.

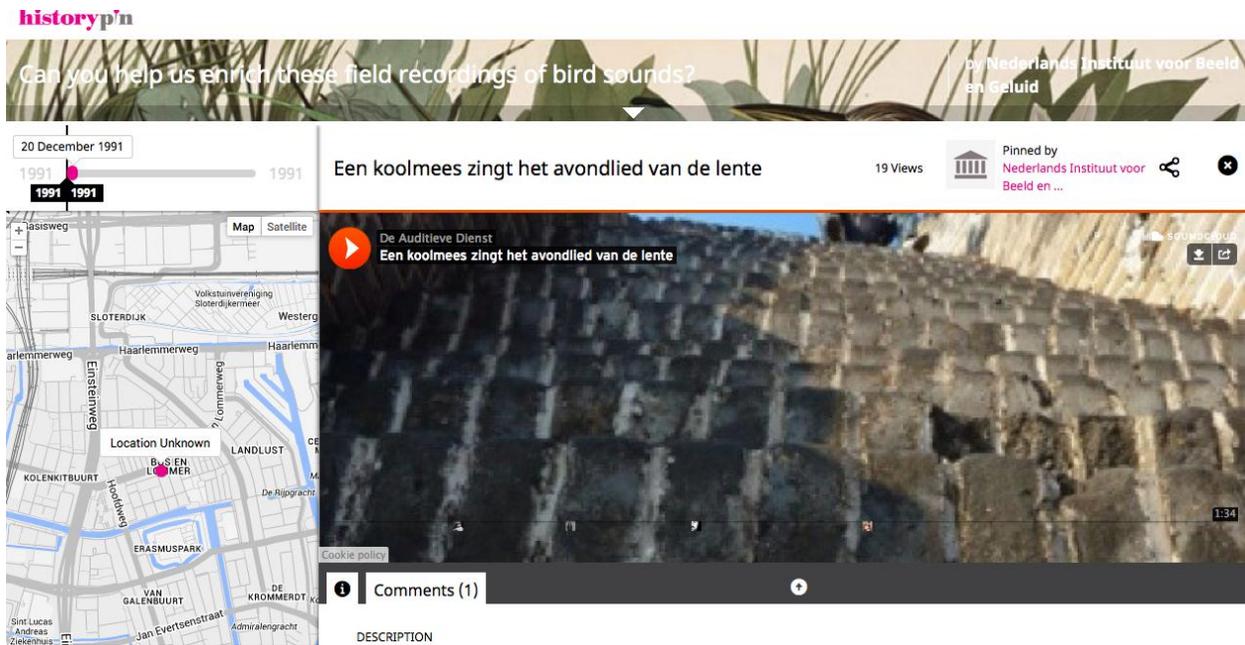


Figure 8: Example of audio file on “Sound Connections”

For example, they can select the sound featured in the image above. The bird is a Koolmees (Great Tit in English). Users could either venture to Europeana to search for an image like the one featured below or if they had an image, sound, or video in their own collection then they would be welcome to link that content to the sound.



Figure: “Koolmees door Marinus Koekkoek”, Public Domain. Source:Europeana

What makes this pilot so unique and valuable for the marketplace is its malleable design. While the backend software is static the front end design can be easily configured to meet the needs of any community. Therefore the tool can be sold or leased for a set amount to any institution or community organisation looking to engage the public and community members without having to make a substantial technical and financial investment. Being able to easily interact with communities and generate content re-use as well as crowdsourced UGC will hopefully entice more engagement on behalf of museums and cultural institutions.

The pilot is currently still in its developmental phase but plans to begin engaging communities start this month (June 2014) and the final prototype will be available at the end of 2014.

3.2.2 Tourism Pilot “VanGoYourself”

The Tourism pilot herein referred to as, “VanGoYourself” 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.

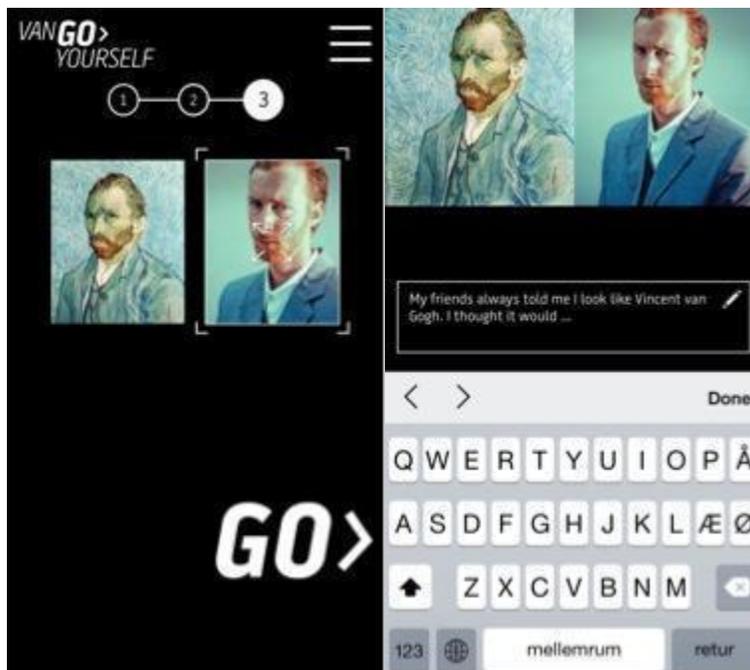


Figure 9: VanGoYourself Design Template

As can be seen in the figure above, individuals are provided with an image of a historical piece of art. They then recreate the image themselves and twin the two for comparison. Users can add a caption and then can share the image via the VanGoYourself website¹⁴, Facebook, Twitter, or Pinterest. The image may also be downloaded but not on the mobile app. The downloaded images are made available under a Creative Commons CC BY-SA license. All twinned images are made available for viewing on the VanGoYourself Tumblr.¹⁵

While VanGoYourself focuses more on creative content re-use and audience engagement, by allowing the twinned images to be shared again on the website and social media, it generates a unique interpretation of co-curation.

The pilot, while still in its early forms has already become very popular. VanGoYourself has already received 700 recreations. The majority of these twinned images are self portraits or

¹⁴ <http://vangoyourself.com/>

¹⁵ <http://vangoyourself.tumblr.com/>

smaller paintings with 2-3 people. There has also been an influx of classrooms recreating larger scale art works.



Figure 10: The Last Supper, Leonardo da Vinci (1452-1519) vanGo'd by Mr Lacey's A2 Biology Class iivi Form, Newport, IOW. Recreation: CC BY-SA, Original: Public Domain

3.2 Europeana 1914-1918

Europeana 1914-1918, discussed in section 1.2 of this deliverable, has produced successful results regarding UGC and crowdsourcing. Europeana 1914-1918 mainly collects content two

different ways, roadshows and via the portal. Roadshows are held at different locations across Europe (over 100 thus far in France alone) and individuals are encouraged to bring in any objects or stories they have from WWI. These stories are recorded on the spot and the images are digitised and metadata is collected and with the permission of the individual, published on Europeana 1914-1918. End-users may also directly upload digitised items or stories directly via the Europeana 1914-1918 portal.

Thus far the amount of content UGC delivered to Europeana 1914-1918 amounts to:

- Total number of published contributions: 9939
- Total number of submitted contributions pending approval: 1553
- Total number of draft contributions: 2955 total

Published contributions consist of 9,939 stories and over 126,959 attached item scans. The draft ones are backlogged items from past collection days. These will be published in due time. Submitted contributions mainly consist of online contributions that must reviewed before publication. Europeana 1914-1918 also allows for annotations and tags to be made. Thus far there have been 309 tags and only 11 annotations.



The screenshot shows a user-generated content entry on the Europeana 1914-1918 platform. On the left, there is a photograph of a man, Harry Mitchell, sitting on a wooden stool. Below the photo is a blue link that says "View full item" and the text "Item 1 of 2". To the right of the photo, the name "Harry Mitchell" is displayed in a large, bold font. Below the name, there are sections for "Alternative title" (Photograph), "Story" (a personal narrative about the man's life and family), "Contributor" (clifford potter), "Person 1" (Harry Mitchell), and "Person 2" (Hilda Hosier). At the bottom left of the entry, there are social sharing icons for Facebook, Twitter, Email, and Print, along with a "Share on" label. Below the sharing icons is a "Translate to ..." dropdown menu and the text "Powered by Microsoft® Translator".

Figure 11: Example of UGC contributed to Europeana 1914-1918

3.3 Europeana 1989

Europeana 1989¹⁶, discussed in section 1.2 of this deliverable is a pan-European project concerning the political and social changes in Central and Eastern Europe in the year 1989; commonly known as the fall of the Iron Curtain. The project aims to create a vivid and complete picture of the revolutionary events in Europe with stories, photos, videos and sound recordings from every country affected. Europeana 1989 also conducts roadshows as well as allowing users to upload content via the portal.¹⁷

Thus far there are 1186 digitised objects in the collection days repository. On the Europeana 1989 website there are now 10.412 objects. We are currently in the process of ingesting both collections in the main Europeana portal.

The screenshot shows the Europeana 1989 website interface. At the top, there are navigation tabs for 'Stories', 'Map', and 'Gallery'. A language selection dropdown is set to 'English' and is powered by Google Translate. The main content area features a story titled 'Years of Euphoria' by user 'piroska62', dated '23 September 1986 - 16 May 1990' from 'Hungary'. The story text describes the author's family's experience during the Hungarian Revolution of 1956, mentioning their father's role in a revolutionary council and their mother's flight to America. To the right of the text is a black and white photograph of a large crowd of people in Budapest, holding a banner that reads 'VALDNY REFORMÁCIÓ'. A 'Next' button is visible in the bottom right corner of the story preview.

Figure 12: Example of a story and image uploaded to Europeana 1989

3.4 GLAMwiki Toolset

The GLAMwiki Toolset is an endeavor to provide tools that make batch uploads of GLAM (Gallery, Library, Archive, Museum) content to Wikimedia Commons more streamlined and easier for institutions. The *motivus* for this project is based on the current tedious and time consuming nature of uploading content to Wikimedia Commons. The project is being

¹⁶ <http://pro.europeana.eu/web/europeana-1989/>

¹⁷ <http://www.europeana1989.eu/en/>

undertaken by Wikimedia Netherlands, Wikimedia France, Wikimedia UK, Wikimedia Switzerland and Europeana.

The GLAMwiki Toolset has two main objectives:

1. Creating a GLAM upload system to Wikimedia Commons
2. Report on requirements for usage and reuse statistics for GLAM content

3.4.1 Uploading to Wikimedia Commons

In order to make batch uploading as easy as possible the GLAMwiki Toolset will enable multiple features. These include supporting the major metadata standards (EDM, Dublin Core etc) and aligning them with the Wikimedia Commons' template, make all metadata template creations more user friendly, create a preview function that would allow uploaders to view their content batches before publication and provide user friendly templates to allow for proper attribution.

Furthermore, the toolset will allow for dual uploads, meaning batches can be simultaneously uploaded to both Wikimedia Commons and Europeana. End users will be able to download GLAM content batches with metadata in standard Commons formats.

3.4.2 Report on Requirements for Usage and Reuse Statistics for GLAM Content

The Report on Requirements for Usage and Reuse Statistics for GLAM Content was written by Maarten Zeinstra from the Amsterdam based thinktank, Kennisland. The report covers the current state of metrics that measure the impact and use of content made available by GLAMs to Wikimedia projects within Wikimedia Commons. The report also gives an overview of requirements of GLAM tools.

The report findings show that there is a high demand for solid tools that provide regular analytics of media content uploaded to Wikimedia Commons.

Survey participants for the report were asked which of the below listed actions they would like to see statistics for. The numbers represent the number of survey respondents who desired the corresponding information on the left.

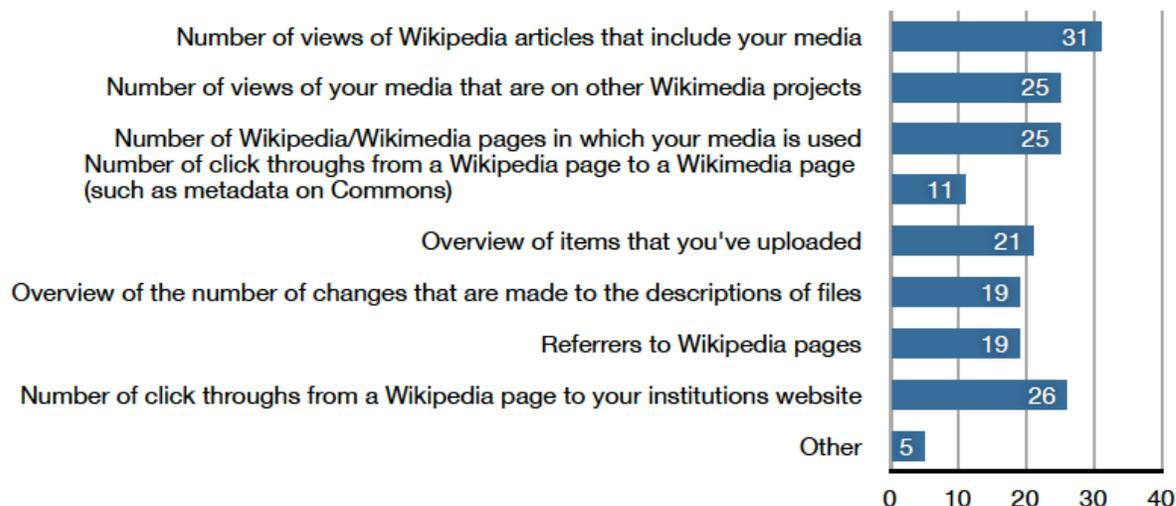


Table 3: GLAMwiki Toolset Report-Requested information

Conclusion and Future work

Facilitating user generated content and crowdsourcing makes cultural heritage an inclusive and participatory aspect of European society. Europeana has made 30 million metadata records available all in one place under a Creative Commons 0 license. The task that this deliverable covers Task 2.1 to *Operationalise tools to enable end user contributions to Europeana content*. Through Europeana Awareness and ancillary activities Europeana and its network partners have been very visible and active within this task. The establishment of multiple platforms that produce UGC via crowd sourcing as well as the software that Europeana is either aggregating or developing have been successful. Through Awareness, and projects that started since then, Europeana is working on including a number of technical components and supporting infrastructure to support various types of content (be it metadata or content) by end users. Awareness is currently in its final stages of development, and parts of the work listed in Chapter 1. are now taken on board in EuropeanaV3 and EuropeanaCreative.

Annex 1: Tools for User Generated Content

As part of the work in Europeana Awareness, Europeana Foundation, Sound and Vision, NTUA and the wider EuropeanaTECH community have been keeping track of technology that is used to gather and manage user generated metadata and content.

The contributions have been shared in EuropeanaPro (the pages on the Europeana websites aimed at professionals) and are also included below.

Annotator

Description:

The Annotator is an open-source JavaScript library and tool that can be added to any webpage to make it annotatable.

Developers:

Open Knowledge Foundation

Project website:

<http://okfnlabs.org/annotator/>

Code repository:

<https://github.com/okfn/annotator/>

License:

MIT / GNU GPLv3

Quality of Documentation:

There is a well structured [Github](#) with well written documentation, including API docs. Also there is a dedicated [website](#) listing information about the tool and providing links to all the important information such as a [tutorial](#), a link to a [live demonstration page](#) where you can test out the tool, a link to a [plugin development page](#), etc.

Ease of Adaptation:

The extensibility of the tool is promoted on the main [website](#). There is a link to a [plugin development page](#) containing a well written comprehensive tutorial.

Code Quality:

The code is maintained in GitHub. CoffeeScript is used as the main language and all files are packaged in a logical structure. In general, the code includes extensive comments. The code is published for download as a minified JavaScript file. A file with the authors of the code is available, listing 20 contributors to the code.

Examples:

<http://okfnlabs.org/annotator/demo/>

Notes

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXQXFSTkE#gid=1

Annotorious

Description:

Annotorious is a JavaScript annotation library for images and zoomable images. Link the Annotorious source files (CSS and JavaScript) into an existing HTML page, and images inside this page will be enhanced with an interactive drawing and commenting tool.

Developers:

Peter Pilgerstorfer and Paul Weichhart

Project website:

<http://annotorious.github.io>

Code repository:

<https://github.com/annotorious/annotorious>

License:

LGPL

Quality of Documentation:

The main website is very informative and it's main menu includes links to: a demo page, a getting started page, a plugin page an API documentation page and finally an about page.

Ease of Adaptation:

There is a [plugin development page](#) available. In combination with a comprehensive [GitHub page](#), it shouldn't be too hard to develop plugins.

Code Quality:

The code is maintained in GitHub. The JavaScript (using Google Closure) code seems clearly structured and includes comments. There is a [useful page](#) with information on how the code was built and can be adapted.

Examples:

<http://annotorious.github.io>

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXQXFSTkE - gid=1

Art4Europe

Description:

Art4Europe is an application which makes easier for the user to gain their knowledge about historical monuments. During our visit at the museum we can take a picture of a painting and the application will download data for us referring to selected painting by using very sophisticated methods of comparing pictures straight from Europeana API.

Developers:

iTraff Technology: Tomasz Grzywalski

Project website:

<http://www.art4europe.com/>

License:

Unreleased prototype

Examples:

<http://www.youtube.com/watch?v=C6PEz2d7OLE>

More information:

https://docs.google.com/spreadsheets/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

D3-cartogram

Description:

Continuous area cartograms with d3 and TopoJSON

Developers:

Shawn Allen

Project website:

<http://prag.ma/code/d3-cartogram/>

Code repository:

<https://github.com/shawnbot/d3-cartogram/>

License:

MIT License

Quality of Documentation:

The main site briefly describes what the tool is and what dependencies the tool has. A prominent, but somehow easy to overlook, link to GitHub is available. On GitHub there is a short readme, which mainly points to an example index.html.

Ease of Adaptation:

The example HTML file includes quite a number of JavaScript functions that need to be examined in order to be able to understand how it all works. Besides this, the code is small and it shouldn't be too difficult to adapt the code to use different (geographical) input.

Code Quality:

There is not much code and it does look readable, but there are not many comments to help you. Because the code uses d3.js and uses geographical algorithms, it takes a while to get in to.

Examples:

<http://prag.ma/code/d3-cartogram/#intlmig/2011>

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

DBpedia Spotlight

Description:

DBpedia Spotlight is a tool for automatically annotating mentions of DBpedia resources in text, providing a solution for linking unstructured information sources to the Linked Open Data cloud through DBpedia.

Developers:

Pablo Mendes (Freie Universität Berlin), Jun 2010-present.

Project website:

<https://github.com/dbpedia-spotlight/dbpedia-spotlight>

Code repository:

<https://github.com/dbpedia-spotlight/dbpedia-spotlight>

License:

Apache License, 2.0, LingPipe

Quality of Documentation:

Well documented, short documentation on Github, further documentation on Wiki, well structured and written.

Ease of Adaptation:

REST Webservices allow integration and adaption.

Code Quality:

Java/Scala based project (Java 48.7% Scala 48.1% Shell 1.7% Python 1.5%), actively maintained, last commit days ago, continous integration, well-structured, maven-based build, tests available. Many branches and contributors.

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

DigitalLabels

Description:

Django app to generate an interactive gallery label interface, optimised for touch screens and tablets.

Developers:

V&A Digital Media team

Project Website:

<http://vanda.github.io/DigitalLabels/>

Code Repository:

<http://vanda.github.io/DigitalLabels/>

License:

BSD License

Quality of Documentation:

Nice presentation, Basic documentation, well written and structured, documentation hosted on github pages.

Ease of Adaptation:

No APIs or documentation about interfaces available.

Code Quality:

Python/JavaScript based (Python 65.2% JavaScript 34.8%), hosted on Github, not being actively maintained, not many comments in code, but readable.

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

From the Page

Description:

FromThePage is an open-source tool that allows volunteers to collaborate to transcribe handwritten documents.

Developers:

Ben W. Brumfield

Project Website:

<http://beta.fromthepage.com/>

Code Repository:

<https://github.com/benwbrum/fromthepage>

License:

GNU AGPL v3.0

Quality of Documentation:

The FromThePage home page provides very little documentation and the Github page also has very little documentation. The documentation that is there is clear and visible but severely lacking in detail.

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

Gathrit

Description:

The concept of Gathrit is to provide a crowdsourcing toolset for cultural collections.

Developers:

James Morley

Project Website:

<http://gathr.it/>

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

ImageMontage

Description:

ImageMontage creates montages of images without the need to load them into a stack first. The source images can be located in one or multiple directories. The images can also have different sizes. To fit images into a montage, the macro scales them vertically to the same user-specified size. Images are added horizontally until they no longer fit on the row and then a new row is started. This process is repeated until all source images are rendered. If you see that the generated montage does not include all source images, change the options to reduce the size of images and/or increase canvas dimensions, and run the macro again.

Developers:

Lev Manovich, Matias Giachino, Jay Chow.

Project Website:

http://lab.softwarestudies.com/2014/03/how-to-visualize-4512-instagram-selfies.html?utm_content=buffer9ee7e&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer

Code Repository:

<http://rsbweb.nih.gov/ij/plugins/image-montage/index.html>

License:

Public Domain

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

ImagePlot:

Description:

ImagePlot is a free software tool that visualizes collections of images and video of any size. It is implemented as a macro which works with the open source image processing program ImageJ.

Developers:

Software Studies Initiative

Project Website:

<http://lab.softwarestudies.com/p/imageplot.html#features1>

Code Repository:

https://docs.google.com/document/d/1zkeik0v2LJmi1TOK4OxT7dVKJO7oCmx_fNP8SYdTG-U/edit?hl=en_US

License:

Public Domain

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

Kartograph

Description:

Kartograph is a simple and lightweight framework for building interactive map applications without Google Maps or any other mapping service.

Developers:

Gregor Aisch

Project Website:

<http://kartograph.org/>

Code Repository:

<https://github.com/kartograph/kartograph.py> and <https://github.com/kartograph/kartograph.js>

License:

LPGL (Kartograph.js) & AGPL (Kartograph.py)

Quality of Documentation:

There are two different version of this application with different specialty features that are clearly stated on the main page. Each version has its own clear, descriptive documentation both on the homepage and Github. No technical support available but there are active demos.

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

MapHub

Description:

Online tool for annotating (historical) maps. Annotations are published and shared.

Developers:

Maphub is being developed at Cornell Information Science, with support from the University of Illinois at Urbana-Champaign and the University of Vienna.

Project Website:

<http://maphub.github.io/>

Code Repository:

<https://github.com/maphub/maphub-portal>

License:

Apache License, Version 2.0

Example:

<http://maphub.herokuapp.com/>

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

Neatline

Description:

The Scholars' Lab designed Neatline as a suite of plugins for the open-source Omeka framework, which provides a powerful platform for content management and web publication. Through Neatline, you can create create rich representations of places, objects, events, narratives, and documents.

Developer:

Adam Soroka

Project Website:

<http://neatline.org/>

Code Repository:

<https://github.com/scholarslab/Neatline>

License:

Apache License v 2.0

Example:

<http://neatline.org/neatline-in-action/>

Quality of Documentation:

Neatline is extremely organized and thorough with their documentation. They provide visible demos of the applications and its capabilities. There is no technical support but the website is updated periodically.

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

Omeka

Description:

A Collection/Exhibition Management System

Developers:

Roy Rosenzweig Center for History and New Media, George Mason University

Project Website:

<http://omeka.org/>

Code Repository:

<http://omeka.org/download/>

<https://github.com/omeka>

License:

GNU GPL

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

OpenScribe

Description:

OpenScribe is an online community where people interested in handwritten material such as letters, diaries, journals etc can join together to help convert these documents into digital text.

Developers:

Google Project Hosting: jcdar...@gmail.com

Project Website:

<https://code.google.com/p/openscribe/>

Code Repository:

<http://openscribe.googlecode.com/svn/trunk/all/admin/readme.txt>

License:

GNU

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

Plumi:

Description:

Plumi is a Free Software video-sharing web application based on Plone

Developers:

EngageMedia in collaboration with Unweb.me

Project Website:

<http://blog.plumi.org/>

Code Repository:

<https://github.com/plumi/plumi.app>

License:

As Plumi pulls together a range of different products, different licenses apply to different elements of the software. However most are covered either by the GNU GPL or the Zope Public License

Example:

<http://blog.plumi.org/sites-using-plumi/>

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

PyBossa

Description:

PyBossa is a free, open-source, platform for creating and running crowd-sourcing applications that utilise online assistance in performing tasks that require human cognition, knowledge or intelligence such as image classification, transcription, geocoding and more!

Developers:

Open Knowledge Foundation

Project Website:

<http://www.pybossa.com>

Code Repository:

<https://github.com/PyBossa/pybossa>

License:

GNU Affero General Public License

Example:

<http://pybossa.com/app/>

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

ResourceSpace

Description:

ResourceSpace is a web-based, open source digital asset management system (DAM)

Developers:

Montala (under contract with Oxfam)

Project Website:

<http://www.resourcespace.org>

Code Repository:

<http://svn.montala.net/svn/resourcespace>

License:

BSD style license

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

TAP

Description:

TAP is a collection of free and open-source tools which support the creation and delivery of mobile tours. Currently TAP consists of authoring tools built on top of the content management system Drupal, a native iOS mobile application, and a web-based mobile application built upon the jQuery Mobile library.

Project Website:

<http://tapintomuseums.org/TAP>

Code Repository:

<https://github.com/IMAMuseum/tap-cms>

<https://github.com/IMAMuseum/tap-ios>

<https://github.com/IMAMuseum/tap-web-app>

Example:

<http://tapintomuseums.org/case-studies>

Timeline JS

Description:

TimelineJS is an open-source tool that enables you to build visually-rich interactive timelines and is available in 40 languages.

Developers:

Northwestern University Knight Lab

Project Website:

<http://timeline.knightlab.com/>

Code Repository:

<https://github.com/NUKnightLab/TimelineJS>

License:

Mozilla Public License, v. 2.0

Example:

<http://timeline.knightlab.com/>

Quality of Documentation:

Very well written documentation for users. Also very well written documentation for programmers (found on the Github page). Frequent documentation updates and live demo.

Ease of Adaptation:

This is a web-application that can be plugged in on websites. Information for plugging in the webapp is available even on the first page of the project's custom website. More extensive information exists on the GitHub page of the project. There is a list of known installations.

Code Quality:

Excellent code quality and package structure. The project is a work of academic level and follows almost every direction towards openness for collaboration and explicitness of descriptions.

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

Tesseract

Description:

Tesseract is probably the most accurate open source OCR engine available. Combined with the Leptonica Image Processing Library it can read a wide variety of image formats and convert them to text in over 60 languages.

Developers:

Ray Smith

Project Website:

<http://code.google.com/p/tesseract-ocr/>

Code Repository:

<http://code.google.com/p/tesseract-ocr/source/checkout>

License:

Apache License 2.0

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

Waisda?

Description:

Crowdsourcing video annotation software.

Developers:

Q42: Kamil Afsar, Martijn van Steenbergen

Project Website:

<http://blog.waisda.nl/>

Code Repository:

<https://github.com/beeldengeluid/waisda>

License:

GPL

Example:

woordentikkertje.manbijthond.nl

More information:

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=1

Annex 2: Strategies for user generated content and crowdsourcing in museums and cultural heritage

Lars Wieneke (CVCE, Luxembourg), Susan Hazan (The Israel Museum, Jerusalem)

Introduction:

User-generated content has become part and parcel of the mainstream Internet experience. Services like YouTube, Flickr or the Wikipedia provide platforms that encourage, enable and build on the creation of millions of users. Today, museums and memory institutions actively explore the inherent potential of user-generated content both to collect and share content with and from their audiences but also as a new tool for museum mediation.

This workshop brought together the dispersed knowledge of practitioners from the digital heritage community as well as researchers from neighboring disciplines to exchange best practices in engaging audiences, managing contributions and creating sustainable platforms for user-generated content. The workshop built on a recent survey initiated by the Europeana task-force on user-generated content that was answered by more than forty museums and memory institutions in Europe and the USA. The results of this survey will be used to contextualize the individual experiences of the participants and to review potential strategies for the integration of user-generated content in the work of museums and memory institutions.

The following chapter presents selected contributions from the workshop and will draw a conclusion with the main findings.

A tale of three terms: User Generated Content, Crowd-Sourcing and Engagement: from prescription to process

Stuart Dunn (King's College London) and Marion Dupeyrat (Ina, Paris)

This paper will explore a series of mappings between different terms: User-Generated Content (UGC), crowd-sourcing and public engagement on the one hand; and the notions of 'expert' and 'amateur' on the other. These terms, all critical to various aspects of the collection, management and presentation of cultural heritage, share a common feature: their meanings have become blurred by the rapid uptake of digital technology in so-called 'memory institutions, especially national museums, libraries and archives. Time and time again for example, the notions of expert and amateur get lost in the larger term 'citizen', as in citizen science, or citizen history. The implied superiority of the one over the other has no meaning when a shared task (whether that task be mechanical, critical or complex) cannot be achieved without collaboration between the two. Similarly, any attempts to define UGC have been overwhelmed by public perceptions informed by Facebook, blogging, Twitter etc, to the point where UGC is assumed to be anything posted online through a web browser in the post-Web 2.0 world. Such un-nuanced terminology is a hindrance to a critical understanding of the role of digital technology in cultural heritage. As cultural collections share more and more content online, and engage with social media, more and more opportunities emerge for the public to engage in more and more sophisticated ways. We tackle this in this paper.

We will draw on two recent studies which have examined this in detail: a 2013 benchmarking study by Ina EXPERT; and a scoping exercise on 'humanities crowd-sourcing' published in the same year, and commissioned by the UK's Arts and Humanities Research Council. These exercises shared a common approach, in that rather than focusing on types of content and engagement, or individual institutions, as previous work in the area have, they sought to define the increasingly complex relationship between the public and organizations (and thus critique those terms themselves) as sets of processes. It is only through processes - and the dynamism that this term implies - that transformations such as audience to participant, data to content and curation to memory can be understood. This is essential, as the digital world enables these processual transformations in new ways. We will stress the importance of process in this paper, and reflect on how a clearer understanding of process in the cultural heritage sector, as viewed through the case studies we have looked at, can help us to better understand the increasing fluidity of terms such as User Generated Content, crowd-sourcing and engagement.

Ancient Lives: Creating Digital Texts and New Technology for Editing and Data Mining Ancient Fragments

Dr James Brusuelas, Oxford

Since August 2011, Ancient Lives has recorded well over 1.5 million transcriptions of ancient Greek papyri (over 7 million characters), the work of over 250,00 online collaborators. The result was not simply the creation of big data, but the inception of an entirely different way of conceiving and interfacing ancient digital texts. Put simply, Ancient Lives has created something that has never existed before: a database of unedited Greek texts. We have strings of Greek characters without word division or any modern editorial convention. However, to access and make full use of that data, as texts actually read in antiquity, new algorithmic methods and digital tools that merge machine and human intelligence are required. The purpose of this paper is to showcase, first, the Ancient Lives' method for data extraction, curating, and producing digital Greek texts from this unique crowd sourced dataset. Then, and perhaps more importantly, to introduce two new Ancient Lives grant funded projects, in which new tools are being developed for the digital editing, data mining, and researching Greek and Coptic fragments through an advanced online interface.

Museums Visitors' Generated Content

Roei Amit, Head of Digital, RMN-GP

In recent years the concept and practices of User Generated Content became an important component of our daily lives: Wikipedia, social networks, video platforms, Web 2.0, for example, and the democratization of Smartphones and the vast use of their connectivity and cameras are participating as well in this vast phenomenon.

The RMN-GP (Réunion des Musées Nationaux – Grand Palais) is experiencing these changes and is updating its audience's strategy and the ways it can interact with them. On the one hand, how can we interact and develop our relations with the public and the content created spontaneously, and on the other hand how we reposition these relations no longer as top down transmission, but as an interactive dialog of exchange.

Over the last year we have changed our visiting regulations to permit photography by the public (unless marked specifically otherwise), we have installed screens in our reception halls and created dedicated spaces on our web sites to make visible and present our visitors photographic and textual creations. We have greatly accelerated our interactions on our social networks and launched several operations, tools and applications that interact with these practices.

“Dynamo”, a thematic show (Grand Palais, Paris, April-July 2013) about the kinetic art of the 20th century, was accompanied with innovative application inviting the public to take pictures and comment on selection of works, publishing them simultaneously and creating a participative kaleidoscopic eCatalogue of the exhibition which was presented on the web, in the application and at the museum video wall.

The exhibition “Grand Atelier de Midi”, (From Cezanne to Matisse, June – October 2013, Marseille - Aix en Provence), moved to an outdoor experience, inviting the public to geo localize a selection of the master pieces presented at the exhibition, taking pictures from the artist point of view, and then publish the diptych. For the “George Braque's” retrospective (Grand Palais, Paris, September 2013- January 2014) the application: La fabrique cubiste, teaches cubism in an amusing and pedagogical way, using cubist filters that can be applied on the photos taken by the public and then published on the social networks.

Last autumn we've organized two textual creative on-line workshop: for our exhibition about “Dreams at the Renaissance” (Musée de Luxembourg, October 2013- January 2014) we asked our twitter community to tell a dream in a twitter format, and for the Felix Vallotton retrospective (Grand Palais, October 2013- January 2014) we asked the We Love Words community to send us short stories inspired by the artist paintings, a selection of ten novels (from the 300 received) were published as an iBook last January.

In this article we will present and discuss these experimentations, trying to portray different ways to develop and interact with the museums visitors' generated content.

Results from the Europeana Taskforce

Christine Sauter, Kings College London

User generated content is a widely but controversially discussed issue, though not yet a broadly nor an always successfully applied approach. The emergence of users generated content is paralleled by the changing situation in digitally archiving and preserving heritage repositories. The ultimate goal of archiving and preservation has gone beyond storing, keeping, caring and preventing decay. Re-use and preserving objects and data for future exploitation by next-generation users is an important shift in focus.

The limitless space puts to question the necessity of selection, and this in turn threatens the role of the curator as informed expert with the closely guarded power of selection.

Both developments put the user into a new perspective, as with the Web 2.0 there is no return from user engagement and user involvement.

User generated content is as welcome and solicited as it is criticised and rejected. The Europeana Network as experts in the field of cultural heritage archiving and re-use endeavored to enquire into the state of affairs within their field and to investigate where support from the Network might be of benefit to the cultural heritage sector. To this end a taskforce was created to conduct a survey in the field and extract recommendations from the responses. This chapter will outline the results of this questionnaire.

User contributions in the Europeana 1914-18 and Europeana 1989 project

Ad Pollé, Europeana Foundation

Europeana is a catalyst for change in the world of cultural heritage. Mission: The Europeana Foundation and its Network create new ways for people to engage with their cultural history, whether it's for work, learning or pleasure. We believe in making cultural heritage openly accessible in a digital way, to promote the exchange of ideas and information. This helps us all to understand our cultural diversity better and contributes to a thriving knowledge economy. www.Europeana1914-1918.eu and www.europeana1989.eu are pan European community collection projects of digitised material across Europe. The projects aim to collect and share the family memories of people from across Europe online and make them available to the world. Europeana 1989 concerns the political and social changes in Central and Eastern Europe in the year 1989, commonly known as the fall of the Iron Curtain. Europeana 1914-1918 documents our shared European heritage, representing every side of the First World War and from every perspective.

Both initiatives aim to create a vivid and complete picture of these historical events with stories, photos, videos and sound recordings from every country affected. Personal stories, memories and experiences can help others to better understand what it was like and to see events from a different perspective. By collecting personal memorabilia and stories from this period, and combining it with institutional collections, we aim to create an engaging user experience. All items are shared online. Contributions can be made online via the project websites or at the collection days by bringing the item(s) so that project staff can digitize the items and record the stories that go with them.

Europeana 1914-1918 and Europeana 1989 offer unparalleled masses of primary source material, from which new interpretations and understanding can be drawn. The great majority of the material is openly licensed so we're inviting teachers, historians, Wikipedians, researchers, publishers, apps developers to re-use the information we provide in new and compelling ways. It is easily accessible via a search widget and API.

The projects will continue to hold digitisation days in more countries over 2014 and beyond, so that the public can go on contributing their stories and documents, adding information and helping identify the places and people in the photographs and pictures on the site. Information about future events can be found on www.europeana1914-1918.eu as well as www.europeana1989.eu.

Conclusion

Lars Wieneke (CVCE, Luxembourg), Susan Hazan (The Israel Museum, Jerusalem), Karin Hannes (KU Leuven)

This section will draw a final conclusion based on the different perspectives sketched in this chapter. This also includes a draft outlook on future challenges and opportunities.

Annex 3: Wiki Loves Public Art 2013

This annex is a response to one of the recommendations from the second year review: *“Recommendation 5 (R5): The involvement of eAwareness and the gains for Europeana from organising the Wiki Loves Public Art campaign should be further detailed, quantified and reported as an annex to D2.7.”*

The goal of the Wiki Loves Public Art (WLPA) contest was to get as many pictures of public art as possible available under a free licence on Wikimedia Commons. The submissions can now be seen and used by anyone, anywhere. These pictures will help Wikipedia to see a boost in its art coverage as photos in the contest can be added to illustrate articles in the online encyclopedia.

From October 2012-August 2013 Europeana, in coordination with Wikimedia Sverige as well as the greater Wikimedia community organised the Wiki Loves Public Art Project. The actual competition took place during May 2013.



Figure 1: Geographical Scope of Wiki Loves Art Project (Image CC BY-SA 3.0 Wikimedia Commons)

Austria	Finland	Israel	Spain	Sweden
2,409	87	311	6,085	360

Table 1: Number of Images uploaded to Wikimedia Commons by country (hyperlink directs to all images from respective country made available via Wiki Loves Public Art 2013)

The work was carried out in three stages. In the first round, each member of the jury had the opportunity to nominate images to the second and third round. The five images that received 3 or more votes in round 1 went directly to the final. The second round included all the images that received two votes in the first round and each jury member had 5 votes. The third and final round included 5 images from round 1 and 4 images from round 2. From the nine images in the final, 3 winners were chosen. The winners were awarded travel gift certificates as well as high-res printed version of the images courtesy of Europeana.

Measureable impact

All in all more than 9,250 images were uploaded as part of the contest by 225 uploaders, of which 57 percent were first time contributors. At least 25 more participated in our events in Sweden but did not upload images in time during the contest month. The reason for this is unknown.

The articles with photos from the contest have been shown a total of 1,353,909 times between May-October 2013, an average of 225651.5 times/month . Nearly half of them in May alone, when an image from the contest was highlighted on Catalan Wikipedia's main page. A significant amount of blog posts and media mentions were produced before, during and after the contest. This was by far the event that had most pieces written about it.

A number of presentations were given at events within the Wikimedia sphere about the successful outcome of the contest and what we could learn from it.

Benefits for Europeana

First and foremost, WMLA was the first large scale pan-European collaboration between Europeana and Wikimedia. It was one of several tasks project partner Wikimedia Sweden was responsible for in Europeana Awareness. Setting up the event took an estimated investment by project partners Wikimedia Sweden, Europeana and Sound and Vision of 5 person months (combined). By far, the most actual time investment came from the volunteers in the various Wikimedia Chapters involved in WMLA.

The following five areas of impact have been identified by the Awareness partners:

1. The campaign resulted in good, press coverage.
2. It proved a good example of our commitment of the Wikimedia Community & Chapters.
3. The Wikimedia community proved to be a trustworthy partner.
4. Europeana visibility on Wikimedia Commons (but not directly on Wikipedia).
5. Campaigns such as WLPA can act as Europeana's hook to local communities.

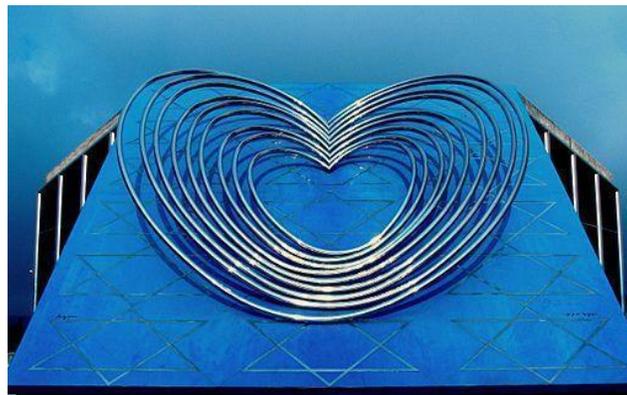
These points have been taken into account in the Chapter 'Analysis of the Wikimedia-Europeana Cooperation' that is included in D2.6: Europeana GLAM WIKI event plus report on Public Art project. This Chapter inductates



Figure 2: Wiki Loves Public Art Winner- 'Ricardo Bofill (Ricard Bofill Levi), Les quatre barres de la senyera catalana', by Coldcreation, CC-BY-SA



*Figure 3: Wiki Loves Public Art Second Runner-Up- Aino-suihkukaivo (1912) – Emil Wikström.
Sijainti: Fellmanin kartanon puisto', by Kartanofoto, CC-BY-SA*



*Figure 4: Wiki Loves Public Art Second Runner-Up- 'Yaacov Agam's sculpture A Beating Heart',
Gila Brand, CC-BY-SA*