D4.1 – Pilots Delivery Plan and Content Sourcing Strategy

Österreichische
Nationalbibliothek

Europeana Creative is coordinated
by the Austrian National Library

Deliverable

Project Acronym: Europeana Creative

Grant Agreement Number: 325120

Project Title: Europeana Creative

D4.1 – Pilots Delivery Plan and Content Sourcing Strategy

Revision: Final

Authors: Lizzy Komen (NISV), Johan Oomen (NISV), Jiří Frank (NMP), Enric Senabre (Platoniq), Vassilis Tzouvaras (NTUA), Nikki Timmermans (KL), Andrew Kitchen (EDC), Petra Newrly (MFG), Nico Kreinberger (MFG), Sanna Marttila (AALTO), Breandán Knowlton (EF), Frank Thinnes (PLURIO.NET), Jon Voss (HP), Steven Stegers (EUROCLIO), Pavel Kats (EF)

Revisions

Version	Status	Author	Date	Changes
0.1	Draft	Lizzy Komen, NISV	March 25, 2013	First draft
0.2	Draft	Jiří Frank, NMP	March 29, 2013	3.2.2
0.3	Draft	Enric Senabre, Platoniq	April 1, 2013	From 3.3 to 4.2
0.4	Draft	Vassilis Tzouvaras, NTUA	April 2, 2013	5.2.1
0.5	Draft	Nikki Timmermans, KL	April 3, 2013	5.1
0.6	Draft	Andrew Kitchen, EDC	April 4, 2013	2.4
0.7	Draft	Nico Kreinberger, MFG	April 17, 2013	6
0.8	Draft	Sanna Marttila, AALTO	April 19, 2013	3.2.5
0.9	Draft	Breandán Knowlton, EF	April 29, 2013	5.2.2
0.10	Final draft	Lizzy Komen, NISV	April 29, 2013	Summary
0.11	Final draft	Katharina Holas, ONB	April 30, 2013	Minor changes
0.12	Final draft	Lizzy Komen, NISV	May 1, 2013	Minor changes
0.13	Final draft	Katharina Holas, ONB	May 2, 2013	Minor changes
0.14	Final draft	Katharina Holas, ONB	May 3, 2013	New template, minor changes
1.0	Final	Katharina Holas, ONB	May 6, 2013	Comments by MK integrated

Europeana Creative Deliverable
D4.1 – Pilots Deliverable Plan and Content Sourcing Strategy

Distribution

Version	Date of sending	Name	Role in project
0.3	March 29, 2013	WP4 Task Leads	Pilot Leads
0.8	April 19, 2013	Andrew Kitchen, EDC Nikki Timmermans, KL	WP5 Lead, WP3 Lead, review of document
0.8	April 24, 2013	Pavel Kats, EF	Technical Lead
0.10	April 29, 2013	Katharina Holas, ONB Max Kaiser, ONB	Project Manager Project Coordinator
1.0	May 6, 2013	Thomas Jaeger, EC	Project Officer

Approval

Version	Date of approval	Name	Role in project
0.8	April 22, 2013	Nikki Timmermans, KL	WP3 Lead
0.8	April 29, 2013	Andrew Kitchen, EDC	WP5 Lead
1.0	May 6, 2013	Max Kaiser, ONB	Project Coordinator

Table of Contents

Figures

Scope

This document is an outcome of Task 4.1 – Pilots Delivery Plan and Content Sourcing.

Specifically, it aims to:

- provide an overview of the foreseen Pilot concepts (section 2);

- describe the interdependencies between the different work packages (section 3);

- define the Pilots Delivery Plan; methodologies and guidelines for successful execution of the Pilots (section 4);

- specify a content sourcing strategy (section 5);

- provide evaluation guidance; methodologies and guidelines that will be used as input for the system- and user-centred evaluation (section 6).

As this deliverable is scheduled to be delivered by project month 3, prior to the first co-creation workshops (month 4), only high-level descriptions of the Pilot concepts can be provided. Detailed planning of the Pilots will take place during and right after the co-creation workshops, by agreeing on the final concept and setting up the development schedule and backlog.

Executive Summary

The Europeana Creative project will actively encourage and promote the creative re-use of digital cultural heritage and associated metadata made available through Europeana. As part of the project, five test applications will be developed as proof of concepts which are being designed together with a number of events to spur innovation and further development by entrepreneurs from the creative industries.

The objective of WP4 is to scope, plan and implement these five Pilots in the thematic areas of Natural History Education, History Education, Tourism, Social Networks and Design.

This deliverable provides guidance and support for the successful execution of the Pilots. More specifically, it provides an overview of the foreseen Pilot concepts; it describes the interdependencies between the different work packages; it defines the Pilots Delivery Plan including methodologies and guidelines; it specifies the content sourcing strategy and it finally provides evaluation guidance.

The chosen approach for the overall planning of the Pilots is a staggered one. This means that activities for the different Pilots will not all start at the same time, but rather at different intervals during the project, namely:

- Months 4–18: Natural History Education and History Education Pilot

- Months 10–24: Social Networks and Tourism Pilot

- Months 16–30: Design Pilot

The knowledge gained from each Pilot project will be brought forward into the next, resulting in a process of continual improvement.

Each Pilot task has a development run time of fourteen months, for which the Agile Scrum method will be applied. The first delivery of each co-funded Pilot is scheduled after eight months of development, coinciding with the "Challenge" events. A further six months is allowed after the Challenge for incubation, refinements and evaluation of the Pilots.

Each Pilot will follow the same phased workflow pattern, from design and scoping into a working prototype application, as shown in the figure below.
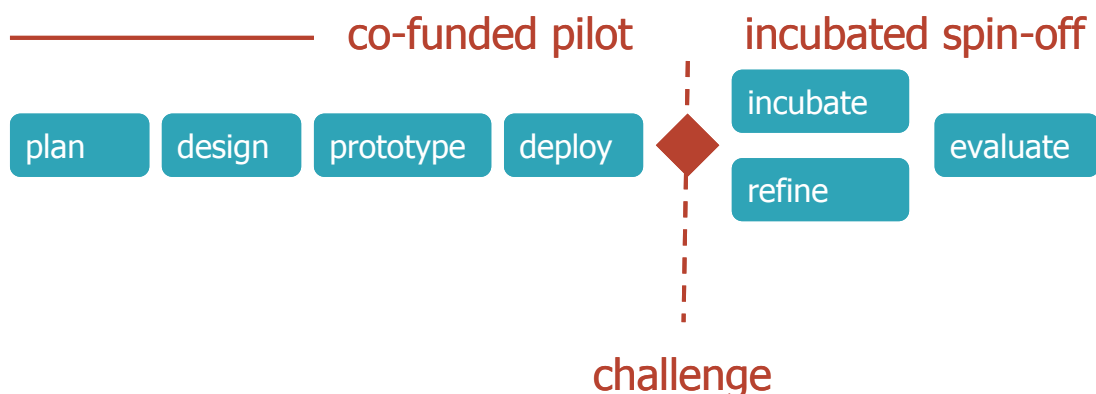
1. Introduction and Context

The main goal of the Europeana Creative project is to demonstrate that Europeana can facilitate the re-use of cultural heritage content.

The project has **eight** concrete objectives:

1. Establish the Europeana Open Laboratory Network as a sustainable environment for experimentation and stakeholder engagement.

2. Develop the Europeana Content Re-use Framework to allow content providers to make their content available for specified re-use scenarios.

3. Implement the infrastructure and services Europeana needs to support creative re-use of European cultural resources and long-term business development.

4. Create five Pilot applications in five thematic areas.

5. Conduct five open innovation challenges, in five thematic areas, to identify, incubate and spin-off five viable projects into the commercial sector.

6. Identify business models that allow key stakeholders within the Europeana ecosystem to develop their own applications and services based on the Europeana Content Re-use Framework.

7. Evaluate the results at key points in the project and measure their success against the strategic objectives.

8. Undertake an extensive stakeholder engagement campaign promoting Europeana's cultural heritage content to the creative industries and the merits of creative re-use to cultural institutions.

The development of the five Pilot applications sits at the heart of the project. The objective of WP4 is to scope, plan and implement the Pilots in the thematic areas of Natural History Education, History Education, Tourism, Social Networks and Design.

The goal of these real-world applications is twofold:

- To serve as useful and engaging applications in their own right, but also serve as a reference platform for the core services and API extensions being developed to support them. In this way, the Pilot projects become true "proofs of concept" for the development of future applications in these sectors.

- To provide the creative inspiration and case study examples for creative industries to begin developing their own applications and services that creatively re-use cultural content, through the challenge activities in WP5.

This deliverable aims to provide further guidance and support for the successful execution of the five Pilots. Input for this document has been provided by WP1, WP2, WP3, WP4, WP5 and WP6.

2. Pilots Work Plan

This chapter provides information on the overall planning of the Pilots and provides a brief overview of their foreseen concepts.

2.1 Overall Pilots Planning

The chosen approach for the planning of the Pilots is staggered. This means activities for each of the different Pilots will not all start at the same time, but rather at different intervals during the project (fig. 1). The knowledge gained from each Pilot project will be brought forward into the next, resulting in a process of continual improvement.
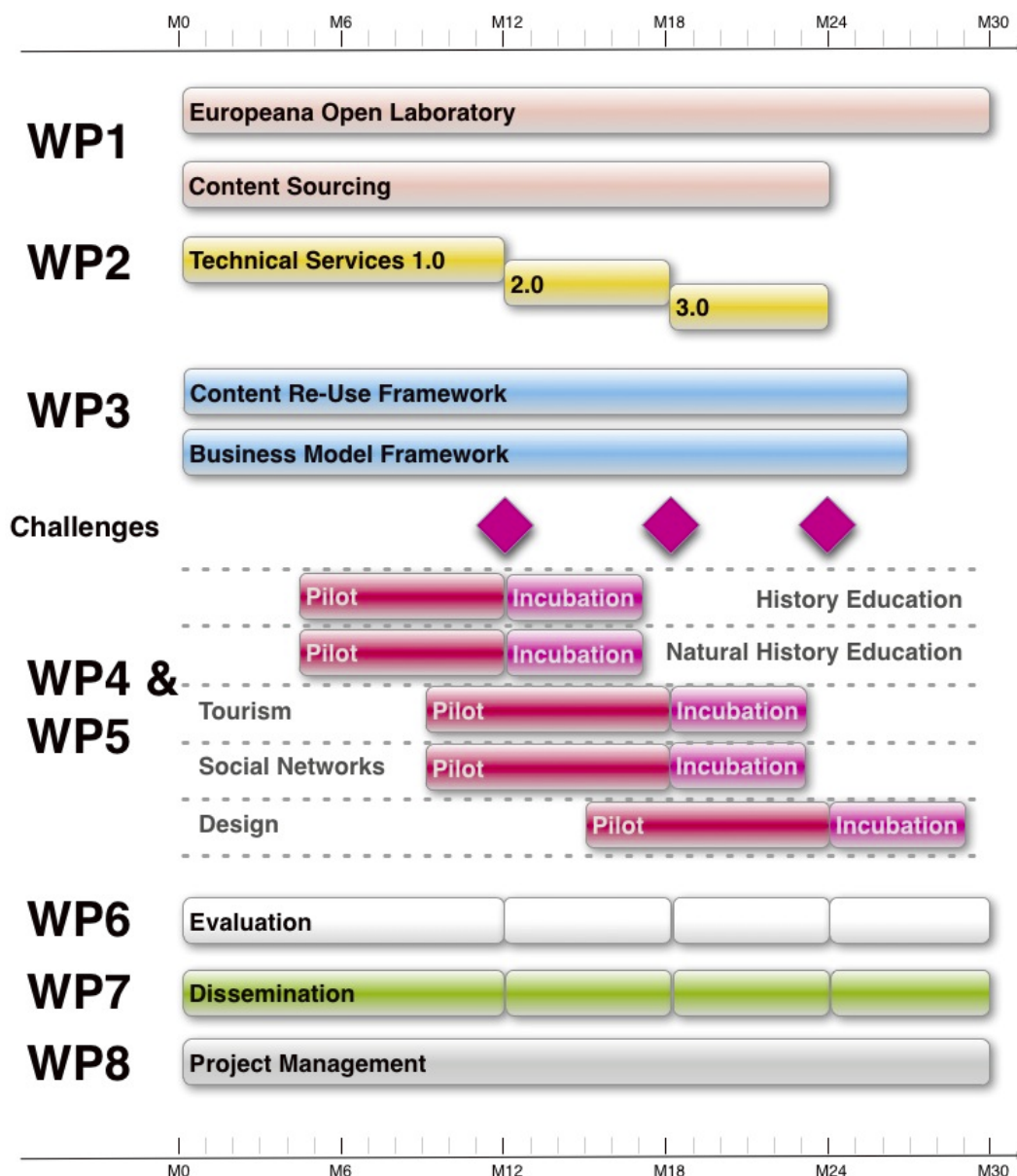

Fig. 1: Europeana Creative work plan timeline

The first activities are planned to kick off in month 4 of the project, with the start of the Natural History Education and History Education Pilots. The second phase will be the Social Networks and Tourism Pilots starting in month 10 and finally the Design Pilot starting in month 16. Each Pilot task has a development run time of fourteen months, which includes a milestone after five months of activity (M9–M13). Pilots should be able to present good progress at given milestones. The first delivery of each Pilot is scheduled after eight months of development, coinciding with the "Challenge" events. A further period of six months is allowed after the Challenge for incubation, refinements and evaluation of the Pilots.

Each Pilot will follow the same phased workflow pattern, from design and scoping into a working prototype application (fig. 2). The first phase of the Pilot can be described as the co-funded Pilot phase, followed by a challenge and an incubated spin-off phase.
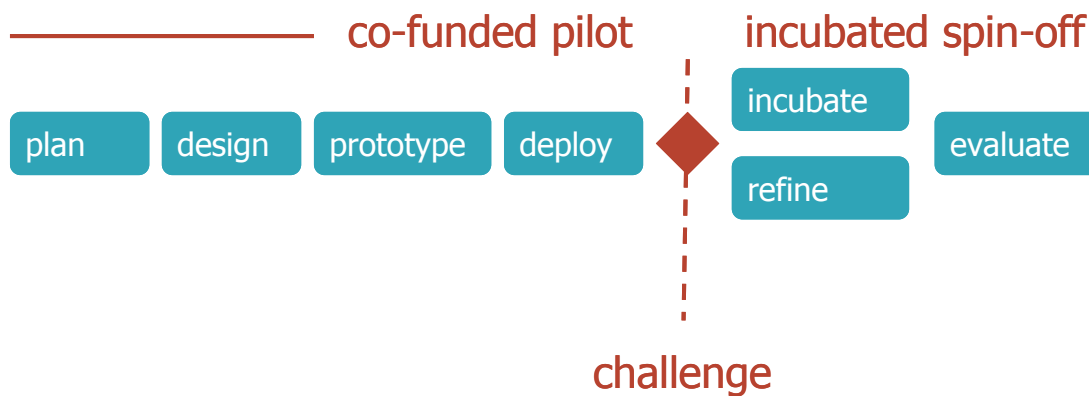
Fig. 2: Europeana Creative Pilot workflow

2.2   Pilot Concepts

The following section provides a short overview of the Pilot concepts. As the actual Pilot activities and first co-creation workshops are not planned to start before month 4, the following text presents the expected outcomes for the Pilots, as known at the time of writing. Detailed planning of each Pilot will take place during and right after each co-creation workshop and will be shared via Google Drive with the project partners.

2.2.1   History Education Pilot (M4–M18)

The History Education Pilot will enhance the Historiana website, which was developed through the Exploring European History and Heritage project[1], by extending access to and stimulating the re-use of cultural heritage content – especially content made available through Europeana – by history educators and their students.

---

[1] http://www.euroclio.eu/new/index.php/work/historiana/1105, accessed April 30, 2013.

In order to achieve this, the History Education Pilot will:

- make it easier for creators of history education content to access relevant sources;

- create ways for history educators and their students to make effective use of Europeana;

- use the Historiana website as a vehicle to access Europeana content.

Purpose

What will be piloted in the History Education Pilot?

In the Pilot the potential for re-use of cultural heritage resources (particularly those made available through Europeana) by history educators will be tested. The challenge is to develop online learning activities that will enable history educators and their students to work with digitised historical sources in ways that are more challenging and more educationally stimulating than those which are currently available to them both online and in printed resources.

The rationale behind this particular concept is that more and more learners, educators and schools have access to IT equipment, but this equipment is rarely used to its full potential (particularly in the case of history education). History websites used by educators and students to access useful background information often lack a multi-perspective approach.[2] Online learning activities are often limited to quizzing students, focused only on the memorisation of a narrow array of historical facts, whilst doing little to facilitate competence-based learning, as promoted by European and international policy makers[3], as well as history education specialists[4].

Outcomes

Which outcomes are foreseen for the History Education Pilot?

- Historiana website as an example of how to integrate Europeana services[5] in an existing website.

- Exemplary eLearning elements that can be embedded in various websites (such as the examples described in the activity section of this task).

- Descriptions of how the eLearning elements that are tested and validated by end users have been developed (to stimulate replication).

- Employment of advanced visualisation technologies to enable the creativity of professional users (i.e., history teachers in this case) to create modern and interactive eLearning presentations (e.g., time maps, serious games).

---

[2] See Recommendation Rec(2001)15 of the Committee of Ministers to member states on history teaching in twenty-first-century Europe, available online at:
https://wcd.coe.int/ViewDoc.jsp?id=234237; accessed April 30, 2013.
[3] See Key Competences for Lifelong Learning – A European Framework, available online at:
http://ec.europa.eu/dgs/education_culture/publ/pdf/ll-learning/keycomp_en.pdf; accessed April 30, 2013.
[4] See http://historicalthinking.ca, http://historicalthinkingmatters.org,
http://www.nchs.ucla.edu/Standards/historical-thinking-standards-1; accessed April 30, 2013.
[5] See Annex I, Services, to see which new functionalities are foreseen in the Europeana Creative project work plan.

Requirements

What is expected of the History Education Pilot?

The History Education Pilot is expected to demonstrate success and innovation in the re-use of Europeana's metadata and content, serving as exemplars for all stakeholders.[6] This will especially target the demonstration of creative possibilities to the creative industries and express the benefits of providing access to cultural heritage content from Europe's memory institutions. Showing non-discriminatory licensing and focused "real-world" applications that can be brought to market will demonstrate the inherent strength of multinational digitisation and federation efforts. The History Education Pilot should be documented and brought to production standard.

Access and User-Friendliness

- The website is easy to use.

- The website can be implemented Europe-wide.

- Students and educators can freely access the Historiana website.

Functionality

- Europeana Services are integrated into the Historiana website, using the Europeana API.

- Contributors (professional volunteers who develop material) can easily find and identify content that they would like to use from Europeana and import the source with all metadata (including attribution) for re-use in the Historiana environment.

- Users can interact.

- Educators can directly download (Europeana) sources that are relevant for their lessons.

- Sources are easy to find because they are presented and searchable in subject-specific ways and relate to topics that are commonly found in curricula across Europe.

Content Requirements

- The most relevant assets relating to historical figures, historical sites and specific source types (including metadata) are made available through Europeana services.

The Elements:

- are engaging for learners and enable them to interact with the content in a dynamic way;

- are easy to use by educators and help them to learn their students to acquire learning outcomes that are specific for history education;

---

[6] See Annex II for a description of the main stakeholders of the Europeana Creative project.

- make best use of learning opportunities offered by open source software and the services developed in WP2;

- are useable (or adaptable) on a variety of learning devices;

- can be integrated on (educational) websites from others;

- can work with (slightly) different content;

- can be adapted by other creative industries stakeholders.

Milestones

What are the milestones to be achieved in the History Education Pilot?

The History Education Pilot has to be delivered in two steps. A prototype has to be available in October 2013 and the release version has to be available in July 2014.

Products

Which products/outputs did we promise to deliver?

Pilot Workplan (in April 2014)

The Pilot Workplan includes a description of the History Education Pilot concept and an explanation of how this will be developed.

Pilot Prototype (in October 2013)

The History Education Pilot will develop a set of functionality and ways of presenting information that stimulate the use of Europeana content by history educators. Whereas the online learning activities and the integration of Europeana services is being prototyped online, new ways of presenting information are usually employed by using PowerPoint first.

History Education Pilot (in July 2014)

See description above.

Deliverable 4.2: Delivery of the History Education Pilot (in July 2014)

The report on the History Education Pilot includes:

- A report on requirements.

- Description of the prototype and the services used.

- Results of the Challenge and the incubation phase of the Pilots.

- An assessment of the added value for the creative industries (for instance, by analysing the ideas gathered in the Challenge).

Areas of Development

What will be the focus of the new functionalities and content created in the History Education Pilot?

The History Education Pilot will focus on those areas where new technologies can be used to offer (online) learning opportunities for history education that are better than learning resources normally used in schools.

During the first team meeting, the following areas of development were identified:

1. Analysing a variety of source types.

2. Dealing with multi-perceptivity.

3. Compare and contrast.

4. How to think historically.

5. Presenting a narrative.

Tasks Setting

What are the tasks done within the History Education Pilot to stimulate the re-use of Europeana content?

Content Selection

1. To identify which content from Europeana is most suited for re-use in the history classes and which content providers hold this content.

2. To select Europeana content that can be used for the prototype and the History Education Pilot.

Building on Practice of Others

1. To collect teacher support material that is freely available and relevant for the re-use of Europeana content.

Creation of eLearning Elements

1. To conceptualise ways of using and presenting this content so that students in history education and their educators can work with this best.

2. To make a selection of the concepts to be further developed using added value for history education (based on feedback from educators and students) and the estimation of the workload (in terms of content development, source selection and web development) as selection criteria.

3. To create prototypes of ways of presenting and using Europeana content.

4. To improve the online learning activities based on the feedback from end users.

<u>Europeana–Historiana integration</u>

1. To develop a way to integrate Europeana content in the Historiana website using the Europeana API.

2. To ensure that the new functionalities and the new content can make use of Historiana as vehicle for Europeana content.

3. To integrate the Europeana content in the Historiana website, refer to existing teacher support material and offer new learning opportunities to interact with the Europeana content online and offline.

<u>Testing and Quality Assurance</u>

1. To present these prototypes to get feedback from end users through a network of Pilot schools.

2. To regularly gather feedback from end users during gatherings with history educators.

2.2.2 Natural History Education Pilot (M4–M18)

According to the Description of Work the aim of this Pilot is to demonstrate two examples of the effective utilisation of Europeana content in the natural history education domain. The objective is to show that those two examples could be a good inspiration for the creative industries, which will also be tested in the Challenge phase. The Pilot team comprises a very effective consortium of partners, which include content providers (NMP and MfN), technical developers as well as creative industries representatives (XZT and SEM).

Content

It is foreseen that the Pilot outcome represents two products/examples: an adventure game and an application. Both outcomes will use the content provided by NMP and MfN to Europeana via the OpenUp! project[7] but also other collections which are not on Europeana yet, but will be uploaded during the product development phase. This content will represent the highlights of the collection and suitable specimens/objects for the game and application. Beside the multimedia content such as images, sounds and 3-D models, the content providers will provide object metadata and extended information according to the game and application needs. This data will also enrich the existing Europeana data. There is the possibility to enrich the content by literature or illustrations provided by the BHL and BHL-Europe project[8], which are already available via Europeana.

Work Plan

According to the Description of Work the development of the Pilot applications will start in month 4. In month 9 is the first milestone to show the progress of the prototype. The prototype will be most probably the alpha version of the game and application demonstrating the basic functional system. The Challenge event is in month 12 which is an unofficial milestone for the beta version of the Pilot products. The Pilot products will be an inspiration for the Challenge and will be finalised as release versions during the incubation/refinement phase. In month 18 the final deliverable for the Pilot, D4.3 (Delivery of the Natural History Education Pilot), is due which will include the documentation and the final version of the game and application and also an overview of the Challenge and incubation phase.

Pilot Applications

In anticipation of the first co-creation workshop in month 4, the foreseen game and application at the moment of writing are:

---

[7] http://open-up.eu/; accessed April 30, 2013.
[8] http://www.biodiversitylibrary.org/, http://www.bhl-europe.eu/; accessed April 30, 2013.

Pilot Application 1: The Natural History Education Application

The concept for this application is foreseen to have effective impact on the usage and attraction of creative industries as well as a prominent business plan. The application will be designed for mobile devices, such as tablets and smart phones, using the iOS and Android systems. There are two different target audiences. The first audience are direct application users which will be primary and secondary school children and young museum visitors such as families. The second audience are institutions such as natural history museums, for example, which will re-use the application structure by filling it with their own content and by using it as some kind of interactive guide.

The basic concept of the application is to provide an educative tour back in time, including information on time periods, geology, geography, important events, fauna and flora. Special attention will be given to extinct fauna and flora, which will be demonstrated through the use of collection specimens from NMP and MfN that are also on Europeana; a connection with the Europeana portal will be provided. The information in the application will be structured in several levels of details to be open to use for children of different age categories. The basic application elements will be the time spiral, a timeline with details and specimens, specimen info boards and a personal notepad/wiki which will be filled with detailed information according to how the user will explore the application. To improve the usage of the application by children and also to encourage them to use the application several times, the achievement and experience system will be implemented including variable tasks according to the application content and information. The application will already include most of the information and content and will be able to be used also in offline mode, but the online mode will provide the information extension and also a connection to the Europeana portal. The application will be developed by SEM.

Pilot Application 2: An Adventure Game for Education in Natural History

The working title for the game is "Night at the Museum". It will be a mix of a point-and-click adventure and a hidden object game. The user will take the role of an explorer who has to solve a lot of puzzles to unveil a big secret. The game has to be played in landscape format; the main character will not be visible in the game (first-person view). The game will also have an educational factor. The user will learn a lot about history, science, evolution, etc. while solving puzzles.

The game will use MfN and NMP content and the environment of the institutions' buildings. The game's purpose will be to explore the exhibition rooms but also the depositories, basements and variable spaces, which are usually hidden to the museum visitors.

The game is planned for tablets such as iPad with iOS system, including 2-D and 3-D elements and using the Unity engine. The target audience is 12+ due to the mysterious and in some moments scary atmosphere. The game can be played online and also in offline mode.

<u>The Story in a Nutshell</u>

The user will take the role of Vincent La Fleur, the son of the very famous explorer Andrew La Fleur. He receives a picture of a part of a very old map which shows the way to a hidden place which keeps a big secret. The missing parts of the map are hidden in the Museum für Naturkunde (MfN) in Berlin and the National Museum (NMP) in Prague.

Fig. 3, Fig. 4: Concepts for the adventure game

Setting

The locations are the Museum für Naturkunde in Berlin and the National Museum in Prague. It is always deep in the night when the user enters a new museum. The locations in combination with little light create a thrilling and mysterious atmosphere.

To make the 2-D scenes more lively, sound, music, animations and particle effects like fire, volume, light, etc. will be added to the scenes.

Puzzles

Some, but not all puzzles will require either natural history knowledge or working with the Europeana database (or a mock-up of it). The audience will learn about topics like taxonomy, evolution and geologic periods. A secondary goal is to encourage users to explore the museums in person on the trail of our protagonist Vincent La Fleur.

2.2.3   Tourism Pilot (M10–M24)

The Tourism Pilot is based on the idea to explore options for new presentations of cultural heritage material for touristic targets.

Cooperating with the tourism and publishing sectors for many years, project partners PLURIO.NET and Culture24 know that it is **crucial to understand the changing needs** of the online tourist in order to be successful. This is a difficult task as online user behaviour is never static and the **online tourism sector is asking for new services and tools all the time**. PLURIO.NET and youARhere have already identified a very strong demand from their clients and project partners for the use of the envisaged solution for **World War I commemoration activities in France, Belgium, Luxembourg and Germany, World War II (2015!) activities in Belgium and Luxembourg (battle of the Ardennes), as well as for the presentation of the heritage of Mons (2015) and other cities, i. e., Luxembourg, Brussels, etc.**

Hence, the Pilot partners will start their activities with a **scoping phase**, in order **to provide meaningful and viable information for a successful co-creation workshop (M10) and the designing phase**. They will focus on better understanding the end users'/personas'/stakeholders' needs (leading to the list of users/operators to be invited to workshop), on identifying meaningful content (to be used in the workshop), and they will map already existing solutions in order to promote a truly innovative approach for the co-creation workshop and the design of the prototype. These scoping activities will not predict the results of the workshop, but, rather on the contrary, help to make it truly innovative, successful and in sync with the needs of the targets and the overall objectives of the Europeana Creative project.

The envisaged solution is an **augmented reality application for mobile phones and/or tablets** that allows visitors to explore the way places looked in the past, discovering cultural objects in their original settings, diving into augmented reality cityscapes, seeing (and hearing) historical events where they happened or "seeing through walls" of museums and archives along their itinerary. This should be achieved by creating a specific tool for creative industries that allows touristic operators and travellers to tap easily into Europeana and other open information databases such as Wikipedia (**"Cultural Travel Scrapbook"**).

Yet, the Pilot partners are open and aware that the co-creation workshop could possibly provide another, even more innovative and creative solution. It's the workshop alone that will validate and **deliver the specifications** for the Pilot's prototype which will then be deployed and tested by the identified audiences, leading to the Challenge (M18), the incubation phase and finally the delivery of the Pilot (M24).

2.2.4 Social Networks Pilot (M10–M24)

The objective of this theme is to attract suppliers of websites with social elements to embed Europeana content in their services, and to create an engagement model of mutual interest to both memory organisations and users. Through digitisation, memory organisations and their users are now beginning to inhabit the same, shared information space, and new services are being launched that explore this fundamentally new paradigm of participation in the cultural heritage domain.

Furthermore, memory organisations around the globe are beginning to explore the potential of crowdsourcing, i.e., outsourcing specific activities to a community through an open call. The impact of user engagement practices in the GLAM domain (galleries, libraries, archives, museums) on work processes is becoming profound. One of the key success factors of these practices is shaping and executing them in a way so that both the users as well as the institutions find them beneficial. Johan Oomen and Lora Aroyo have clustered different crowdsourcing models and created the following classification of six types of crowdsourcing (see table below).[9] The Social Networks Pilot belongs to the category "Complementing Collections".

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.

Fig. 5: Crowdsourcing types

Approach

The Social Networks Pilot will extend the experiment of crowdsourcing and enable content holders to better reach audiences through social networks.

- Firstly, the project will enable Europeana metadata to be used by location-based services, whilst increasing the visibility of digitised content in popular location-based search.

- Secondly, the project will support national crowdsourcing initiatives, enabling them to collaborate with a transnational service.

---

[9] See Johan Oomen and Lora Aroyo, "Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges", in: *Proceedings of the 5th Communities and Technologies Conference*, Brisbane, Australia, June 2011.

Content and Platforms

The British Library (BL) and the Netherlands Institute for Sound and Vision (NISV) will supply a significant amount of high-quality sounds with geo-locations to Europeana. These sounds have been geo-tagged as part of national crowdsourcing campaigns – UK Soundmap[10] and Geluid van Nederland [The Sound and the Netherlands][11] – and are made available under conditions that allow re-use within the Content Re-use Framework.

Geluid van Nederland makes a sound archive from the NISV collection available to a wide audience. The sounds are available through a "sound map" (fig. 6). This interface provides an overview of the national coverage of the sound archive, both in space and time, and is used to encourage users to add their own sound (either via web-based upload or the Sound Hunter app) to complement the archive.[12] SoundCloud [2] is used as the underlying technical infrastructure.

Fig. 6: Geluid van Nederland sound map

The UK Soundmap (fig. 7) was launched in July 2010, asking people to record the sounds of their environment, be it at home, work or play. Since then, over 2,000 recordings have been uploaded by some 350 contributors.[13] The British Library makes use of the external platform AudioBoo[14] for sound contributions.

---

[10] http://sounds.bl.uk/Sound-Maps/UK-Soundmap; accessed April 30, 2013.

[11] http://www.geluidvannederland.nl/; accessed April 30, 2013.

[12] See Oomen 2013.

[13] See http://britishlibrary.typepad.co.uk/archival_sounds/uk-soundmap/; accessed April 30, 2013.

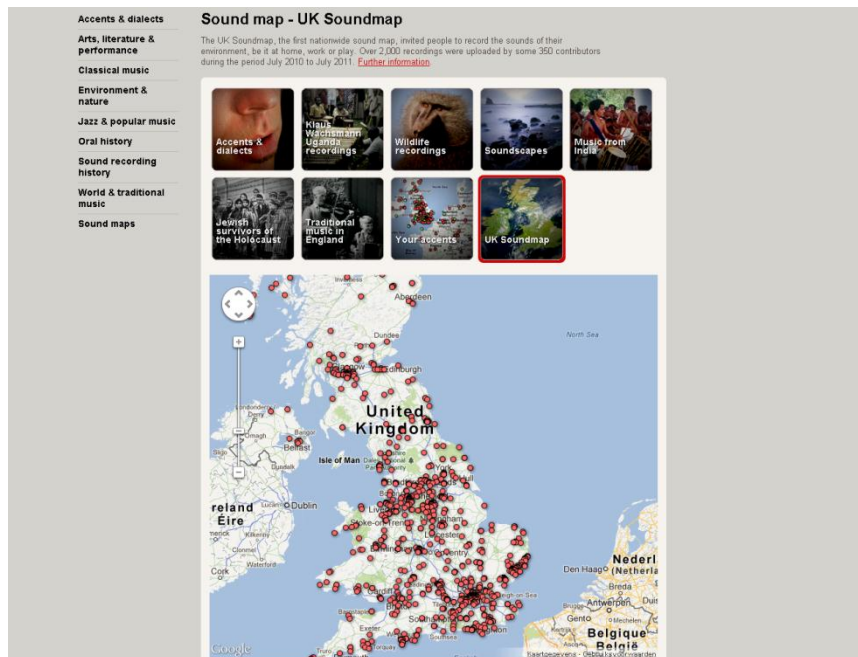[14] http://audioboo.fm/; accessed April 30, 2013.

Fig. 7: UK Soundmap

In addition to the sound material, a large body of user-contributed content is gathered as part of specific campaigns within the Europeana Awareness project[15]. Specifically, this is content from the projects:

- Europeana 1914–1918[16]
- Europeana 1989[17]

The Europeana 1914–1918 project collects memorabilia and stories from the period of World War I. The backend for this project is currently utilising RunCoCo.

Europeana 1989 is a pan-European community collection of digitised material concerning the political and social changes in Eastern Europe around 1989. This project is in coordination with Historypin (HP) and is utilising the HP backend.

NISV and BL will, in collaboration with Historypin and Europeana, specify a sound workflow for content delivery between all platforms involved. Furthermore, WP1 will conduct an inventory of the available content from the content providers for this Pilot and identify other (potential) content-providing institutions in addition (see chapter 5).

---

[15] http://www.pro.europeana.eu/web/europeana-awareness; accessed April 30, 2013.
[16] http://www.europeana1914-1918.eu/; accessed April 30, 2013.
[17] http://www.europeana1989.eu/; accessed April 30, 2013.
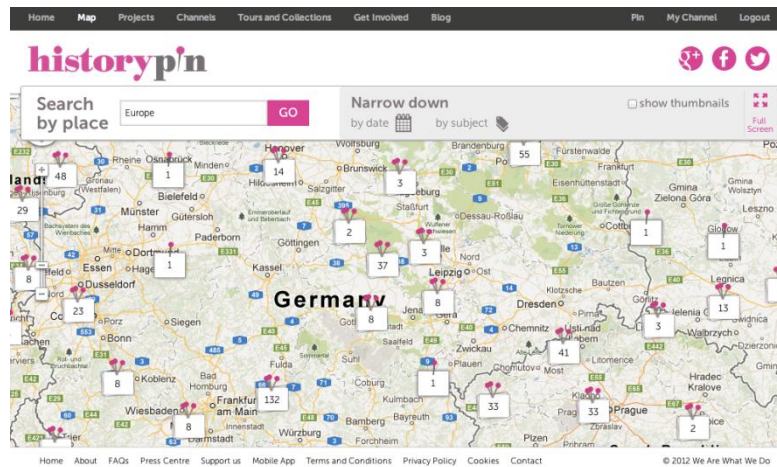
Fig. 8: Historypin

Historypin will ingest selected metadata about the audio recordings from BL and NISV as part of the sound workflow in the HP backend. This metadata will include the URLs for the hosted audio files which will be embedded within the user interface of the HP frontend, where the audio content will also be highlighted as audio.

Ontotext (ONTO) will provide services for geo-referencing of the data contributed to this Pilot based on entity extraction.

Expected Deployment of this Pilot

While the end result and thrust of this Pilot is a shared interactive and social experience of historical audio content, the preceding development of the workflow pertaining to audio content that is compatible with the Europeana Data Model (EDM) is critical, and where we will interface significantly with WP2.

2.2.5 Design Pilot (M16–M30)

The Design Pilot connects the cultural heritage offered by Europeana to the open design communities of craft and media designers who wish to use cultural heritage objects as sources for new, derivative designs, such as embroidery, textile patterns, 3-D printed objects, media art, etc., and are also interested in sharing their designs back to the community in re-usable forms.

The Design Pilot will amplify the usefulness of Europeana for designers and the creative industries that are connected to such design activities; it will sponsor the revival of cultural heritage objects, reborn as elements in contemporary designs, and aggregate this new, emerging cultural heritage into collections in a format that could in the future become available through Europeana.

The Design Pilot will be run within the Open Lab situated in the Aalto Media Factory in Helsinki, and utilise its Fab Lab and other digital media and design resources. The Design Pilot is conducted by Aalto University (AALTO), AIT Austrian Institute of Technology (AIT) and Spild af Tid (SAT).The core objective of the Design Pilot is to understand the creative process of producing a design or an art artefact that utilises and re-uses existing open cultural heritage resources. How can the elements of the open creative process, such as searching relevant content, sharing both artefacts and good practices (e.g., attribution mechanisms, developing sharable design building blocks) back to the cultural commons be supported by digital tools and co-developing social practices? The Design Pilot is carried out in close collaboration both with creative communities (practitioners, designers and artists) and the creative industries, leading in the tradition of the Participatory Design (PD) approach.

The Design Pilot will work around three thematic concepts related to the creative process using existing open cultural goods and materials available: *interpretation*, *appropriation* and *reinvention*. Two services will be developed for use in this Pilot by AIT based on the collaborative creation and design activities:

1. A new advanced search mechanism to support various types of visual search among its visual contents (e.g., color, shape, etc.).

2. A new tool that manages the appropriate crediting and transformation/derivative history by maintaining links between originals and their derivatives.

2.3   Planned Workshops and User Testing Schedule

Co-creation workshops will be based on the following sequence, adapted to venue and schedule requirements when needed:

1) Context Walkshop: (1:30h)

- Content scouting
- More context and thinking "out of the box"

2) Eliciting Exercise (1:00h)

- End users and personas
- Content to aggregate
- Learning/teaching opportunities
- Exploring possibilities and needs

3) Mapping and Evaluating (2:30h)

- Connecting users, actions and content
- Potential ideas and processes for piloting
- Rapid evaluation session
- Narrowing and evaluation of options

4) Flow Design and Rapid Prototyping (2:30h)

- Second layer of details to the process (2-D or 3-D)
- Online features and content as well as offline aspects (agents, places of use, abstract concepts, etc.).
- Define a first version of the minimum viable product

5) Presentations (1:30h)

- What's optimum? From there, what can others add or do?
- Prototypes/mock-ups/diagrams
- Public Q&A session
- Feedback for starting development

6) Implementation Meetings (1:00h)

- First definition and prioritisation of initial backlog

- Linkages in each project and dependencies

- Deployment and evaluation issues

- Planning of next steps and implementation as needed

3. Interdependencies

One of the key activities of Task 4.1 is to manage the interdependencies between WP4 and the other work packages, specifically WP1, WP2, WP5 and WP6. Interdependencies with WP1 and WP2 relate to the availability of the Open Labs. Interdependencies with WP5 relate to releasing services to be used as a basis for the competitions and providing the baseline for the incubation phase of the work plan. Finally, WP6 depends on WP4 to be able to perform system- and user-centred evaluation. The diagram below shows the dependencies between the Europeana Creative work packages.
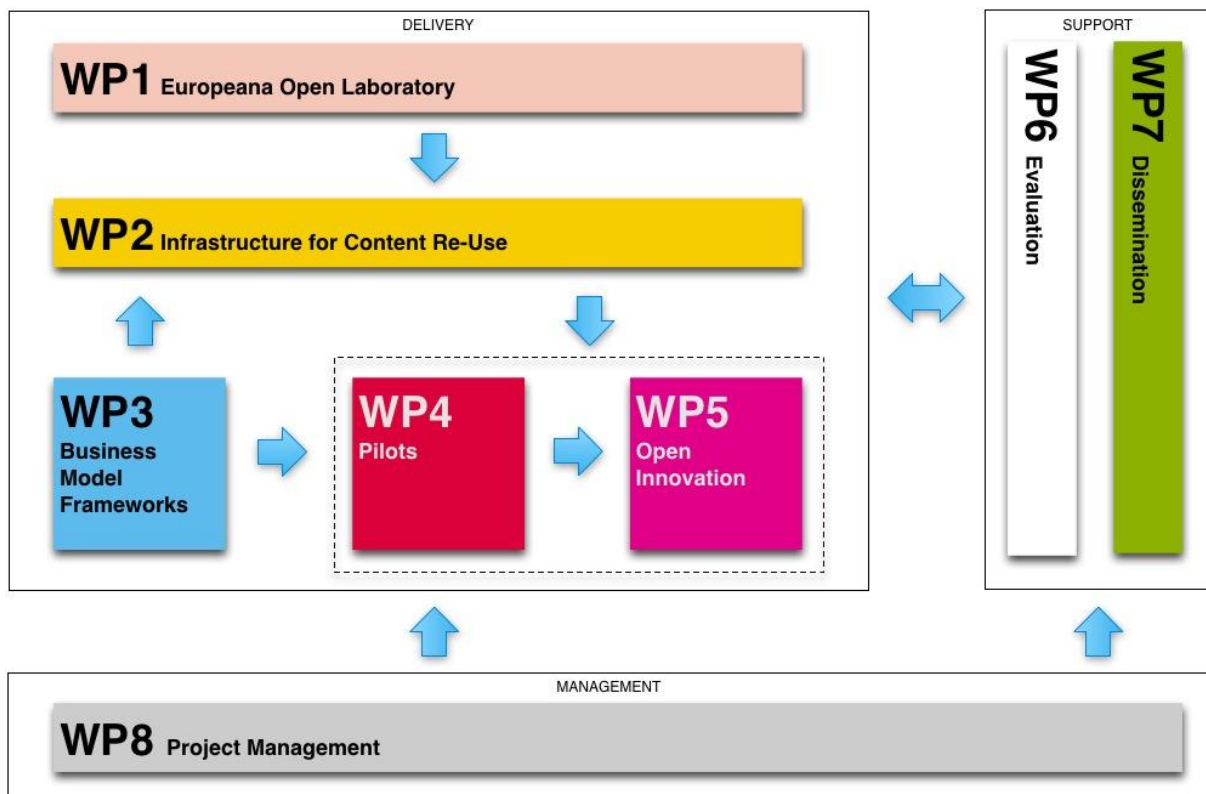

Fig. 9: Interdependencies between work packages in Europeana Creative

3.1 Interdependencies WP1 and WP2

The Pilots will depend on the input of WP1 for the scheduling and content of the workshops and the programming of activities. This will be communicated by the workshop leader (Platoniq) to the WP4 leader (NISV).

The Pilots depend on the input of WP1 for content sourcing. The responsible person at Europeana Foundation (EF) will coordinate the sourcing of content needed for each Pilot and communicate this to the WP4 leader (NISV).

The Pilots depend on the input of WP1 to define the detailed Pilots Delivery Plan for each Pilot, covered in chapter 4 of this document. The co-creation workshop leader (Platoniq) will facilitate a discussion about implementation as one of the concluding stages of each co-creation workshop. This delivery plan will then be further elaborated in writing by the Pilot partners, and the plan will be implemented by the WP4 leader (NISV).

The Pilots depend on WP1 and WP2 for the availability of the Open Labs. In the first instance the physical lab spaces need to be available for the Pilots' activities, and to serve as hubs for the provision of incubation services. During the co-creation workshop each lab is the place for the different activities planned, providing the basic technical infrastructure for its development. Afterwards, it can be also a "physical hub" with information displays showing the work done during the workshop and to showcase selected ideas that had been identified for potential development in specific areas. The main dependency with the Pilots development at that stage would be to provide more on-demand information, if needed, to potential individual or institutional partners about the stage of the Pilot.

In the second instance, the online tools (from OpenCultureLab.eu) will need to be developed and made available in a public state before the Pilot themes are presented to the public. The WP1 leader (EF) will make sure that these activities are staged appropriately for each Pilot theme, and make sure that the various tools are in a public-ready state in time to support the development and deployment of each theme. In this way, the online tools will enable the availability of content for the development of Pilots, provide assistance when needed for its re-use and implementation, whilst showcasing the Pilots when they are released.

WP2 will provide the technical services needed to support the re-use of Europeana cultural resources. These services will provide granular functionality enabling the best rapid application development within the Open Labs environment, and shall directly support the development of the five themed Pilots and Challenge spin-off projects within WP4 and WP5. To support the Open Culture Lab environment of WP1, WP2 will develop a working online console to experiment with the use of APIs and sample data sets. The use of WP4's Pilot projects as a basis for requirements and development of these technical infrastructure services will ensure a "real-world" approach to development that can be made immediately useful.

3.2 Interdependencies WP5

Interdependencies with WP5 relate to releasing services to be used as basis for the competitions and providing the baseline for the incubation phase of the work plan.

At a high level of interpretation both WP4 and WP5 are targeting the delivery of "impact 4", noted in the Europeana Creative Description of Work as follows:

"Inspire and support creative industries to re-use Europe's cultural heritage."

In order to support the delivery of this impact, WP5 shall run a series of Challenges for the creative industries to embrace creative and innovative re-use of cultural heritage resources through utilisation of the technical infrastructure produced by the project. The Challenges shall take the form of online events, offline events (including a hackathon) and various promotional activities. Each Challenge will result in a "winner" being identified as the most viable for wider market entry – termed as a *spin-off project*. Each of the five spin-off projects that emerge from the Challenges shall be provided with hands-on incubation and support to improve their chances of success. Additional incubation and support guidance materials will be made available to any other interested parties, whether or not they entered a Challenge, to help them in the process of creating viable and sustainable businesses.

The delivery of impact 4 will depend upon many factors though fundamentally, according to its wording, the Pilots will provide the bulk of inspiration and the Challenges the support to the creative industries. Therefore WP5 is quite dependent upon WP4 providing the inspiration for the creative industries.

It is expected that each Pilot project will be based on the subject-specific skills of sector experts and industry stakeholders, following each of the five common themes, in building real applications or services that show the commercial potential for creative re-use of digital cultural heritage material. These applications or services must utilise and exemplify the technical infrastructure provided by the project. This is vitally important for the Challenges, as the Pilots need to be able to actively demonstrate use of the extended APIs, the available content and how the Content Re-use Framework has been implemented. This demonstration will be of the greatest importance during each Challenge for two reasons. Firstly, there will be online events where each Pilot project can be presented to the audience with a question and answer session. Secondly, there will be an offline event where each Pilot project can be presented to the audience, again with a question and answer session, and where Pilot project staff will be expected to actively participate in 1:1 or group-based sessions at the hackathon.

The goal of these activities will be to help potential Challenge entrants develop their understanding of the API, the wealth and breadth of content made available by it and the license framework – and provide advice and guidance concerning the Pilot project's use of it. Providing inspiration and encouragement to the potential Challenge entrants will be key elements in this process, helping them to innovate and begin developing their own applications and services that creatively re-use cultural content, thus meeting the expected Europeana Creative impact.

3.3 Interdependencies WP6

The interdependencies between WP4 and WP6 Evaluation are mainly based on the following tasks, deliverables and milestones:

Task 4.1: Pilots Delivery Plan and Content Sourcing

Within this task clear guidance will be provided on the key milestones and the concrete methodologies to realise the Pilots. This guidance will have concrete influence on the system and user-centred evaluation to be executed within Task 6.2, Pilot and infrastructure testing and evaluation, which will be led by Platoniq. The Pilot's infrastructure testing will start in project

month 4 and will last until the project's end. Furthermore the guidance and Pilot development structure will also be an important element to set up the overall evaluation strategy and framework (D6.1)

Tasks 4.2–4.6: Pilots Development

WP6 has also close synergies with the overall Pilots development as their development process, outputs and impacts will be evaluated by WP6 (MS19, MS21). This will mainly be realised by Task 6.2 and Task 6.3 and generally by an evaluation group set up within WP6.

The evaluation results will all be summarised in a Pilot and infrastructure evaluation report (D6.3, M29) which will again highlight the Pilots' usability and acceptance.

Further details for the evaluation tasks are described in chapter 6 of this deliverable.

3.4   Linkage between WP1, WP2 and WP5

WP5 is dependent upon WP1 and WP2 in a similar way that WP4 is dependent upon them. The goal of WP5 is to spin-off five projects into the commercial sector, each of which will use Europeana Creative's content and technical infrastructure provided by WP1 and WP2. Essentially this is the same dependency that each of WP4's Pilot projects have with WP1 and WP2, though in each case the Pilot projects will require content and technical services at an earlier stage in the delivery of Europeana Creative. The key difference between the relationship of WP4 and WP5 with WP1 and WP2 is that WP5 will not be involved in the specification and formulation of the infrastructure and content provided by it. By contrast WP4 will be intrinsically involved in this process.

4. Pilots Delivery Plan

4.1 General Methodology

For the general methodology the following core principles will apply:

1) Co-creation workshops at the beginning of every Pilot's conception:

- Reorder, iterate and improve with modularity.

- Adapt from scratch or for previous developments.

- Expand or fork for Challenges and Open Labs.

- Focus on deliverables (diagrams, stories, mock-ups, etc.).

- Integrate observation, documentation and evaluation.

- Intersections between offline, digital and online.

Co-creation workshop methods > eCreative Pilots

Fig. 10: Co-creation workshop methods

The general methodology for the Pilots is described in a presentation which forms Annex IV to this document.

2) Adapting agile methodologies from Scrum during the development and evaluation of each Pilot:[18]

- The highest priority is the early and continuous delivery of valuable software.

- Welcome changing requirements, even late in the development process.

- Agile processes harness change for the Pilot's competitive advantage.

- Deliver working software frequently, with a preference for the shorter timescale.

- Partners and developers must work together regularly throughout the project.

- Build Pilots around motivated individuals. Give them the environment and support they need, and trust them to get the job done.

- Although the most efficient and effective method of conveying information to and within a development team is face-to-face conversation, we will adapt to online channels.

- Working software is the primary measure of progress.

- We should promote sustainable development.

- The partners, developers and users should be able to maintain a constant pace indefinitely.

- Continuous attention to technical excellence and good design enhances agility.

- Simplicity – the art of maximizing the amount of work not done – is essential.

- The best architectures, requirements and designs emerge from self-organising teams.

- At regular intervals, the team reflects on how to become more effective, then tunes and adjusts its behaviour accordingly.

4.2 Guidelines

The general approach for the delivery of the Pilots will be defined immediately after each co-creation workshop, then refined by the Pilot partners. The strategy is to begin with a list of general questions about each Pilot. These questions will be considered as part of the implementation phase of each co-creation workshop. A rough idea of the solution to each of these questions will be created on-site, then elaborated in writing and turned into a detailed delivery plan by each sub-group working on developing and communicating the thematic Pilot.

sHow will these plans relate to the co-creation workshops?

At the conclusion of each workshop the task leader (Platoniq) will conduct a structured session to answer questions in the following areas, based on the feedback of the workshop participants.

Each Pilot delivery plan will include:

---

[18] Adapted from http://morepm.com/tag/scrum-principles/; accessed May 2, 2013.

4.2.1 Content Strand

Questions of content needed for the Pilots Delivery Plan will be broken into two areas: input and output.

On the input side, the following questions need to be considered:

- What content is needed for this Pilot in its initial form? Where will this content be accessed from when the Pilot is in production?

- Where can we get the content that is needed? (This is covered to a large extent in Task 1.4, Sourcing Content for Pilots. But the Pilots Delivery Plan will need to include a brief inventory of collections needed, including the need for local or distributed storage of content assets.)

- As the Pilot is deployed and continues to evolve, what other sources of content might need to be made available to the Pilot application?

- Which license issues arise? How to address them?

On the output side, the following questions will need to be considered, for each Pilot:

- Will any content be explicitly created or generated by the Pilot application? This could include user comments, annotations, metadata enhancements or corrections, photos.

- Will any data be implicitly created by the Pilot? For example, usage data, tracking of user activity, account information, log-in history, audits, geographical tracking, etc.?

- Where will this information be stored, and how might it be accessed? Which institution (internal or external to the consortium) will be responsible for the ongoing maintenance of this information? Will this information be accessible through the Pilot itself (thus becoming an input)?

- Will there be in the case of activities/actions in each Pilot (for example, educational Pilots, but it can be applied to the rest of them) some type of evaluation, badges or activity streams related to users' behaviour?

- Will any content be re-used during the user experience in the application?

In relation to the output side, the following should also be considered: Europeana has terms for user-contributed content (that is content and metadata not contributed by data providers who have signed the DEA) that ensure that such content can be used within Europeana and by the users of Europeana.[19] These terms need to be accepted by end users / Pilot users who end up creating content/metadata that goes back to Europeana.

---

[19] See Europeana Terms for User Contributions, available online at:
http://www.europeana.eu/portal/rights/terms-for-user-contributions.html; accessed May 2, 2013.

4.2.2 Community Strand

Delivery of a Pilot doesn't exist in a vacuum – we will need to consider the needs of users and audiences to make sure that we are deploying in a way that can be used.

- What is the audience for this Pilot? Is there a standard persona that we can define, or perhaps a job or demographic that describes the community of interest for this Pilot?

- Do we give access to collections / source material to the community while we are developing Pilots through the Open Lab website?

- Do we invite the community to work with similar content?

- What's the community segmentation (developers, final users, students, educators) of the Pilot?

4.2.3 Publicity and Dissemination Strand

There are two kinds of publicity and dissemination to be done for each Pilot: In the first case we need to communicate with the end users of the Pilot – the communities who will value it and what it does. In the second case, we need to communicate with those who would be interested in developing the ideas for this content further. In that second category, we might also distinguish between the software and technical developers who might want to continue expanding on the ideas represented by the Pilot, and the content holders or memory institutions who might be interested in continuing to contribute content or ideas related to the Pilot theme. Each of these sub-communities will need to be approached somewhat differently.

- How will we talk about the Pilots?

- When will we disseminate things? In which stage of development?

- In which format(s)?

  - Designs, mock-ups, diagrams, prototypes and other "raw" outputs or documentation from the workshops (as work in progress materials).

  - Qualitative information like detailed posts or descriptions or interviews with end users, partners or developers.

  - Access to functional versions of the Pilot in repositories, along with instructions for use or test.

- Who can do it and from which perspective? (Cultural, technical, educational, strategic, etc.)

4.2.4 Access Strand

How will we be able to access this Pilot? Does it exist:

- Online as a website?

- Installed in a place (museum, etc.)?

- Classroom experiment happening in a school?

- As a demonstration in an Open Lab?

- As an online functional mock-up?

4.2.5 Testing Strand

How will we check things? Think about:

- How will we test this Pilot?

- Who should do it?

- Should the labs organise demos and testing sessions during the development phase?

- How should conclusions from the testing be shared?

- Which features are critical to test more times or deeply?

4.2.6 Evaluation Strand

Does the Pilot achieve things and meet its objectives? How do we measure it? Checkpoints related to these areas:

- Sustainability/business opportunity

- What's technically doable

- What's legally doable

- Level of engagement

- Usability

- Level of appropriation for other partners at local levels

- Specific values (depending on the theme):

    - Learning and teaching

    - Playability values

    - Storytelling

    - Design

4.2.7 Sustainability Strand

How will the Pilots be sustained in the future, with regard to their finance? Asking questions about:

- Is any there any formula, such as freemium, pay per use, updates or inclusion in other services?

- Are there crowdfunding or other distributed mechanisms that need to be in place for this Pilot to succeed?

- How can they relate to the plans or other partners' sustainability strategies?

4.2.8   Adaptation Strand

We would expect the Open Labs to consider these questions, based on their knowledge, networks and expertise:

- How will this Pilot inspire the external incubated projects for them to possibly re-use parts?

- How can this Pilot be localised? Can it work in other languages or places?

- Can it work with other collections of content?

- Does it depend on any infrastructure that only exists during the lifetime of the project?

4.2.9   Development Strand

Sprints will be four weeks long, beginning at the co-creation workshop, and there will be about eight or nine sprints per Pilot. Themes will be staggered by two weeks so that demos / sprint starts do not overlap.

Each sprint will include in the first week:

1. Demo of previous sprint (product owner and developers).
2. Retrospective (with WP6).
3. (Re)Prioritise backlog (product owner).
4. Addition of any user testing bugs/problems to the backlog.
5. Define focus for next sprint (which stories will be covered).

Each sprint will include, in the final week, a demo to all consortium members of progress in each area. This demo will focus on:

- Content

- Code

- Deployment

- Evaluation checklist

- Dissemination

- Documentation

- Any other areas need

Main roles during development:

- Each Pilot Lead as product owner. E.g., Theme 1 (History) is EUROCLIO (Steven Stegers), Theme 2 (Natural History Education) is NMP (Jiří Frank), Theme 3 (Tourism) is PLURIO.NET (Frank Thinnes), Theme 4 (Social Networks) is NISV (Lizzy Komen, Johan Oomen), Theme 5 (Design) is AALTO (Sanna Marttila).

- WP1 (Platoniq) is Scrum Master for all themes.

- NISV will provide guidance as WP4 Lead.

An outline of the suggested Scrum adaptation for the Europeana Creative Pilots can be found in Annex III.

5. Content Sourcing Strategy

5.1 Content Re-use Framework

The Content Re-use Framework, developed within Europeana Creative (WP3), builds on the current Europeana Licensing Framework that was developed within the EuropeanaConnect project[20] and is currently being refined as part of the Europeana Awareness project[21].

The current Europeana Licensing Framework provides a unified set of terms of use that enables access to (a mandatory minimum set of) metadata and (non-mandatory) thumbnail images / previews on Europeana. Europeana requires the following terms of use:

- The mandatory minimum set of metadata is released under CC0 by the data providers. Currently the minimum set of metadata is a Dublin Core-based set of fields with twelve additional Europeana-specific elements.

- The metadata records contain a rights statement about the digital object (and preview) they refer to.

The current Europeana Licensing Framework does not cover access to the content resources themselves. Europeana does not hold any content / digital objects.[22] It only provides links to digital objects that are made available by its data providers on their own websites. Currently, Europeana provides links to the following types of digital objects:

- The majority of objects referred to are under copyright and therefore not re-usable outside of the exceptions and limitations provided by national copyright laws. Currently (as of March 2013) 10,424,723 objects are labelled "all rights reserved" and are not re-usable. A total of 7,485,462 objects are re-usable as they are either labeled as "Public Domain" (5,201,067 objects) or as "Creative Commons licensed" (2,284,395 objects). A further 7,987,367 objects in Europeana are unmarked.

- The majority of objects that are re-usable and referred to are textual objects. Currently, there are mostly textual objects available under a "Creative Commons license" or marked as "Public Domain" (3,793,830 objects). This is followed by images (902,444 objects), audio (10,559 objects) and then video (1,809 objects).

The development of re-use scenarios for the creative industries stands and falls with access to high-quality metadata as well as high-quality content that is available in formats suitable for re-use and under conditions that allow re-use scenarios. Europeana Creative will build a framework that fixes these issues and can guarantee reliable access to significant amounts of high-quality content that can be re-used by third parties.

The Content Re-use Framework will consist of a permissions infrastructure that allows data providers to communicate conditions for re-use of their content (including commercial re-use scenarios) and will be voluntary for data providers (except for project partners in Europeana

---

[20] http://www.europeanaconnect.eu/; accessed May 2, 2013.
[21] http://pro.europeana.eu/web/europeana-awareness; accessed May 2, 2013.
[22] The only exception is when in some cases (community collections projects like Europeana Collections 1914–1918), Europeana acts as a data provider to itself.

Creative that hold content) to participate in. This means that content providers both from within and outside the consortium can opt in to give access to richer content or extended metadata sets.

The Content Re-use Framework will establish a number of minimum requirements and allow for a range of different re-use scenarios within the Europeana Creative project. The Content Re-use Framework needs to set minimum requirements for digital objects that will be made available. Following the requirements gathering workshop, there should be minimal requirements:

- in terms of rights status (it must allow re-use);

- in terms of technical quality (minimum size / quality of the digital object);

- there should be direct access to the digital object.

These minimum requirements must be extended with specific access rules for special groups and re-use scenarios (for example, educational groups). These have to specify conditions under which these groups can have access to content.[23] The specific re-use scenarios and conditions will be further gathered in the project (M1–4) and produce a first working implementation of the Content Re-use Framework by month 8.

By implementing a Content Layer on top of the Europeana Licensing Framework it will be possible to handle access to content that is compatible with the rest of the Europeana ecosystem. The Europeana Cloud project[24] – which runs in parallel with the Europeana Creative project – deals with the development of a cloud infrastructure that will enable (scalable) access and storage of (rich) metadata and content via the framework.

5.2   Content Sourcing

5.2.1   Platform

In the framework of WP2 Task 2.1, a platform will be created for the retrieval of Europeana metadata and related digital content objects in order to enable re-use from applications. It is crucial for all applications to have access to content on terms that reflect the terms of the Content Re-use Framework agreed by the content provider, and this service will enable this function. On the technical implementation level, different approaches will be followed for retrieving metadata and digital objects.

Enabling access to digital objects (rather than metadata) requires the implementation of the Europeana Content Re-use Framework being specified in WP3. Some content for experimentation will be stored by the project centrally (capacity will be provided for the content that will be used by many Pilots), or within content provider repositories. The Content Re-use Framework will be tightly integrated with the metadata available via Europeana and all publicly available content will be accessible via the Europeana portal.

During the first phase there is an option of utilising a commercial cloud storage provider as this will allow the content framework to become operational quickly. After the end of the project,

---

[23] For example, for educational re-use certain users can have access to certain content in exchange for something (for example, data enrichment or attribution).
[24] http://pro.europeana.eu/web/europeana-cloud; accessed May 2, 2013.

storage requirements can either be met by a cloud storage infrastructure that is developed by Europeana, via commercial cloud storage providers or by a system that relies on the existing repositories of the content providers. Final storage options will be considered as part of the evaluation task in WP3 (Subtask 3.1.3). Access to the content will be realised via APIs and – where possible – via the Europeana portal.

Semantic queries will also be explored using commercial and non-commercial semantic triple stores. If it proves useful to query metadata using linked data SPARQL-style syntax, and if such query engines can be shown to be sufficiently performant, the project will deploy such open source or commercial platforms, including OWLIM.[25]

5.2.2   Content Acquisition Strategy

The content needed for the Europeana Creative project will be sourced using a variety of channels, based on the needs of the Pilots and the current status of the content required. Because the content focus within the Europeana Creative project is specifically to source only "content needed for experimentation", the strategy followed will be to use a demand-led model of content acquisition, rather than a more comprehensive volumetric approach. Where a trade-off of effort must be made in the project context, the preference will be for the provision of high-quality content under terms that provide the maximal amount of re-use potential, rather than the sourcing of larger quantities of metadata, or the sourcing of content too encumbered by licensing or copyright constraints to be suitable for wide re-use within the project's Pilot applications.

It is important to note the distinction between *content* and *metadata*: In this context, *content* is used to refer to a rich media file of type image, text, audio, video or 3-D model representing an item of cultural heritage. *Metadata*, of course, refers to the description of such cultural heritage objects, and in the context of Europeana is sometimes also used to refer to preview or thumbnail representations of a larger media file.

The Europeana repository has practically no content fitting this definition at present, but it is clear from the nature of the envisioned creative re-use that content of this type will be required. Sourcing such content will follow a standardised process. In general, the project will use the following sequence:

1. Evaluate the specific needs of a given thematic Pilot application to determine the content needed to support the application. This will be developed into the form of an inventory. For example, if the History or Natural History Education Pilot will explore expedition diaries, then full or partial scans of such diaries will need to be available to the Pilot developers. Note that this step will commence only at or after the initial co-creation workshop for each theme.

   a) Revise this inventory as needed throughout the project, with a view to a complete review of needed content after the initial requirements are defined for each Pilot theme.

2. Evaluate what metadata is currently in the Europeana repository that both matches the needed inventory, and which describes a suitable quantity of rich media content under terms suitable for re-use.

---

[25] http://www.ontotext.com/owlim; accessed May 2, 2013.

a) Of the metadata records identified in step 2, evaluate which records refer to media files that are already directly linked in such a way that the media file can be directly returned by the API. For these metadata records, they are probably already in a state suitable for immediate use by the project.

b) For those needed and suitable content items identified in step 2 but for which media files are not currently available directly, investigate with the content provider how these links might be enabled, either through an update to the metadata, through screen-scraping or other extraction scripting, or some other mechanism. The assumption for these records is that the media files are available, but will reside within the environments of the data provider.

c) For the identified records for which media links could be provided except for legal or licensing concerns that can be addressed by the project, investigate with the content provider how access to these media files can be provided, and what assurances of licensing adherence or other legal conditions must be undertaken in order to enable this content for re-use within the project. Because the purpose of the Pilot approach is to develop best-practice experiments, it is acceptable (though not ideal) for content to be made available under license terms that apply only to the project. Of course, longer-term contextual licenses would be preferred in the interest of sustainability.

3. Evaluate the new content on offer from consortium members that matches the availability and re-use requirements for the project. Ensure that the relevant metadata is ingested into the Europeana Repository via standard aggregation channels in good time to be used by the Pilot development activities.

    a) For example, the consortium includes content-holding partners such as the British Library (BL), the Museum für Naturkunde (MfN), the Austrian National Library (ONB) the National Museum, Prague (NMP) and the Netherlands Institute for Sound and Vision (NISV), each of whom have pledged to provide content suitable for experimentation to the project to enable the requirements of the Pilot themes. Section 2.2 of this document lists the initial assumptions made about the content items needed for each of the thematic Pilot applications. These assumptions will be examined and validated as part of the content evaluation tasks following the co-creation workshop for each theme.

    b) All consortium members are expected to make metadata available to Europeana using existing aggregation channels. For new content for which media files can be linked to a location at the providing institution using standard metadata, this will be the preferred approach.

    c) For content which can be described by standard metadata, but for which the media files cannot be reliably linked to locations within the hosting institutions, the technical partners of the project will enable a simple content-hosting solution to store and make available these media files.

4. Evaluate the rich media content that has been collected by Europeana through user-generated content campaigns and similar mechanisms (e.g., Europeana 1914–1918 campaign, Europeana 1989 campaign), for which the content meets the needs of the project and the rich media files are directly held by Europeana or its partners.

a) For these content items, verify that rich media files are being held by Europeana or a partner in a suitable format for re-use and that the user has provided needed permissions for re-use. If there are alternate versions of the media required (lower resolution image files, for example), ensure that these digital expressions can be created and stored.

5. Evaluate what content sources are required for the project that can be sourced from outside the project consortium, and what institutional and technical connections must be made to enable this channel.

a) Like content described in step 3, it may be necessary to use project resources to store media files when they are freely available but cannot be easily stored and accessed via data-providing institutions.

5.2.3 Content Analysis Starting Points

Informing the process of content acquisition in the Europeana Creative project will be other work that has already commenced showing human-readable analysis of collections. The following sources will be drawn upon.

Collections and Metadata Analysis, Strategy and Plan 2013

http://pro.europeana.eu/documents/866067/983522/D3.8+%26+MS12+Content+and+Metadata+Strategy+and+Plan+2013

This analysis was performed as part of the Europeana v2.0 project.[26] It does not go into much detail of actual content by subject, but there is a higher level analysis of content by type, country of origin and also the quality of metadata that may provide complementary information.

Europeana Rights Analysis

https://docs.google.com/spreadsheet/ccc?key=0AvqKqfVhb_m2dFMwZ3pfWUlRcjFqWklmckZaT1dmSnc#gid=0

This 2012 document was created as part of the analysis of rights statements within Europeana metadata. There is some partially relevant information for content analysis in the "Analysis Human" tab.

Europeana Content Analysis

A detailed analysis of content in Europeana, performed in the summer of 2010, is attached as Annex V to this document.

---

[26] http://pro.europeana.eu/web/europeana-v2.0; accessed May 2, 2013.

Free-Text Collection-Level Subject Metadata in Large-Scale Digital Libraries

http://dcpapers.dublincore.org/pubs/article/download/3630/1856

Paper reviewing collection level descriptions at The European Library (TEL).

Europeana Collections, 2013

Additional information on collections will be available from various databases Europeana holds. Europeana Operations Officer Francesca Morselli is investigating how best to permit access to this, as well as work out how to develop a collections spreadsheet for 2012 (as was undertaken in 2010). Information on the content of the dataset can currently be obtained by questioning the portal, the API, the open data set and the (temporary URL) SPARQL end point.

The European Library Collections, 2013

The portal and the API allow questioning of the dataset. TEL presents some collection level metadata to end users, but this is varied in quality (dependent on what was given to TEL by depositors). It is available as open data at http://data.theeuropeanlibrary.org/download/, but an easier to analyse spreadsheet should be available in April 2013. Additional information on collections should be available from TEL's customer relationship management system SugarCRM. The Europeana Library is investigating how best to enable access to this data.

6. Evaluation

The main objective of WP6 is to evaluate the results at key points in the project and measure their success against the strategic objectives.

Based on the overall Pilots planning from chapter 2.1, this means that the provided evaluation concept concentrates on a time frame of fourteen months for each Pilot (see pp. 10–11, fig. 1).

The main ambition is to assess the user acceptance and usability of the Pilots via mixed method design, whilst reflecting on the Pilot infrastructure functionality. In addition permanent feedback shall be provided to each partner involved in the development process (WP1, WP2, WP3, WP4 and WP5).

The main challenges for the evaluation are the complexity and diversity of all five Pilots and acting with consideration regarding the evaluation resources. To ensure a holistic point of view, a multi-perspective approach based on several methods like focus groups, different types of usability and user acceptance tests (UAT), expert interviews, monitoring by diary-keeping and online surveys will be chosen (fig. 11). All evaluation types are designed to be realised in an online format.
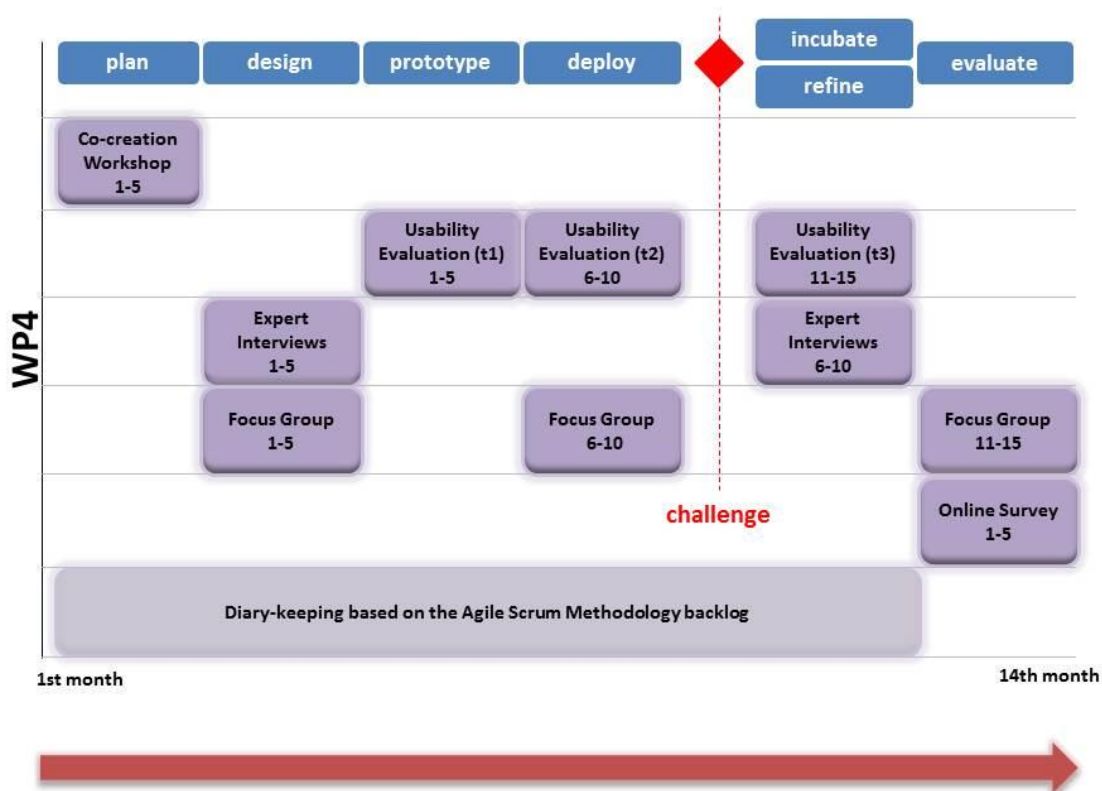

Fig. 11: Evaluation methods in Europeana Creative

The reasons for choosing this methodological approach will be explained and described in detail in the following sections.

6.1 Focus Groups

This qualitative method can be used either online or offline. An online realisation enables an evaluation with a smaller resource requirement. The planned design will be a synchronous online focus group. To suit the project purpose it should be arranged as a video conference instead of an online focus group via chat room. For the Pilot evaluation fifteen focus groups shall be conducted with focus groups 1–5 scheduled in the early Pilot phases. Based on the main interest of generating success criteria, key quality indicators and gathering feedback, the method can be adapted as a focus group via video conference tools.

The expected time frame per focus group will be two to four hours, including breaks. The focus groups will be designed as outlined below:

- **Focus Groups 1–5** will consist of eight to ten participants, thereby being composed of one representative from each Pilot and three to five external experts and stakeholders. The objective in each focus group is to define stakeholder success criteria as a basis for key quality indicators. To ensure the feasibility of comparison between the five Pilots it is important to develop standardised evaluation criteria within the evaluation group. These indicators shall enable on a meta-level the measurement of progress and success during the iterative evaluation process and also later on in the Challenges. Based on the experience within the co-funded Pilot phase, an adaption prior to the incubated spin-off is possible. Depending on the output of these focus groups the evaluation approach has to be extended with an online survey or Delphi method.[27]

- **Focus Groups 6–15** will support the monitoring of each Pilot. Thereby two additional focus groups are planned within the Pilots in chronological order with the above-mentioned kick-off focus groups. Overall ten workshops will be realised to gather feedback about progress, obstacles and need for improvement, resulting in an overall project fitness from every Pilot. The output of focus groups 6–10 will be a last input before the Challenges start. The several focus groups will be composed of six to eight stakeholders who are related to the specific Pilots.

6.2 Usability Evaluation

As already mentioned, it is not possible to apply a unified methodology for usability evaluation. Here, the chosen method depends strongly on recommendations from the Pilot Task Leads. According to specific requirements, tools like Google Analytics, whatusersdo.com, usabilitysciences.com, etc. can be used and evaluated in an iterative process (t1, t2, t3). The frequency of these evaluations can be adapted according to the needs of the different Pilots. Furthermore the user acceptance can be captured with short and concise online surveys based on the Bruce Tognazzini's principles,[28] for example.

---

[27] See Gregory J. Skulmoski, Francis T. Hartman and Jennifer Krahn, "The Delphi Method for Graduate Research", in: *Journal of Information Technology Education*, Vol. 6, 2007, available online at: http://bern.library.nenu.edu.cn/upload/soft/0-article/+03/JITEv6p001-021Skulmoski212.pdf; accessed May 2, 2013.
[28] First Principles of Interaction Design, available online at: http://www.asktog.com/basics/firstPrinciples.html; accessed May 2, 2013.

6.3 Expert Interviews

At different stages of the project it will be helpful to get in-depth knowledge from external experts involved. This can be ensured by in-depth interviews. During the entire project this method will be used. For the co-funded Pilots the objective is to get a detailed impression of each Pilot's progress from a technical perspective via one to ten expert interviews. A minimum of ten interviews are planned and shall be processed either personally during a workshop, via video chat or via phone. The participants will be identified by the WP4 Pilot Task Leads and suggestions will be collected from other WPs.

6.4 Diary-Keeping (e. g., Impediment Backlog)

This approach is supposed to serve as a backup for feedback during the development process. The co-funded Pilots are set up with the Agile Scrum methodology where a continuous backlog can be submitted to the Scrum Master. It is intended to arrange eight or nine sprints for each Pilot which include a backlog prepared by the product owner (see 4.2.9, Development Strand). Overall the benefit will be a minimum of fourty backlogs (eight sprints per Pilot) with hints for improvement and troubleshooting. Such backlog reports are convenient to document technical obstacles within the WP stages. Furthermore it is suitable as feedback instrument for other WPs and Pilots. Therefore the backlog will be provided to all WP Leads in a similar way to a science journal; a final form will need to be determined.

6.5 Online Survey

During the evaluation phase of every Pilot five online surveys conducted via SurveyMonkey or LimeSurvey shall be realised.[29] The objective is to get a response on how the involved stakeholders rate the implementation and incubation of the Pilots. By using expert interviews and focus groups it is not possible to get a wide range of feedback. The information collected with both methods will be the basis for a standardised questionnaire used in these surveys. The results will be an input for D6.2 and D6.3.

6.6 Examine the Extent of Success and Delivery

During the several stages of the project, ideally a high number of stakeholders will follow the development process. The measurement in different stages and on different levels enables the observers to give hints and suggestions for improvement which can help the Challenge participants to optimise their approaches later in the project. The objective for evaluation is to get a holistic impression of the Pilot delivery process and to discuss barriers and possibilities to avoid these. The methods can be adjusted in case they are not constructive. The co-creation workshop phases at the beginning of the Pilot development are supposed to be used for final alignments of the evaluation.

---

[29] http://www.surveymonkey.com, www.limesurvey.org; accessed May 6, 2013.

Annex I: Services

From the Description of Work:

Five key added value services will be produced in WP2 that support the Pilots delivered in WP4. It is envisaged that these services will provide uses by creative industries within the specific Themes plus expanded use if applied to datasets external to that Theme in the future.

- *Pattern recognition and image similarity service* – pattern detection software will be integrated within the Open Lab infrastructure. This software will have its capability extended beyond standard image format to detect patterns in other forms of multimedia such as video, through image segmentation. A graphical user interface will be configured to enable the user to select a photo or area of a photo as a parameter to find other similar images, with the option to filter based on the enriched data fields, such as material.

- *Geographic mapping service* – a rich tool for linking Europeana digital objects and other linked data with a geographical location will be developed to support the concept of a "cultural route" used for the Tourism and History Education Themes. A user will be able to create a route and then search Europeana using the geographic mapping tool for available resources to link to the route, applying filters such as location, historical time period or Theme.

- *User generated content management service* – the metadata and content held in Europeana and its partner institutions can be vividly enriched through linking to user generated content. This user generated content may come in various forms and the service will be implemented to provide translations for the existing linked data, text annotations and feedback facilities for the Europeana objects. This service will be utilised by the Education and Tourism Pilots.

- *Data transformation service* – metadata within Europeana is represented in the EDM format, which is generally inappropriate for direct usage within systems from outside the cultural heritage sector. The data transformation service will take XML Schema Definition (XSD) as input and output formats and provide a visual mapping editor to enable easy manipulation of data transforms. This service will be utilised by the Education Pilots for transforming IEEE LOM and serious game XML formats.

- *Linked data service* – this service will link Europeana data sources to external SPARQL endpoints using string-based and knowledge-assisted matching strategies. Entity/term extraction and/or natural language processing frameworks will be evaluated to expand the number of suggested links. The proposed service will automatically extract entities from text and will provide links to related information provided by external linked data sources, such as Freebase, DBpedia, Wikipedia, VIAF, Getty and Geonames. This service is expected to be utilised by all Pilots to varying extents.

APIs

The Pilot applications will access the infrastructure services through well defined service APIs which are designed to provide access to content by applying the access rules defined by the Content Re-use Framework. Messaging protocols will be developed and used for remote service invocation. The input and the output of each service will be represented by business messages that standardise the interaction within the Open Lab environment by making an abstraction of the application context (e.g., type of the application, user session, authentication and authorization information).

The APIs will be formulated around the requirements of the Pilots but designed and implemented in a appropriate manner for future expanded use and scalability.

Annex II: Stakeholders of the Europeana Creative Project

From the Description of Work:

Europeana Creative focuses upon two **primary stakeholder audiences**. These are:

- **Creative industries** are a primary audience as they provide enormous potential to increase access to cultural heritage resources for members of European society. They bring the capacity, capability and appetite to re-use the resources as elements in the construction of innovative applications and services for their clients. The project's main aim for this target audience is to increase awareness of the availability of cultural heritage resources and promote the benefits of using the infrastructure provided to create value and economic growth.

- **European cultural heritage institutions** will be engaged by the project as a primary audience in order to increase the volume of resources available for the creative industries to re-use, by widening the network of institutions providing content. The Europeana Network represents more than 2,200 institutions, each of which would be a valuable addition to the project once they join. The project's main aim for this target audience is to increase the awareness of new business models and benefits of working with the creative industries, thus obtaining commitment to release their cultural heritage material under the terms of the Content Re-use Framework.

As the project focuses upon various themes there will be a group of **secondary stakeholder audiences**, as follows:

- Education sector (the education sector's user base includes the teaching workforce, the learners themselves and all the educational content and service providers).

- Tourism sector

- Design communities

- Social networks

- Independent software and application developer communities

Annex III: Scrum Adaptation for Europeana Creative Pilots

Roles and Actions

Product Owner: Pilot Leads

- Responsible for the Pilot Backlog. Prioritising work around the Pilot, knowing what is required of it and which requirements to prioritise.

- Actively involved in order to supervise the Pilot frequently and guide its development at every step.

- Defines priorities and requests during the development sprint.

- + info: http://www.mountaingoatsoftware.com/scrum/product-owner

Scrum Master: Platoniq

- Responsible for supporting the development team, guiding them through the process and removing impediments blocking their work.

- Facilitates the creation of the Pilot Backlog, focusing on its dependencies and deployment plan.

- Defines checkpoints/test to start in sprint planning, as involving evaluation at the start helps to clarify requirements.

- Responsible for facilitating the Scrum meeting. Keeping it focused, timely and "on topic".

- Idem with the regular Scrum meetings during the Pilot sprint.

- Define calendar of sprints and checkpoints for the Pilots.

- Define Sprint Planning Workshop as a recurring appointment before every sprint.

- Provide and monitor online collaborative tool for whiteboard (Pilot Backlog, Tasks To Do, Work In Progress, Ready To Be Verified and Done)

- Define "Done" (in relation to WP6) and indicators to help evaluation/testing at checkpoints; http://www.allaboutagile.com/definition-of-done-10-point-checklist/.

- + info: http://www.mountaingoatsoftware.com/scrum/scrummaster

Team: Developers, Content Providers, Other WP Partners

- Depending on their role and involvement according to the Description of Work: participate in the sprint planning, decide and help to prioritise features, develop or work on specific tasks and report regularly via regular Scrum meetings.

Guidance: NISV

- Gives support, guidance, coaching and assistance if needed.

- Helps guiding the Product Owners through the process and removing impediments blocking their work.

- Co-responsible for facilitating the Scrum meeting at the intersection of each development sprint. Keeping it focused, timely and "on topic".

#1 Define Backlog (1ˢᵗ Week)

The Pilot Backlog is (re)prioritised every four weeks with the list of things collectively decided as needed for the Pilot development. They are defined separately (so they can stand alone as discrete, deliverable pieces of work). Anyone can add anything to it. But *only* the Product Owner can prioritise things there. The type of inputs can be diverse (expressed in the less technical language possible):

- Content needs

- Features

- Bugs

- Enhancements

- Issues

- Risks

- Technical work

- Knowledge acquisition

Priority is determined simply by the order of items in the list. The Pilot Backlog evolution should be completely visible for anyone involved or interested in the Pilot.

The definition of items and prioritisation will take place in the last activity of the co-creation workshop, and afterwards shared and updated online openly using Trello (with admin access only for Product Owners, the Scrum Master and WP1 leaders). For example, for the Natural History Education Pilot see: https://trello.com/board/natural-history-education-Pilot-ecreative/51506359fd3e017166008b9a.

#2 Estimate Sizes in Backlog (1ˢᵗ Week)

We will estimate the Pilot Backlog in points, not in units of time. Rather than asking "How long will it take?" we are asking "How big is it?". Taking this as an approximation, and not as a premature commitment.

We will use any of these numbers for indicating each size: 1, 2, 3, 5, 8, 13, 21. Negotiating the size of each Backlog item as a team, with the main input of the Product Owner where needed. After that, the Product Owner has another quick look at priorities, reallocating things if he/she thinks it is important.

#3 Sprint Plan (I): Requirements (1ˢᵗ Week)

As the first activity for the sprint planning meeting (attended via Skype or similar by the whole team, developers and even end users / testers included if possible) we need to:

1. Select target items from the Backlog for the sprint: an objective that sums up the goal for the next sprint, from a section of items from the top of the Pilot Backlog that the team thinks can be achieved (including a bit more than we think can be done, just in case).

2. Clarify sprint requirements: the Product Owner presents each item and explains how he/she sees it working from a functional perspective. The whole team discusses the item in detail, asking questions about the feature in order to establish what it should do and how it should work. Those questions include also testing and evaluation issues for each item.

The outcomes of this discussion are captured on a pad and transferred as minutes of each Pilot to Basecamp, writing requirements feature by feature before they are developed, in a way that is as lightweight and as visual as possible.

#4 Sprint Plan (II): Tasks (1ˢᵗ Week)

In order to break the requirements into tasks and estimate the time required to complete them, as a second round in the meeting we will work on how the selected backlog items will be delivered:

1. Calculating the available number of days/hours the team has to work on the sprint.

2. Breaking the requirements into tasks: design, development, unit testing, system testing, documentation, etc. Trying to state tasks as deliverables, if possible.

3. Estimating tasks in days/hours for each item: keeping them small and discussed.

4. Adding up all the tasks estimated from the Pilot Backlog to the "To Do" one, in order to define the Sprint Backlog. If needed, it should be the lower item(s) on the Backlog that are removed from the sprint.

5. Identifying extra items as stretch tasks for the sprint, in case of an over-cautious approach to the estimates.

#5 Regular Scrum During the Sprint (2ⁿᵈ & 3ʳᵈ Week)

Every Tuesday and Thursday morning via Skype chat (or similar), each team member reports back to the rest of the team in turn, during a 15–20 minutes session. Reports should be concise and focused, addressing three key questions:

1. What have we achieved since the last meeting?

2. What will we achieve before the next meeting?

3. Is anything holding up our progress? (Impediments)

In case of important questions related to a task, they should be raised but avoiding to discuss them in detail until the end of the Scrum meeting (when those deserving discussion can stay back to be discussed together with WP1 and WP4 leaders, after the rest is back to work).

#6 Reviewing and Repeating (4th Week)

1. Sprint review meeting, open to all partners: demoing what has been completed (using Google Hangouts or similar).

2. Retrospective (only among WP1, WP4 and other WPs if needed):

    1. Did the team deliver what they committed?

    2. What went well?

    3. What could have gone better?

    4. What to do differently in the next sprint?

3. Third part of the session: We repeat the process (prepare new sprint from step #1)

Principles

(Adapted from http://morepm.com/tag/scrum-principles/)

- The highest priority is the early and continuous delivery of valuable software.

- Welcome changing requirements, even late in development.

- Agile processes harness change for the Pilots competitive advantage.

- Deliver working software frequently, with a preference to the shorter timescale.

- Partners and developers must work together regularly throughout the project.

- Build Pilots around motivated individuals. Give them the environment and support they need, and trust them to get the job done.

- Although the most efficient and effective method of conveying information to and within a development team is face-to-face conversation, we will adapt to online channels.

- Working software is the primary measure of progress.

- We should promote sustainable development.

- The partners, developers and users should be able to maintain a constant pace indefinitely.

- Continuous attention to technical excellence and good design enhances agility.

- Simplicity – the art of maximising the amount of work not done – is essential.

- The best architectures, requirements and designs emerge from self-organising teams.

- At regular intervals, the team reflects on how to become more effective, then tunes and adjusts its behaviour accordingly.

References

https://en.wikipedia.org/wiki/Scrum_%28development%29

http://www.scrum.org/Scrum-Guides

http://www.allaboutagile.com/category/how-to-implement-scrum-in-10-easy-steps/

http://www.allaboutagile.com/what-is-agile-10-key-principles/

http://www.amazon.com/Agile-Software-Development-Scrum/dp/0130676349

http://www.mountaingoatsoftware.com/topics/scrum

Co-creation workshop methods > eCreative Pilots

CORE PRINCIPLES:

- **Reorder, iterate and improve with modularity**
- **Adapt from scratch or for previous developments**
- **Expand or fork for challenges and labs**
- **Focus on deliverables (diagrams, stories, mockups, etc)**
- **Integrate observation, documentation and evaluation**
- **Intersections between offline, digital & online**

1. Get inspired + Ice-breaker around context > Walkshop

2. Elicit expectations > Scenario forecast

3. Play with samples of content > Mapping & sorting

4. Shape & process > Flow design / Rapid prototyping

5. Polish & share > Presentations + Distributed Q&A

6. Evaluation + Next steps > ExpertsXpress meetings

Content Analysis
Content available through Europeana.eu – Summer 2010

List of figures and tables

Introduction
This content analysis gives an overview of the representation of subjects, time periods, languages and countries related to the collections available through Europeana.eu during summer 2010.

Its purpose is internal; it helps to determine target audiences for Europeana's marketing strategy and is closely related to Europeana's Content Development Strategy. It starts out with a summary and a description of the used methodology, followed by an overview of the results. To increase readability, not all tables and charts are used in the report. Additional information is available in Appendix C (separate document) which follows the same structure. The paper is concluded with suggestions for further research.

Summary
This analysis aims to give an impression of the subjects, time periods, languages and countries represented by the 12 million items available through europeana.eu.

Image and text dominate clearly (63% and 35%) and there is a great lack of sound and video material (each less than 1%).

Overall, the four most popular subjects are:
▪ Books and articles: manuscripts, rare books, literature, poetry and ephemera.
▪ (art-)Historical artefacts: postcards, ethnographic material, folkloristic objects and medals.
▪ Photography: historical photographic collections of certain regions as well as ethnographic collections, and portraits.
▪ Art: paintings and drawings.

The content is relatively recent: over half of the collections holds content from between the 18th and 20th centuries. Prehistory and the Middle Ages, as well as contemporary times are underrepresented with around 25% together.

36 languages are used in the collections. Most common are German, French & English. From the European languages Bulgarian, Estonian, Icelandic, Latvian, Lithuanian, Slovak and Romanian are used least often, with each less than 1%.

When looking at the country of origin for content, the largest contributors are France, Germany, Spain and Sweden (between 8-13%). Almost half of the total list of countries score very low with under 1% (especially Iceland, Latvia, Macedonia, Monaco, Turkey and Malta).

A rough overview of the display of thumbnails on europeana.eu was made manually, which shows that half of the collections in europeana.eu displays colour or mixed colour & black and white thumbnails; roughly one third of the content does not show a thumbnail.

The following table shows the overview of the strong and weak areas within each category according to characteristic. However, the audiovisual categories are very small and therefore the conclusions are not as reliable as for the other two categories. These two need further development in general.

europeana
think culture

	TEXT		IMAGE		VIDEO		SOUND	
	Strong:	Weak:	Strong:	Weak:	Strong:	Weak:	Strong:	Weak:
Subject	Books & Articles (manuscripts, literature, etc), Periodicals (newspapers, magazines), Archives.	Music (scores and lyrics), texts about performing arts & film, political documents, UGC	Photography (art-) Historical artefacts Art Historical maps	Technical Archeology Medical Biology Economy	TV broadcasts Recordings (no other specification) Documentary	(Silent) film News Interviews	Dialects & accents Folk Radio	Music (Jazz, Contemporary, Classic) Wildlife sounds Ethnographical recordings
Time Period	16-20th centuries	Prehistory – 14th & Contemporary	17-20th centuries	Prehistory – 14th century	Contemporary & 20th century	Prehistory – 19th century	Contemporary & 20th century	Prehistory – 19th century
Language	English French German Spanish Latin	Bulgarian, Icelandic, Latvian, Lithuanian, Maltese, Portuguese, Romanian.	German French English Swedish	Bulgarian, Czech, Danish, Estonian, Finnish, Greek, Hungarian, Irish, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak.	English German Spanish French	Bulgarian, Estonian, Latvian, Lithuanian, Maltese, Polish, Slovak, Slovene.	English French German	Bulgarian, Czech, Danish, Dutch, Estonian, Greek, Hungarian, Latvian, Lithuanian, Maltese, Portuguese, Romanian, Slovak, Spanish.
Country	Spain Ireland France Germany	Belgium, Bulgaria, Cyprus, Estonia, Iceland, Latvia, Lithuania, Luxembourg, Poland, Romania, Slovakia, Sweden, Switzerland.	Germany France Sweden	Cyprus, Czech Republic, Finland, Hungary, Latvia, Lithuania, Luxembourg, Macedonia, Monaco, Norway, Poland, Portugal, Serbia, Estonia, Iceland, Slovakia, United Kingdom.	Austria Europe	Belgium, Bulgaria, Cyprus, Estonia, Finland, Greece, Hungary, Iceland, Latvia, Lithuania, Luxembourg, Norway, Poland, Romania, Slovakia, Slovenia, Sweden.	France Austria	Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Greece, Hungary, Iceland, Italy, Latvia, Lithuania, Luxembourg, Poland, Portugal, Romania, Slovakia, Spain, Sweden.

Table 1 – Complete overview of strengths and weaknesses

1. Methodology

This research was done during summer 2010: it incorporates all content available through Europeana.eu up to 19th August 2010. During this time, Europeana gave access to **268** collections, together making up for almost **12 million** items.

The following information of each collection was gathered: name of provider, country of provider, whether it came in through a project, The European Library or other aggregator, type of content was (text / image / sound / video and totals), the language used, how the material could be described, the time period that is covered and location (both country of origin of the material as well as the country of the provider). The comparison was in some cases difficult: some providers offer their entire collection in one single set (i.e. Scran) others split thematic collections up in separate sets (i.e. Swedish Cultural Heritage).

The first step was analysing the 'analysis' HTML files from the collections – these sum up metadata fields used in the collection, i.e. temporal, spatial, subject, language and location. 157 collections out of the 268 had files that could be analysed (in some cases they were in a difficult format or relating to a collection too large to make sense). The comparison was in some cases difficult: some providers offer their entire collection in one single set (i.e. Scran) others split thematic collections up in separate sets (i.e. Swedish Cultural Heritage).

Missing information was gathered from files held by Europeana's content ingestion team, such as old questionnaires and information about the collections in the CMR-system. Information on content coming from the projects comes from their descriptions of work, or in some cases more up to date sources, i.e. EuropeanaLocal's Event Log. To make this research accurate, only already ingested content at the time of the research has been taken into account. The final step was doing searches on europeana.eu using the syntax allowing view on separate collections by their identifier.

Caveat
The percentages used in this report are estimates, and relate to 'presence' in stead of a hard number. With the information available, it could only be made clear that in a certain collection there is material from for example, a certain country (the same is applicable to subject and language). Unfortunately this does not say anything about the weight of that percentage, or the relation to the exact amount of items from that specific country, only that there is an X amount in that collection. Unless research is done with more information available using a different approach (automatically and not manually for example) this could not be prevented.

The percentages in this report therefore relate to the amount of times a characteristic (subject, time, language, country) was featured in a collection. For example, France was mentioned as country of origin of content in collections 35 times as opposed to Latvia, which is mentioned only once (there is one Latvian collection in Europeana, but there are several French collections as well as other collections featuring French content). Adding this up for all countries, gave France a percentage of 10,4% and Latvia 0,3%. For subject categories this is slightly different, since 1, 2 or 3 categories could be allocated per collection. This is explained in more detail in the next paragraph.

Initially the sizes of the collections were taken into account, but the Europeana content ingestion team provides quantitative overviews, showing how many objects from which country or type are available through Europeana. Some information on the sizes of collections is still featured in this paper; however, for numbers that are more exact the reports of the content ingestion team are preferable.

Finally, within the categorisation, there was no distinction made between actual categories and subjects (for example, 'Art' as a category and 'Medical' as a subject). This benefited the detail of allocating subjects, enabling defining a collection for example not only as art, but also on what

subject (if relevant). However, this made a comparison between the several categories somewhat unequal.

Categorisation
The information that was gathered is vast and differs greatly, so in order to draw conclusions, it was divided into several categories. The categories were time period, country, language and subject.

The time periods covered by the collections (usually partially) were:
A. Prehistory: until 5th century
B. Medieval: 5th-14th century
C. 14th century
D. 15th century
E. 16th century
F. 17th century
G. 18th century
H. 19th century
I. 20th century
J. First World War
K. Second World War
L. Contemporary (from 2000 onwards)

Prehistory, Medieval times, the First and Second World War are specified separately because they are often treated separately by the content holders as well and thus specified as such in the collections.

The locations related to the content can be interpreted in two ways: first there is the country the content provider is from and second there is the country the content originates from. These two overlap often of course, but there are exceptions; most notably in ethnological collections where the content is not from the country the museum displaying it is at.
The country the provider is from was easier to determine than the country content originates from: for the second, the provider had to have supplied this information in the metadata which was not always the case. In this current research, it was decided to use the country the content originates from for the more detailed research. However, both are compared in the first chapter too.

The languages were abbreviated in the same way they are in the metadata and on europeana.eu (ISO standard). On europeana.eu, topline pages are translated in all languages used by full partners. These languages are all the official languages of the European Union supplemented by Icelandic and Catalan. Outside of this group, there were several other languages detected which were taken into account. The full list is available in Appendix B.

By doing a series of quick searches for each collection, a rough overview was made of the display of the content: showing thumbnails or no thumbnails, and in colour, black & white and mixed. These results are estimates, and are also subject to change, since Europeana's content ingestion team has worked on the display of thumbnails since.

Subjects
The categorisation into subjects of the collections was more laborious than other elements such as language and country. Because of the short time frame available to perform the analysis, it was decided to not use an existing standardised categorisation (i.e. Dewey Decimal or Library of Congress) but to develop it along the way according to what was available in the collections. For example, when a collection consisted of illuminated manuscripts, that category was added into the categorisation, which built up the list of subject categories incrementally. The complete categorisation is attached as Appendix A.

The lowest level of categorisation is made up of the four basic Europeana categories: Text, Image, Sound and Video. These were detailed further into up to three levels; for example: 1. Text → 1.3 Books & Articles → 1.3.1 Manuscripts and rare books → 1.3.1.1 Illuminated.

This caused some overlapping: some content is strictly speaking the same type, but is categorised differently across all providers, because of a choice made by the content holder to file them in a certain way. For example, an illuminated manuscript can be categorised as either text or image. Therefore in this case, there are categories for illuminated manuscripts in both the Text and Images categories.

Each collection was analysed and then categorised into up to 3 subject categories. The decision to allocate a subject (or 2 or 3 subjects) was based only on whether these covered the content of the collection sufficiently as far as could be judged with the information available. Of all collections, 134 were not given a second & third category and 61 did not need a third category. There was no possibility to oversee the quantity of each subject category in the collection, so there is no order of importance in the amount or order of subjects allocated. For example: a collection from the European Film Gateway has both images (film posters) and video files (trailers) – so the collection was given the subject categories 2.5.8.1 Film stills/posters and 4.1.3 Film. This does not mean there are more film posters in the collection than video files; it simply means both categories are present.

How many times subject categories are mentioned does not add up to the total amount of collections. Bigger collections have of course more 'vague' descriptions than smaller ones. For example, the collection from Scran is everything in one set, so it received the most encompassing subject categories such as 'Art'. On the other hand, a small specialised collection such as a political poster collection could be given one detailed subject category.

All the information is collected in a spreadsheet, available here. To make the spreadsheet somewhat readable, codes were used for the different characteristics.
Besides the collected results in this report, it also now gives the possibility to look up specifics from a certain collection, provider, country, etc. For example: collection nr. **05601** is National Library of Finland, has **188494** texts, which makes it a **Large** collection. The **Swedish, Finnish and German** languages are used, it is the **Historic Newspaper Library** with All Finnish newspapers published between **1771-1890** (Newspapers, periodicals, ephemera) giving it the time code **GH** (18th and 19th century) and the categories 1.1 **Periodicals**, 1.1.4 **Historical Newspapers & Journals**, 1.3.8 **Ephemera**.

2. Results

2.1 Topline results

This first section discusses the results on the highest level. It shows the most popular subjects across all the collections, and the languages, countries and time periods attached to those. Besides this, it shows the most used languages, countries and time periods across all collections separately.

Overall

Summarising the categorisation of all collections, the following charts show the division of content in total. It shows that images and text dominate clearly. The left chart shows the amount of collections with one of the categories (for example, there are 13 collections with video content). The right chart is based on the actual amount of items categorised as text, image, sound or video – which shows an even lower percentage of audiovisual content.
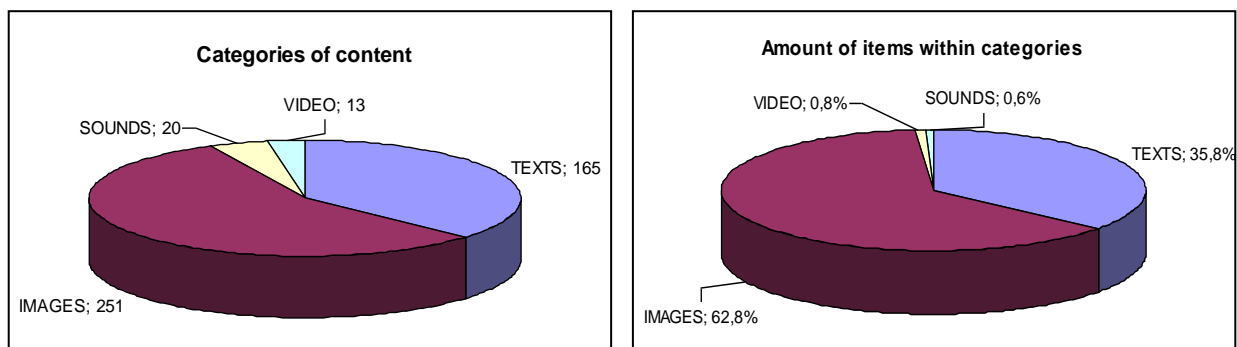

Figure 1 – Overview of categories of content in totals

Overall, the most popular subjects are 'Books and articles' (23%, mostly manuscripts, rare books, literature, poetry and ephemera), '(art-) historical artefacts' (10%, mostly postcards, ethnographic material, folkloristic objects and medals), 'Photography' (17%, mostly historical photographic collections of certain regions as well as ethnographic collections, and portraits) and finally 'Art' (12%, mostly paintings and drawings). Together these four categories make up for 62%.
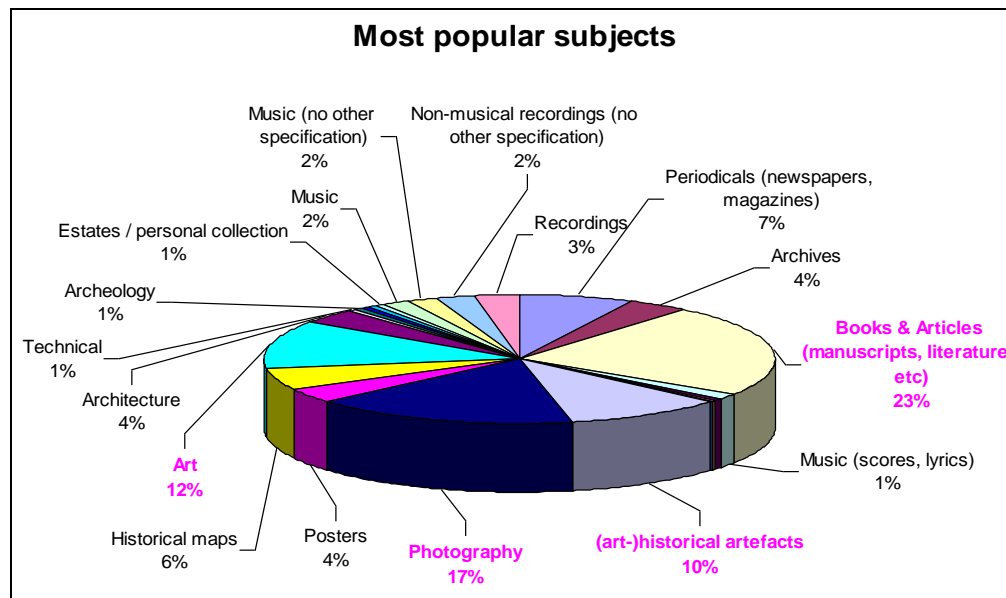

Figure 2 – Overview of most popular subjects

The following table shows which language, country and time period can be attached to these four most popular subject categories:

Subject	Language used most in this category:	Country content originates from mostly:	Time period content comes from mostly:
Books & Articles	English, French, German, Spanish/Catalan and Latin	Spain, Ireland & France.	15th-20th century
(art-)Historical artefacts	German, English, Swedish and French.	Germany, Sweden & France.	16th century – now
Photography	German, English, Swedish and French.	Germany, Sweden & France.	19th century – now
Art	German, English, Dutch and French.	France, Germany, Switzerland.	Prehistory – 18th century

Table 2 – Most popular subjects

Time periods
The content in Europeana is relatively recent: over half of the collections (55%) holds content from between the 18th and 20th centuries. The earlier ages are underrepresented: the Prehistory & Middle Ages only account for 8% together. Also underrepresented is contemporary material (2000 onwards) with 7%. Early modern times (14th till 18th century) have a quite good representation with 31% together.
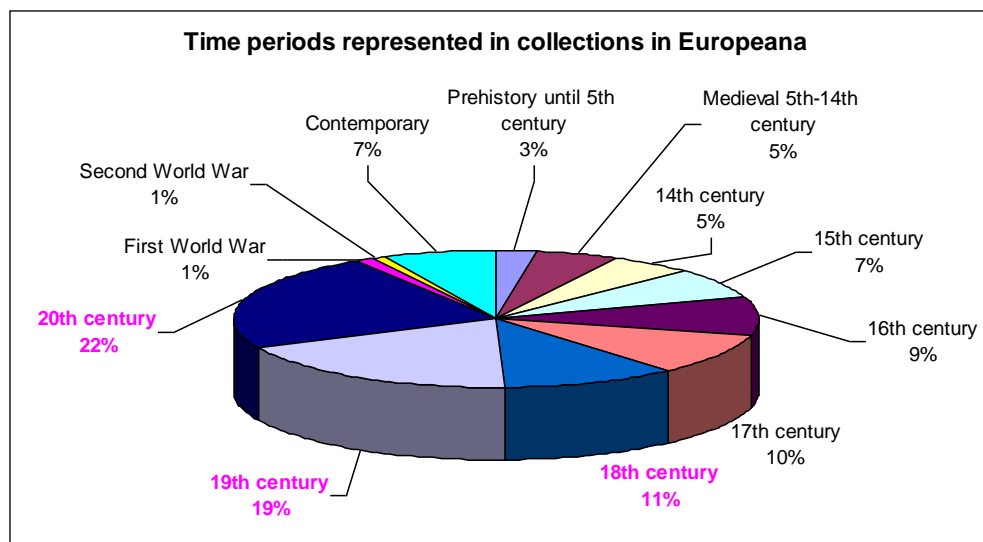

Figure 3 – Overview of time periods

Languages
Within the collections, there are 36 different languages used, of which 10 are not used in the Europeana.eu interface (see Europeana's language policy). The languages most commonly used are German, French & English (each account for approximately 15%). See for a full list of used languages Appendix B. If a language does not show in this graph, it does not mean it was not used, but that the percentage was too low to show (under 1%).
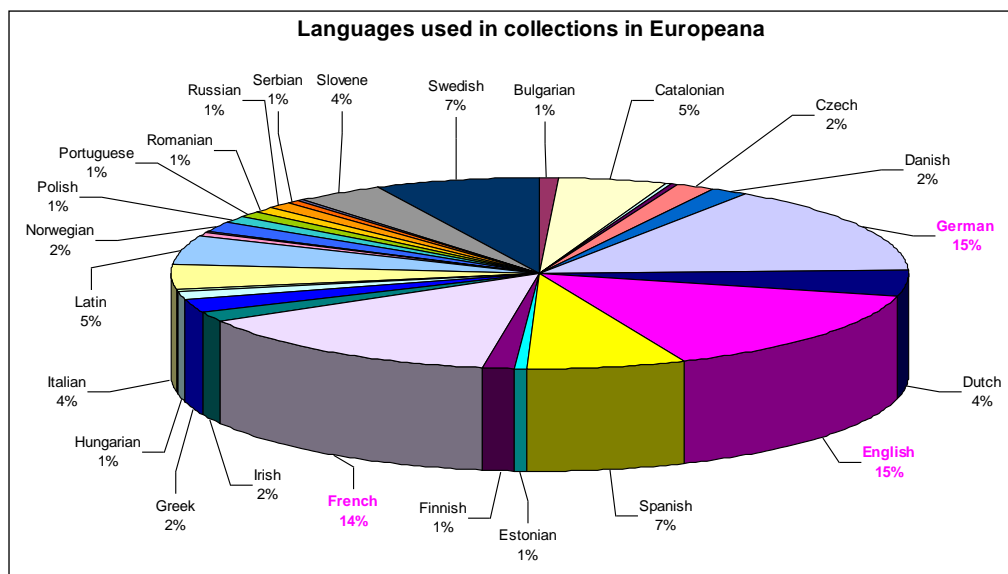
Figure 4 – Overview of languages used

From the European languages (not taking into account minor languages or non-European languages such as Coptic or Hebrew) Bulgarian, Estonian, Icelandic, Latvian, Lithuanian and Romanian are used least often, with each less than 1%.

Countries
When comparing the two ways countries are represented in the collections (on the one hand countries the content originates from and on the other hand countries the content provider is from) there are only small differences. The main difference is that of course there are no content providers from outside of Europe, as well as no providers from Macedonia, Monaco and Turkey (who are not members of the European Union).

	Countries content originates from	Countries the content providers are from
Austria	3,6%	3,0%
Belgium	2,7%	3,0%
Bulgaria	1,8%	1,9%
Cyprus	0,6%	0,7%
Czech Republic	0,9%	0,7%
Denmark	3,0%	3,0%
Estonia	0,6%	0,7%
Europe	2,1%	3,4%
Finland	1,8%	2,2%
France	10,4%	10,5%
Germany	13,0%	13,5%
Greece	1,5%	1,1%
Hungary	1,8%	2,2%
Iceland	0,3%	0,4%
Ireland	6,5%	7,9%
Italy	3,6%	2,2%
Latvia	0,3%	0,4%
Lithuania	0,6%	0,4%

Luxembourg	0,6%	0,4%
Macedonia	0,3%	
Monaco	0,3%	
Norway	1,2%	1,1%
Outside Europe	6,2%	
Poland	0,6%	0,7%
Portugal	0,9%	0,7%
Romania	0,9%	0,7%
Serbia	0,6%	0,7%
Slovakia	0,6%	0,7%
Slovenia	3,6%	4,5%
Spain	10,1%	12,0%
Sweden	7,7%	10,5%
Switzerland	3,0%	3,7%
The Netherlands	3,3%	1,9%
Turkey	0,3%	
United Kingdom	5,3%	4,9%

Table 3 – Overview of representation of countries

Comparing both options, France, Germany, Spain and Sweden are dominant (between 8-13%).
Almost half of the countries score very low with under 1%: looking at the right column in the table
above (which in this case is most representative, since this excludes the small categories of
countries that are not European and therefore are logically underrepresented) that is the case for
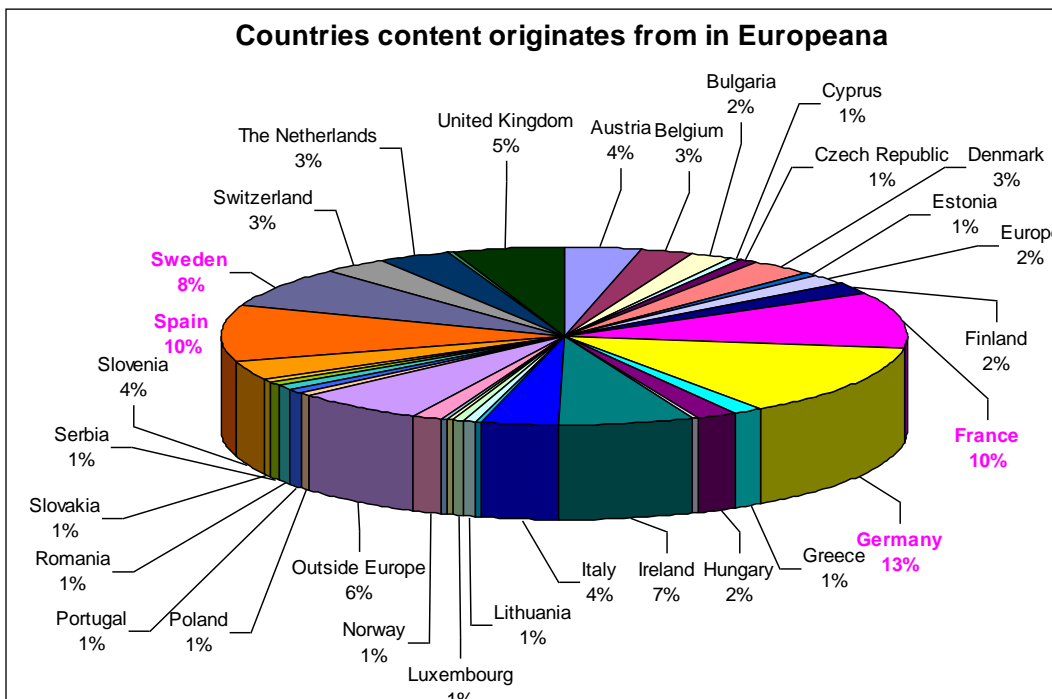12 of the 31 countries.

Figure 5 – Overview of countries content originates from

The category 'Outside Europe' (in this context, Europe is meant in the broadest definition: the 27
European Union countries, including candidate countries and also the rest of European countries)
can be further specified (see graph). Because of the small size of this category, the numbers are
actual values in stead of percentages – for example, this means that there
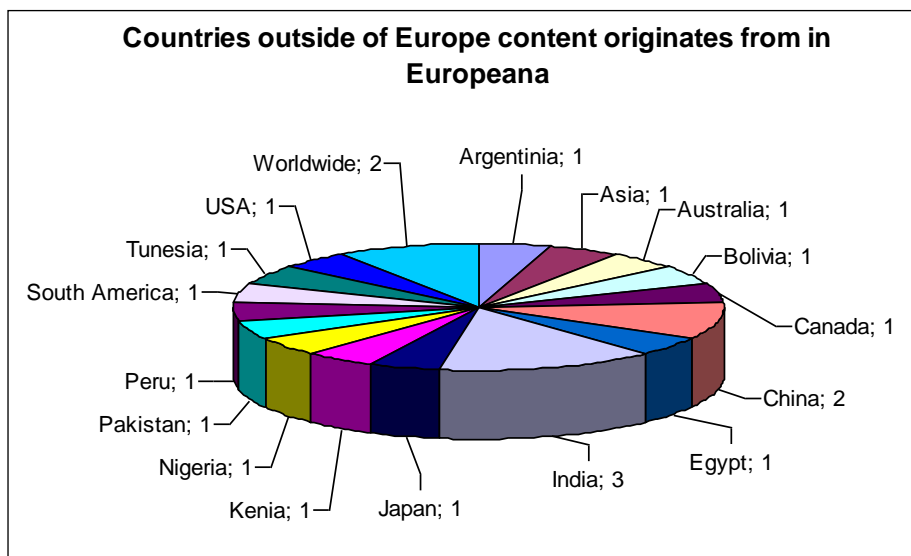is content from India in three collections.

Countries outside of Europe content originates from in Europeana

Worldwide; 2
Argentinia; 1
USA; 1
Asia; 1
Australia; 1
Tunesia; 1
Bolivia; 1
South America; 1
Canada; 1
Peru; 1
China; 2
Pakistan; 1
Egypt; 1
Nigeria; 1
India; 3
Kenia; 1
Japan; 1

Figure 6 – Overview of non-European countries content originates from

Thumbnails

By doing a series of searches of each collection, a rough overview of the display of thumbnails on europeana.eu could be made. It shows that half of the collections in europeana.eu displays colour or mixed thumbnails. This also shows that roughly one third of the content does not show a thumbnail.
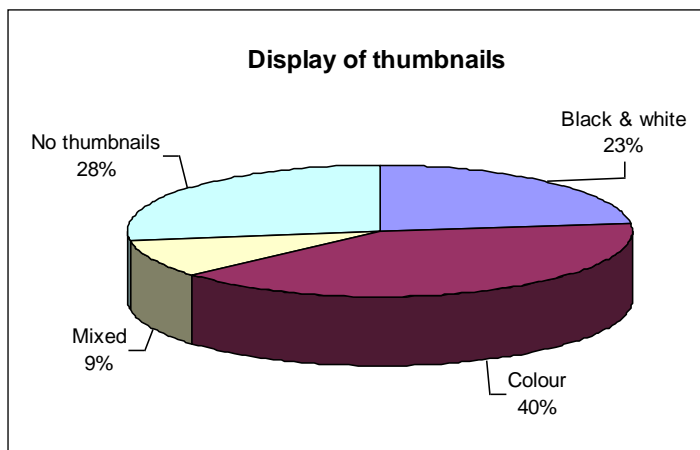
Display of thumbnails

No thumbnails 28%
Black & white 23%
Mixed 9%
Colour 40%

Figure 7 – Display of thumbnails

The following four paragraphs continue the analysis by following the first level of categorisation: **text, image, sound & video***. The choice to ingest the content as one of these four main categories is made by the content holder. Therefore there are some overlapping categories between the four main ones.*

2.2 Text

The content that is categorised as text can be divided up into 'Periodicals', 'Archives', 'Books & articles', 'Music', 'Performing arts & film', 'Political' and 'User-generated content'.
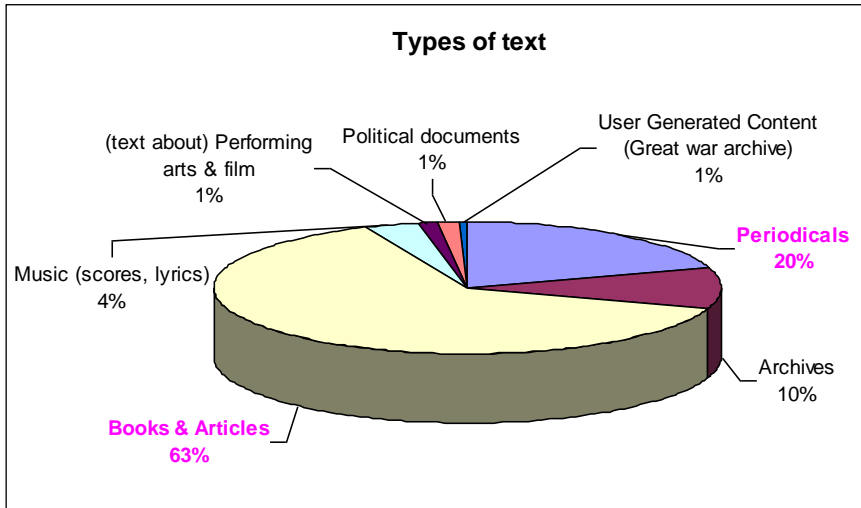

Figure 8 – Overview of types of text

Besides a rather good representation of 'Periodicals' in Europeana (20%), the most common category is 'Books & articles'. This can be specified further into 21 categories, relating to both types and subjects. Within those subcategories, the main one is 'Manuscripts and rare books' (present in 23 text collections).

Most of the text collections originate from the 19th and 20th centuries, with the 16th, 17th and 18th centuries also well represented.
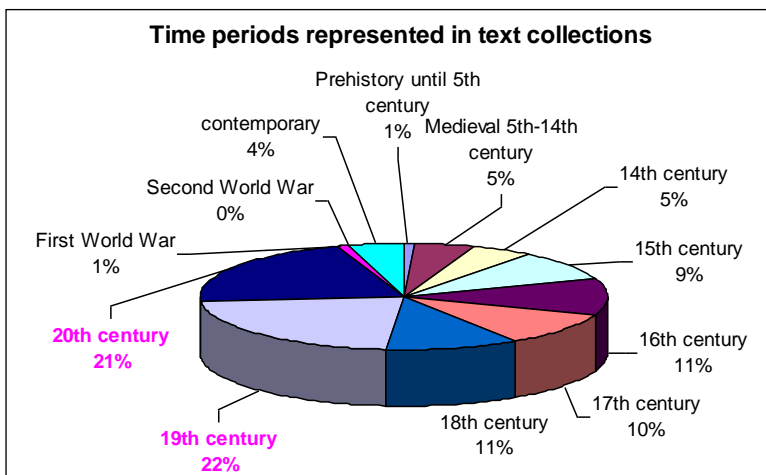

Figure 9 – Time periods represented in text collections

This category, the text collections, has of course the highest diversity in languages used – this even includes Latin, Coptic and Slavonic, used in old manuscripts mainly. With the exception of Latin (used in 24 collections) these are not used often. The most popular languages are the usual suspects: English (16%), French (11%), German (11%) and Spanish (9%).
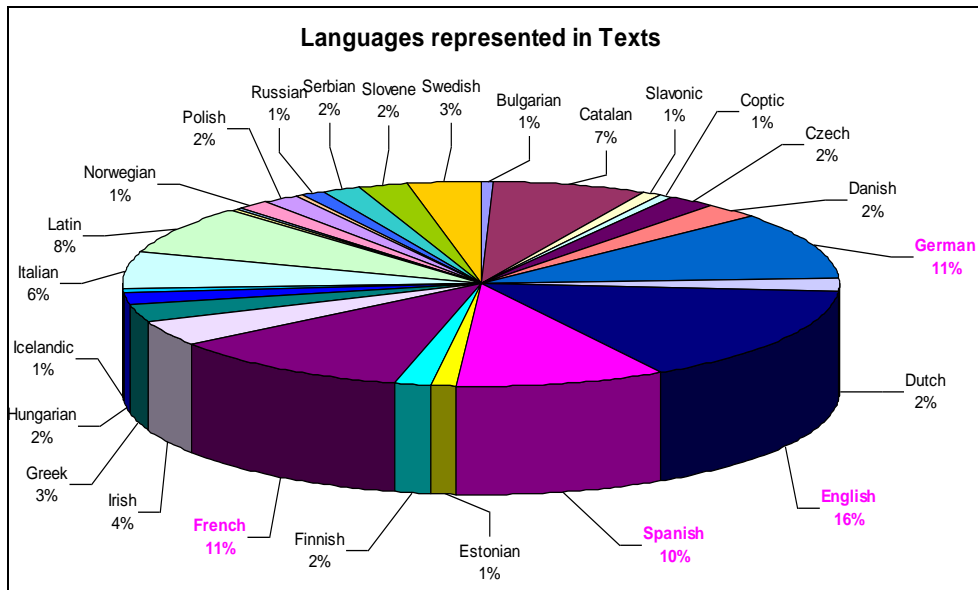
Languages represented in Texts

Figure 10 – Languages represented in text collections

This can be broken down a bit further when looking at the three biggest categories within the text category and their top 3 of languages used:

		Amount of collections the language is used in
1.1 Periodicals	German	7
	Slovene	7
	French	6
1.2 Archives	English	10
	German	4
	Irish	4
1.3 Books & Articles	English	31
	Spanish	24
	French	22

Table 4 – Languages used in main three categories of text collections

These top threes coincide partly with the countries best represented in text collections. Spain (14%), Ireland (13%) and France (10%) are represented the best. The popularity of Slovene in 'Periodicals' comes from several collections of journals and newspapers from the National and University Library in Ljubljana.
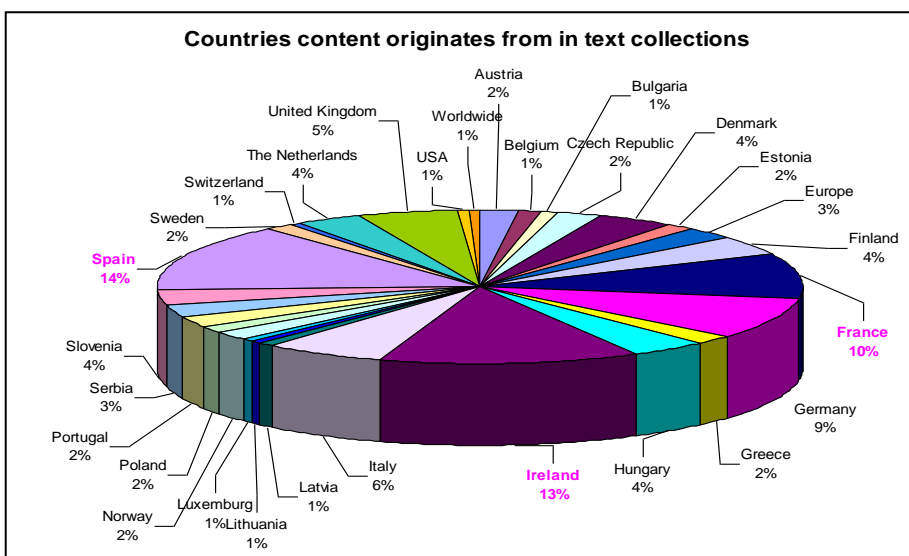

Figure 11 – Countries content originates from in text collections

2.3 Images

The content categorised as text can be divided up into '(art-)historical artefacts', 'Photography', 'Posters', 'Historical maps', 'Art', 'Estates / personal collections' and the subjects 'Architecture', 'Technical', 'Archaeology', 'Medical', 'Biology', 'Economy' and 'Music'.
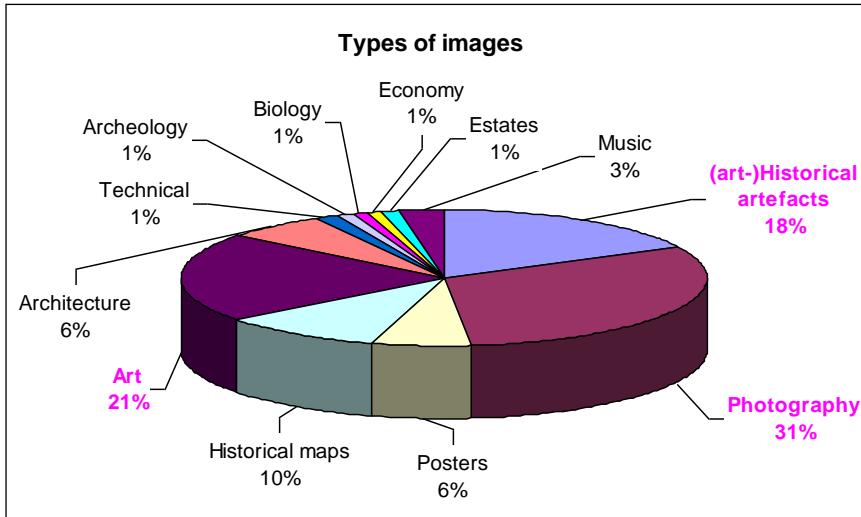

Figure 12 – Types of images

'(art-) Historical artefacts' and 'Art' are large collections, as to be expected in a portal for cultural heritage. However, the biggest category of images is 'Photography', which can be broken down further into 9 categories: 'Aerial', 'Ethnography / folklore / travel', 'History / regional', 'Portraits', 'Archaeology', 'Military', 'Maritime', 'Nature / animals' and 'Sport'. Within those subcategories, the main one is 'History / regional' (present in 32 image collections). Many of the national aggregators, archives or providers to EuropeanaLocal have photographic collections of a region, consisting of 'documentary' photos of the buildings, people and surroundings of certain areas.

Coinciding with the majority of photographic collections in this category, the 19th and 20th centuries are best represented across all image collections.
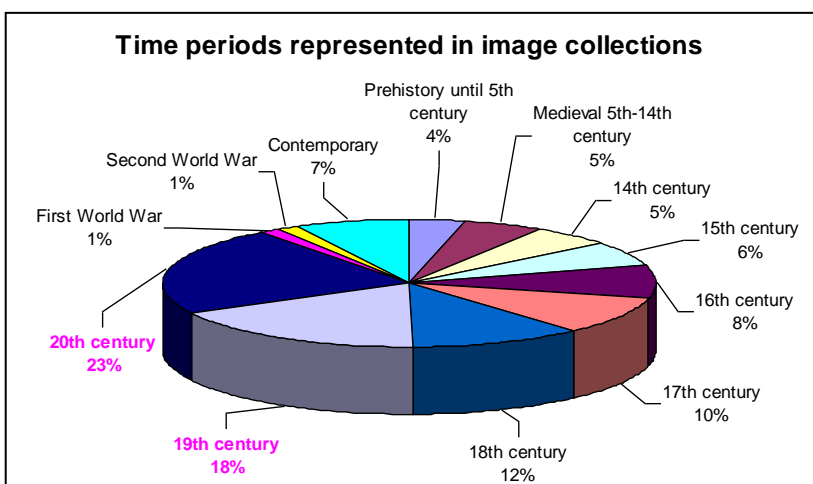

Figure 13 – Time periods represented in image collections
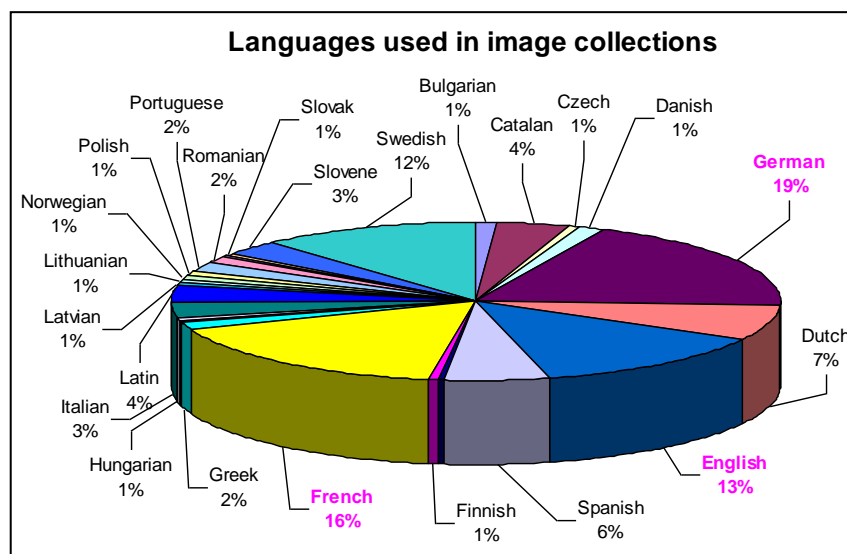
Languages used in image collections

Figure 14 – Languages used in image collections

The languages in image categories are more conventional than those in text collections, as to be expected. The most popular languages are the usual suspects, German, French and English, with Swedish following closely. Latin is represented quite often because of the large amount of historical maps in this collection, as well as coat of arms and treaties.

This can be broken down when looking at the three biggest categories within the image category and their top three of languages used:

		Amount of collections the language is used in
2.1 (art-)historical artefacts	German	19
	French	16
	Swedish	15
2.2 Photography	German	28
	Swedish	18
	English	11
2.5 Art	French	19
	English	13
	German	12

Table 5 – Languages used in main three categories in image collections

The top threes are very similar. The fact that Swedish scores this high is because of the many collections the Swedish Cultural Heritage aggregator has made available through Europeana. These consist mainly of photographic and art-historical collections (museum objects).

This same comes up when looking at the countries the content originates from; Germany, Sweden and France score the highest.
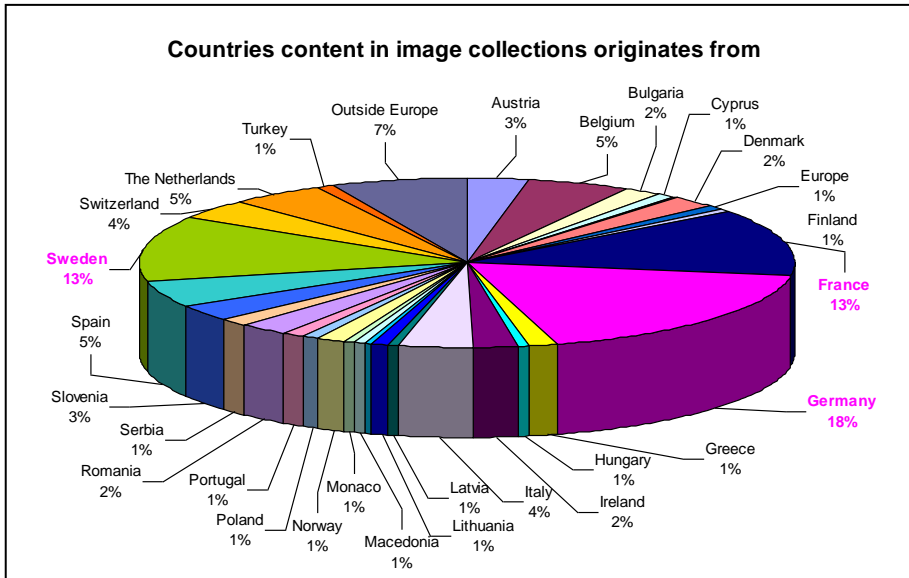
Countries content in image collections originates from

Figure 15 – Countries content in image collections originates from

2.4 Sounds

As was mentioned in the beginning of this report, the audiovisual collections are underrepresented in Europeana. As opposed to 165 collections with text and 251 collections with images, there are only 20 collections with sounds. Therefore, these aren't categorised in detail as text and image. The category is basically too small to make general conclusions out of.

Sound collections in Europeana can be divided in two: 'Music' and 'non-musical recordings'. The first can be divided up further into 'Folk', 'Jazz', 'Classical/opera' and 'Contemporary'. 'Non-musical recordings' can be divided up into 'dialects & accents', 'Wildlife sounds', 'Ethnographical recordings' and 'Radio'.
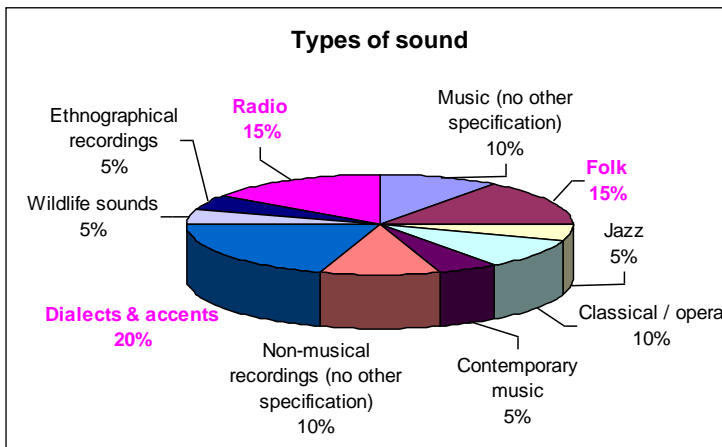
Types of sound

Figure 16 – Types of sound

The largest category is 'Dialects & accents', with audio content from 4 different providers and countries.

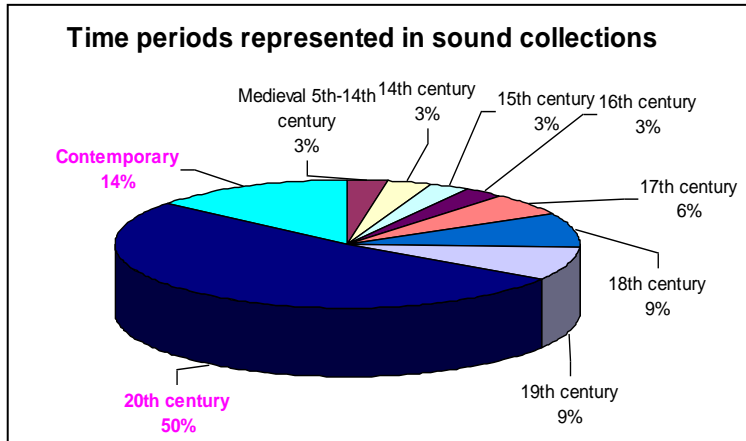**Time periods represented in sound collections**

Figure 17 – Time periods represented in sound collections

Of course, the main period the sound collections are from is the 20th century and after. The older centuries are also represented. These are representing the moment when the content was created in stead of when it was recorded, for example classical music.

The languages used – again – come down to the three usual suspects, German, English and French. The fact that German is used a lot in these collections can be traced to two Austrian sources (Österreichische Mediathek and Dismarc).

Countries content originates from in sound collections
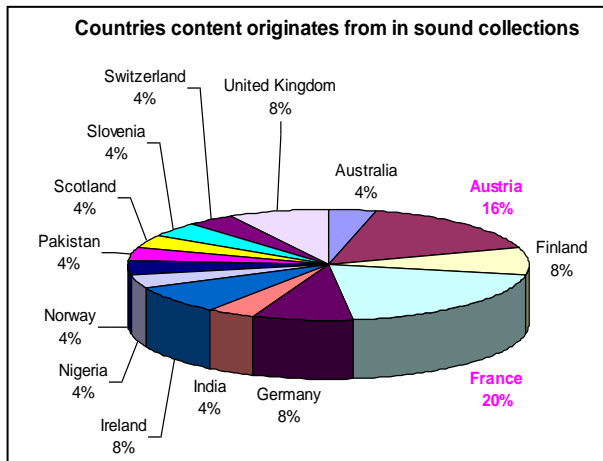
Figure 19 – Countries content originates from in sound collections

Languages used in sound collections
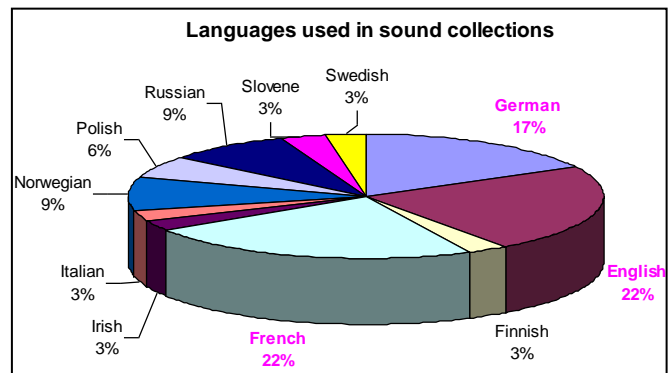
Figure 18 – Languages used in sound collections

2.5 Videos

Types of video

Figure 20 – Types of video

Again, as opposed to the good representation of text and image, there are only 13 collections with video material in Europeana. The main categories are 'TV broadcasts' and 'general recordings' with each 4 collections; however, this collection is also too small to make real general conclusions.

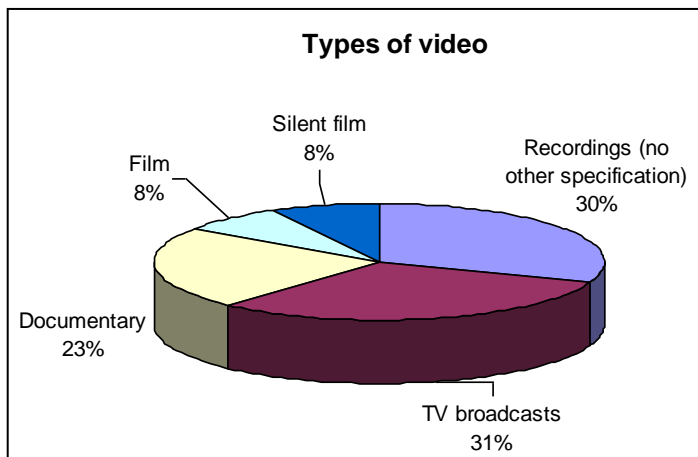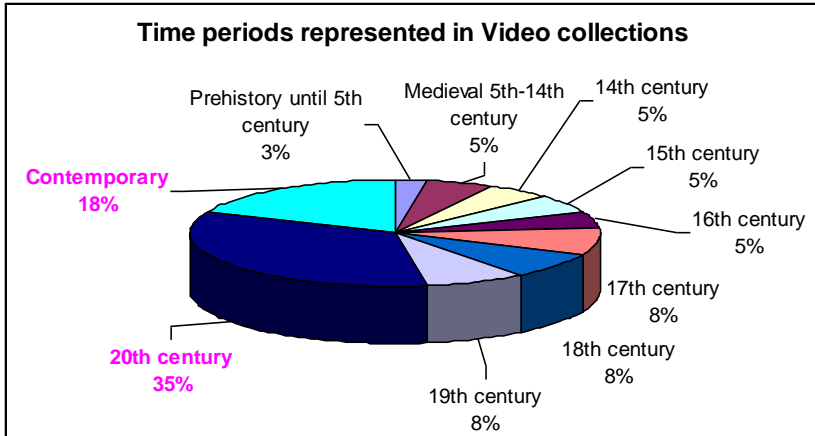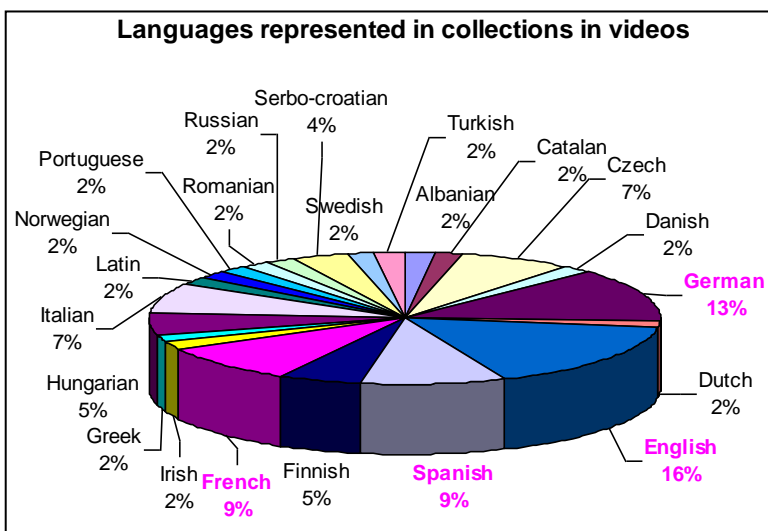**Time periods represented in Video collections**

Figure 21 – Time periods represented in video collections

Because of the nature of the material, video also mainly comes from the 20th century and after.

Languages represented in collections in videos

Figure 22 – Languages represented in video collections

Again, the biggest languages are best represented in this category: English, German, French and Spanish.
The great diversity in languages besides these four main languages can be traced down to three providers. These three are European projects, making available content from providers throughout Europe. These are the European Film Gateway, their predecessor Midas (Association des Cinémathèques Européennes) and VideoActive.

Countries material originates from in video collections

Figure 23 – Countries content originates from in video collections

This also explains why 'Europe' is mentioned as country content originates from – the collections from the Association des Cinémathèques Européennes and VideoActive are categorised as such.

3. Further research

The content analysis of Europeana is an ongoing action and therefore this document is fluid. It will be updated regularly. The next update is scheduled for spring 2011.

This analysis was done manually and within a certain timeframe, which made some flaws unavoidable (as described in paragraph caveat). In order to have a complete overview on what is available through Europeana there should be a system automatically keeping track of what is ingested and what is removed, continuously. The fact that this analysis was made based upon data from August 2010 makes it out of date from the moment the content ingestion team does something with the content: either ingesting new content or improving the already ingested content. This document therefore is already partly outdated. With the size of Europeana growing as rapid as it is, this is not feasible manually.

As was already mentioned in the first chapter under *caveat* the percentages used in this report are estimates, and relate to 'presence' in stead of a hard number, since it could only be investigated that in a certain collection there is material from for example, a certain country. Of course, the analysis would profit greatly if it could be based on actual numbers of items in a certain language, of a certain subject, etc.

Currently there is very little information available on the collections. An option would be to have content providers answer a questionnaire on their collection before submitting it and keeping this in a database. This could take the same shape as the collection descriptions The European Library maintains.

The categorisation that is used in this research could be further detailed and cleaned up, or mapped into one of the existing standards.

A second phase in this content research should be a gap analysis: it can be concluded from this research what is still missing in Europeana (depending on what the goal is) and what we should target in yearly goals in the Product and Service Plan.

Another next step is a correlation analysis with the user research done within Europeana, mainly the logfile analysis and results from Google analytics, to see what Europeana's users are actually searching for most.

Appendix A – Categorisation

Categories			
1.**TEXT**	1.1 Periodicals	1.1.1 Literature magazines	
		1.1.2 Music magazines	
		1.1.3 Other magazines	
		1.1.4 Historical Newspapers & Journals	
		1.1.5 Travel	
		1.1.6 Yearbooks	
	1.2 Archives	1.2.1 Finding aids	
		1.2.2 Estates of persons	
		1.2.3 Correspondence	
		1.2.4 Registers	
	1.3 Books & Articles	1.3.1 Manuscripts and rare books	1.3.1.1 illuminated
			1.3.1.2 diaries
		1.3.2 Literature & poetry	
		1.3.3 Philosophy	
		1.3.4 Religious / hagiographies	
		1.3.5 Watermarks	
		1.3.6 Book covers	
		1.3.7 Incunables	
		1.3.8 Ephemera	
		1.3.9 Technical	
		1.3.10 Astrology	
		1.3.11 Folklore & fairytales	
		1.3.12 Medical	
		1.3.13 Scientific / Biology	
		1.3.14 Travelling / etnografic	
		1.3.15 Economy	
		1.3.16 Immigration	
		1.3.17 Political	1.3.17.1 Pamphlets / propaganda
		1.3.18 History city / region	
		1.3.19 Music	
		1.3.20 Architecture	
		1.3.21 Printers	
	1.4 Music	1.4.1 Scores	
		1.4.2 Lyrics	
	1.5 Performing arts & film	1.5.1 Theatre	
		1.5.2 Film	
	1.6 Political documents		
	1.7 User Generated Content		
2. IMAGES	2.1 (art-) historical artefacts	2.1.1 Coat of arms	
		2.1.2 Charters	
		2.1.3 Treaties	
		2.1.4 Medals / coins	
		2.1.5 Correspondences	
		2.1.6 Maritime	
		2.1.7 Documentary	
		2.1.8 User-generated content	
		2.1.9 Itineraria	
		2.1.10 Postcards	

		2.1.11 Politics	2.1.11.1 Pamphlets
		2.1.12 Etnography / folklore	
		2.1.13 Religious	
		2.1.14 Military	
		2.1.15 Money	
		2.1.16 Clothes / fashion	
		2.1.17 Book binding	
		2.1.18 Descriptions of objects	
	2.2 Photography	2.2.1 Aerial	
		2.2.2 Etnography / folklore / travel	
		2.2.3 History / regional	
		2.2.4 Portraits	
		2.2.5 Archaeology	
		2.2.6 Military	
		2.2.7 Maritime	
		2.2.8 Nature & animals	
		2.2.9 Sport	
	2.3 Posters	2.3.1 Propaganda / politics	
		2.3.2 Commercial	
	2.4 Historical maps	2.4.1 Atlas	
	2.5 Art	2.5.1 Sculptures	
		2.5.2 Decorative & applied arts / design	
		2.5.3 Photography	
		2.5.4 Paintings	2.5.4.2 Portraits
			2.5.4.3 Watercolours
		2.5.5 Drawings / sketches	2.5.5.1 Nature & animals
		2.5.6 Engravings / old prints	
		2.5.7 Vedutas	
		2.5.8 Performing arts & film	2.5.8.1 Film stills/posters
		2.5.9 Illuminated manuscripts	
	2.6 Architecture	2.6.1 Prehistorical	
		2.6.2 Classical objects	
		2.6.3 Drawings / building plans	
	2.7 Technical		
	2.8 Archeology		
	2.9 Medical		
	2.10 Biology		
	2.11 Economy	2.11.1 Patents	
		2.11.2 Commerce	
	2.12 Estates / personal collection		
	2.14 Music	2.14.1 Instruments	
		2.14.2 Scores	
3. SOUNDS	3.1 Music	3.1.1 Folk	
		3.1.2 Jazz	
		3.1.3 Classical & opera	
		3.1.4 Contemporary music	
	3.2 Non-musical recordings	3.2.1 Dialects & accents	
		3.2.2 Wildlife sounds	
		3.2.3 Ethnographical recordings	
		3.2.4 Radio	

4. **VIDEO**	4.1 Recordings	4.1.1 TV broadcasts	
		4.1.2 Documentary	
		4.1.3 Film	
		4.1.4 News	
		4.1.5 Interview	
		4.1.6 Silent film	

Appendix B – Languages

cat	Catalan
bul	Bulgarian
cze	Czech
dan	Danish
deu	German
dut	Dutch
gre	Greek
eng	English
esp	Spanish
est	Estonian
fin	Finnish
fre	French
gle	Irish
hun	Hungarian
ice	Icelandic
ita	Italian
lav	Latvian
lit	Lithuanian
mlt	Maltese
nor	Norwegian
pol	Polish
por	Portuguese
rom	Romanian
slo	Slovak
slv	Slovene
sve	Swedish

Languages that are outside Europeana's language policy:

scc	Serbian
alb	Albanian
tur	Turkish
cop	Coptic
chu	Slavonic
sh	Serbo-croatian
ltz	Luxembourgish
heb	Hebrew
rus	Russian
lat	Latin

Appendix C – Tables (separate document)