assets.

assets.

D3.1.2 REPORT ON USER EVALUATION OF FUNCTIONALITIES

Advanced Search Services and Enhanced Technological Solutions for the European Digital Library

Grant Agreement Number: 250527

Funding schema: **Best Practice Network**

Deliverable D3.1.2 WP3.1

Change History

Version	Date	Status	Author (Partner)	Description
0.1	06-02-2012	Draft	Cristina Martinez (LABEIN)	First version
0.2	11-02-2012	Draft	Anne Marie van Gerwen (Europeana)	Revision and comments
0.3	13-02-2012	Draft	Cristina Martinez (LABEIN)	Changes in Planning, Tasks, Data Collection, Test preparation and Appendixes.
0.4	14-02-2012	Draft	Preben Hansen (SICS)	Comments, some additions
0.5	14-02-2012	Draft	Cristina Martinez (LABEIN)	Revision of comments. Figure numbers.
0.6	16-02-2012	Draft	Massimiliano Nigrelli (ENG)	Added a comment and revised the introduction.
0.7	09-03-2012	Review	Cristina Martinez (LABEIN)	Integrate contributions from reviewers: SICS (Oscar Täckström), EXALEAD (Amar Djalil Mezaour) and ATC (Nikos Sarris) Conclusions

| 0.8 | 26-03-2012 | Review | Anne Marie van Gerwen (Europeana) | Some general pointers and text changes for better readibility |
| 0.81 | 26-03-2012 | Review | Cristina Martinez (LABEIN) | Feedback for previous revision |

Table of Contents

1. Introduction

One of the objectives of the 'ASSETS for Europeana' project (hereafter 'ASSETS') is to improve the usability of Europeana. ASSETS services are tested on real end users (general public, academic researchers, learners, teachers, etc.) to ensure that they comply with user requirements and are easy to use. Evaluations had to be user-centric and have been conducted by an external contractor to ensure objectivity.

This document reports the activities involved in the user evaluation of the ASSETS project along the whole task, beginning with the plan, continuing with the tests and finishing with the conclusions. This document has been prepared within Task *T3.1.2 User evaluation activities*, and it relates to the *Evaluate Designs* step of the Human-Centred Design (HCD) process followed in the project for the design and evaluation of user interfaces.

Services to be evaluated, user scenarios and target groups are described in the *D3.2.1 Community Service Use Cases* document.

Other types of evaluations, which do not involve end users, are planned for each technical module. For example, acceptance testing is to be performed in *Task T3.1.4 Software Quality Control. ,*These evaluation activities are outside the scope of this document.

Apart from this end-user evaluation, there will be a separate evaluation of the professional services; those services equipped with a GUI will be evaluated by professional users, such as content providers, while the evaluation of those services equipped with APIs, will instead be performed by technical users (such as software programmers).

The above mentioned evaluations are not included in end user testing; their outputs are reported in the deliverables D2.4.4 *Service Testing* and D3.1.3 *Final Report of User Evaluations*.

In order to select the contractor for the end user evaluation, a tender was submitted by Europeana. The contents of the tender can be found in Appendix 1 .

Out of six proposals, User Vision[1] was selected as the contractor to conduct the usability testing of the ASSETS portal (http://assetsdemo.atc.gr/portal/) site.

The testing took place between the 29th of November and the 15th of December 2011 at three different European locations: Barcelona (Spain), Stockholm (Sweden) and Nicosia (Cyprus).

In the next sections we present an overview of the ASSETS services to be evaluated, the Evaluation plan, the performed Expert review, the Usability tests results and the main conclusions drawn from the tests.

---

[1] www.uservision.co.uk

2. End-user services to be evaluated

In this section we describe the ASSETS services to be evaluated and their respective graphical user interfaces (GUIs), including objectives, use case (sample of use that illustrate how the service can be tested in the ASSETS Portal) and target user groups (types of users that would be more suitable to use each service).

Also, some Storyboards have been created by the Service developers to test each of the services. They can be found in Appendix 3.

The end-user services that have been developed within ASSETS Stream2 and Stream3 and include:

- Query suggestions
- Metadata based ranking
- Image based similarity search
- 3D based similarity search
- Audio based similarity search
- Video based similarity search
- Video summarisation
- Semantic cross-linking and browsing
- Taxonomy based notification
- Preference query (Personalisation)

2.1 ASSETS Search services developed in Stream 2

Below is a brief description of the end user services developed in Stream 2.

2.1.1 Query Suggestions

Query Suggestions service	
Objectives	Provide a list of related queries for any submitted user query (historical information on the user interaction with the system will be used both to build and to validate the service).
Use Case	Given the text query "Gutenberg", a set of query recommendations are presented to the user as clickable links. One of them is the interesting recommendation "Gutenberg bible". By submitting this query we get a list of results relevant to that print.
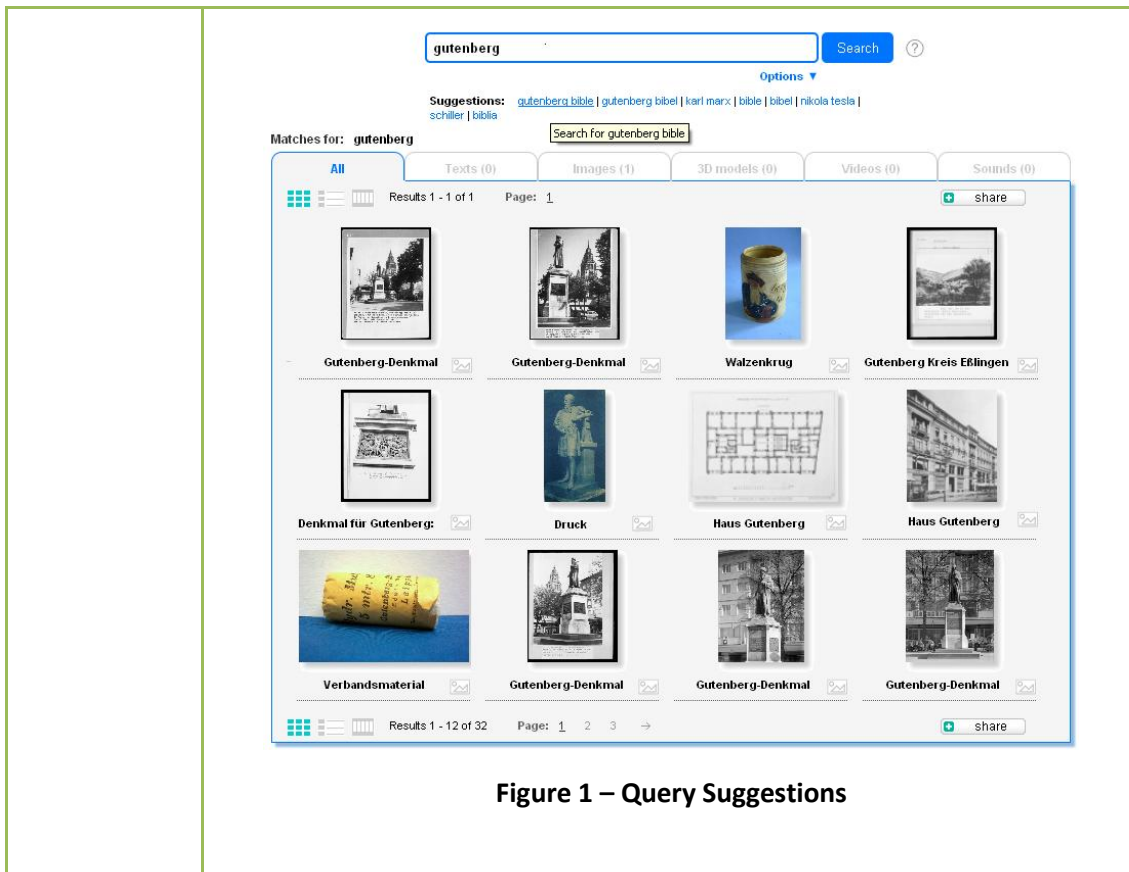
Figure 1 – Query Suggestions

Target user groups	This function works for users who are less likely to make detailed precision searches and prefer to search first and then filter/narrow down search results step by step.
Service Responsible	Claudio Lucchese (CNR)

2.1.2 Metadata based Ranking

Metadata based Ranking service	
Objectives	Provide a better, more relevant ranking of the results returned after a user query. We will compare the original results from Europeana against the result provided by ASSETS after applying the new ranking function.
Use Case	Go to the Demo page in ASSETS ThoughtLab: Based on the query "colosseo", the Europeana portal (on the left hand side) provides a list of videos regarding very specific events that happened near the Colosseum, but that do not provide an actual description of the Colosseum itself. ASSETS (on the right hand side) provides a list of documents all regarding the Colosseum itself, or a description of the Colosseum and other nearby buildings (top 2 results).

Figure 2 – Metadata based Ranking

Target user groups	Any type of user will benefit from this service.
Service Responsible	Claudio Lucchese (CNR)

2.1.3 Image based Similarity Search

Search Similar Images service	
Objectives	Search for visually similar images, allowing a user to find images based on other images. A link for searching similar images should be added near each returned image, independently from the type of query.
Use Case	The user first searches for "plato italiano" as text query and then selects one of the results as a query for searching similar objects with respect to the visual content. When the mouse is over the selected thumbnail a Tool zoom window is displayed. The toolbar button labelled '~' allows the user to make a similar search based on the selected image. **Figure 3 – Similar search button in the Zoom Tool** When the user clicks the '~' button the system displays a list of similar

	images. The image used for the similarity search query is displayed at the top of the results.
Figure 4 – Results of Similar search
A similar search can also be launched in the Details page, by clicking the "Similar search" link located at the Actions menu. |
| Target user groups | Finding similar images by similarity might still not be mainstream search behaviour. However, a similar feature does exist, e.g., in Google's image search.
Many younger users are using Google's image search and for those users this is a rather well established method. The question is in what context this pattern is used and what level of refinement is needed by the user. Searching for a windmill and then searching for similar images may be a rather straightforward search behaviour. |
| Service Responsible | Fabrizio Falchi (CNR) |
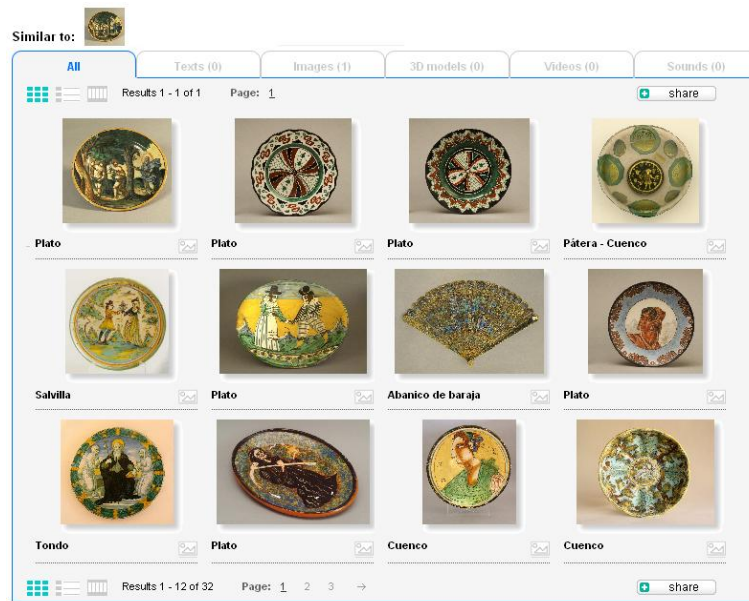
Upload and Search Similar Images service	
Objectives	Upload a locally available image or identify an image accessible through the web, as the base for a similarity search.
This facility is a modality of 'Similar Search' in which the input object for search is provided by the user, either by selecting a local file or by providing an URL address.
Use Case	An amateur historian is looking for images of medieval church statues. He already has a number of photos from Wikipedia and some that he has taken himself on his hard-drive. He wants to find images in ASSETS showing similar statues.
So he goes to ASSETS to upload example images of his own in order to see

whether the system has images of medieval church statues matching his interest. He's aware that ASSETS only has limited metadata and query translation capabilities and believes he will get better quality search results by comparing images rather than typing in queries.

First the user clicks the **Options▼** link. Then he clicks the "Upload and Search" link.
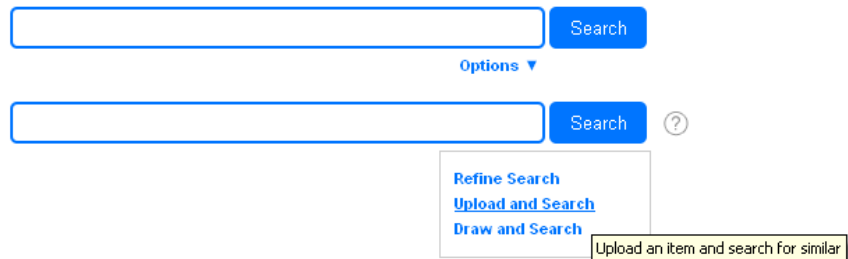
Figure 5 – Upload and Search link

Then the user can either select to upload a file (default) or enter a URL address:

- If the 'Upload File' option is selected then the user should click the "Browse" button. When the user selects a file in the dialog box and clicks the "Open" button, a similar search is launched and the results are displayed.

Figure 6 – Upload and search controls – Upload File option

- If the "Upload URL" option is selected then the user should enter the URL address in the edit box. When the user clicks the "Search" button a similar search is launched and the results are displayed.

Figure 7 – Upload and search controls – Upload URL option

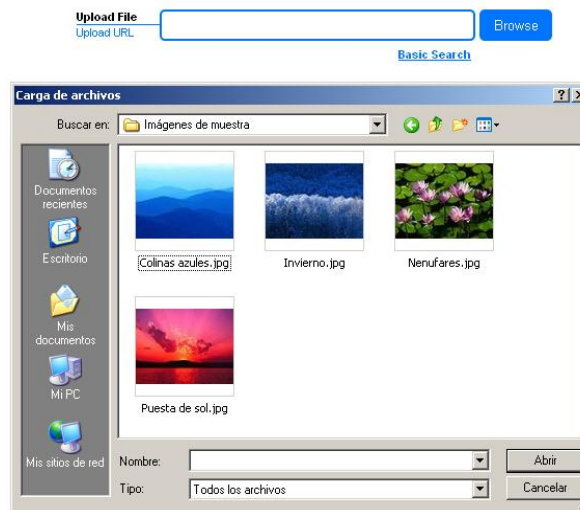|	
	Finally, the user could cancel the 'Upload and search' facility and get back to basic search by clicking the 'Basic Search' link.
Target user groups	The target user groups could be of several kinds requiring different levels of functionality and UI support: • First, at a basic level, a user wants to make an image search. Or a group of users, such as school children doing an assignment. For example, a person wants to see similar images of artist "X", "type of landscape". The user wants to use an image taken by her/his camera (e.g., from a vacation 6 months ago), downloaded from the web or a picture taken visiting a museum. • Second, at a more advanced level, a similarity image search could be performed involving more background knowledge on a certain item/context, such as an expert/specialist user (e.g. amateur historian) that is very familiar with the object and the context. This user group may look for a certain detail or characteristics in similar images. This user also knows how to judge the relevance of images in a larger result list, while in the first user group, they are satisfied with the first "best" image.
Service Responsible	Fabrizio Falchi (CNR)
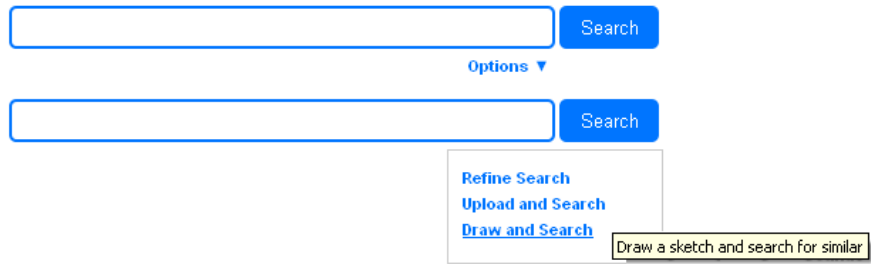
2.1.4   3D based Similarity Search
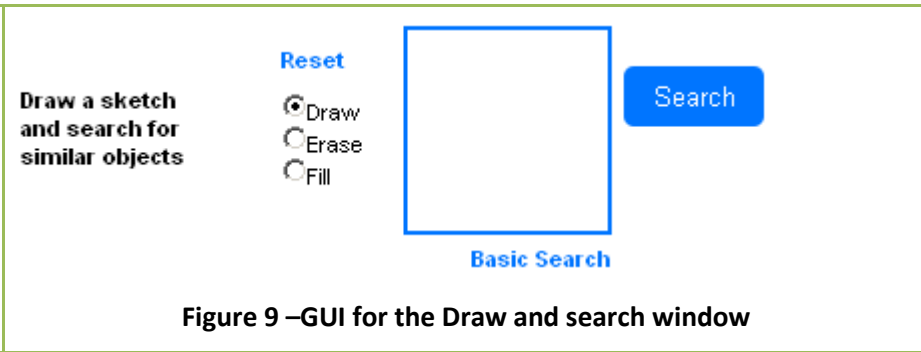
Search Similar 3D models service	
Objectives	Having found a 3D model in the system, search for visually similar 3D models that ASSETS could store and are not shown..
Use Case	An archaeologist searches for megalithic monuments (keyword = stadium) and finds a 3D model of a dolmen. As he is aware of the multi-lingual search limitations of Europeana and that each Europeana language has its own word for dolmen, he clicks the similar Zoom tool button rather than performing more text-based searches. ASSETS returns the results ordered by similarity level. The user interface for this scenario follows the same design pattern as the 'Search similar images' use case (2.1.3).
Target user groups	Searching for 3D models similar to existing ones must be considered a "power user/specialised user" behaviour. The user groups for this may be found in very specific professional technical (design of aircrafts and motorcars) and cultural domains (archaeology, architecture, films, etc.). A high level of domain knowledge and search knowledge is required.
Service Responsible	Michalis Lazaridis (CERTH)

Upload and Search Similar 3D models service	
Objectives	Upload a locally available model or identify a model accessible through the web, as the base for a similarity search. This facility is a modality of 'Similar Search' in which the input object for search is provided by the user, either by selecting a local file or by providing an URL address.
Use Case	An architectural historian has a 3D-model of a Roman triumphal arch and wants to see if there are similar models of triumphal arches in the system. He activates the search by upload function and uploads his model. ASSETS returns similar models ranked according to the degree of similarity. The user interface for this scenario follows the same design pattern as the 'Upload and Search similar images' Use Case (2.1.3).
Target user groups	Searching for similar 3D models by uploading existing ones must be considered a "power user/specialised user" behaviour The user groups for this may be found in very specific professional technical (design of aircrafts and motorcars) and cultural domains (archaeology, architecture, films, etc.). A high level of domain knowledge and search knowledge is required.
Service Responsible	Michalis Lazaridis (CERTH)

Sketch query and search service	
Objectives	Create a sketch and search for similar models. This facility is a modality of '3D models Similar Search' in which the input object for searching models is a sketch drawn by the user.
Use Case	First the user clicks the **Options▾** link. Then she clicks the "Draw and Search" link. **Figure 8 – Draw and Search link below the search box** The Draw and Search window opens, allowing the user to make a sketch from scratch. After clicking the "Search" button, the sketch is uploaded and the search similar function returns a list of similar models.

Figure 9 –GUI for the Draw and search window

Target user groups	Searching for similar 3D models by drawing a sketch may be considered a "power user/specialised user" behaviour, but is can also be used by most general interest users.
Service Responsible	Michalis Lazaridis (CERTH)

2.1.5 Audio based Similarity Search

Search Similar Audios service	
Objectives	Having found an audio file in the system, find more audio items that are similar based on audio characteristics.
Use Case	A user has found a piece of music she likes through a standard keyword search. She cannot identify the genre and the metadata does not state one or at least not one in a language that she understands. So to find more pieces of music similar to the one she has found she clicks the similar Zoom tool button.
	The user interface for this Use Case follows the same design pattern as the 'Search similar images' Use Case (2.1.3).
Target user groups	This search pattern is most common on popular on-line music sites, but is uncommon in general search engines as well as in specialised academic/research search engines. People using this search pattern may be frequent music listeners. either professionally or at home and often via a mobile device (smartphone).
Service Responsible	Oscar Paytuvi (BMAT)
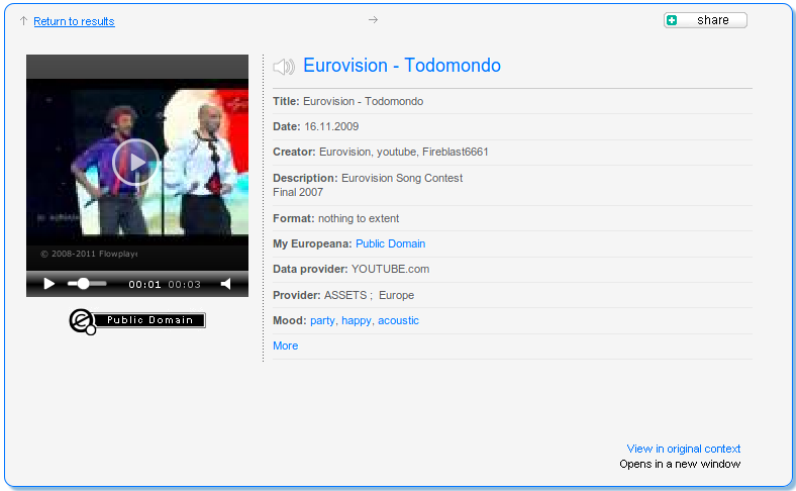
Upload and Search Similar Audios service	
Objectives	Upload a locally available music track or identify a music track accessible through the web, as the base for a similarity search.
Use Case	A user owns a piece of music she likes and she wants to find other music similar to it. So she uploads the music file to ASSETS to find more music like it.
	The user interface for this Use Case will follow the same design pattern as the 'Upload and Search similar images' Use Case (2.1.3).

Target user groups	Specialist users that may need a compiled list of similar music for a certain task (e.g. research). Some advanced search skills may be needed.
Service Responsible	Oscar Paytuvi (BMAT)

Find tracks by music descriptors service	
Objectives	Find music tracks by mood, tempo amongst other music descriptors, such as rhythm, key or timbre.
Use Case	A user finds a piece of music with a very happy and powerful sound. When she opens the **Details window** of that music item, its music descriptors tags will be shown on the right hand side of the window. The user can click the "happy" link to launch a new search based on that value. ASSETS returns audio items that have been indexed and tagged as "happy". **Figure 10 – Search by Audio Descriptor in Details page**
Target user groups	Both general users and researchers. This search pattern based on tags is common on popular on-line music sites, but is uncommon in general search engines or specialised academic/research.
Service Responsible	Oscar Paytuvi (BMAT)

2.1.6   Video based Similarity Search

Search Similar Videos service	
Objectives	Search for visually similar videos, i.e. provide solutions to allow the user to carry out searches of video content based on extracted video features. The input for the search can be some video found in the system or even a specific video shot.
Use Case	1. Search videos similar to a given video: The user interface for this Use Case follows the same design pattern as the 'Search similar images' Use Case (2.1.3).

	2. Search videos similar to a certain keyframe: The user is consulting the Details page of a video result he finds interesting. The user finds one of the keyframes shown very interesting, and wants to retrieve more videos with a content similar to it. He uses the similar button embedded in each video keyframe (see Details page picture below), to perform a similar search of videos based on the selected keyframe.
	Figure 11 – Details page for video
Target user groups	Both general interest users and researchers; the service can be considered as not demanding very strong search skills.
Service Responsible	Fernando López (UAM)

Upload image and Search Similar Videos service	
Objectives	Upload a locally available image or identify an image accessible through the web, as the base for a video similarity search.
Use Case	Search videos similar to a local keyframe uploaded by the user. The user interface for this Use Case follows the same design pattern as the 'Upload and Search similar images' Use Case (2.1.3).
Target user groups	Both general interest users and researchers; the service can be considered as not demanding very strong search skills.
Service Responsible	Fernando López (UAM)

2.1.7 Video Summarization

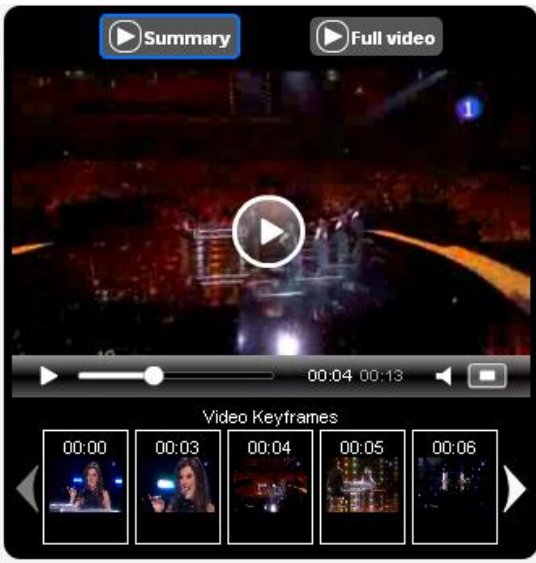| **Video Summarization service**	
Objectives	Provide condensed versions of complete videos (video summaries or video abstracts) for easing the users' browsing process. Video summaries will allow the user to overview the original video content
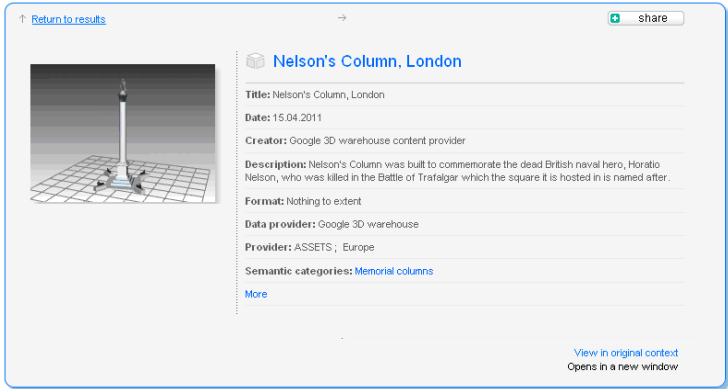
	without downloading/watching the complete original video.
Use Case	A university lecturer searches a the ASSETS portal and finds a number of early 20th century films that are of interest to him. In order to evaluate which films to include in his lessons he would first like to look through "trailers" or summaries of the films in order to assess their suitability. The user performs a search and opens the Details page of one of the video results. If the user clicks on the ▶ **Summary** button**,** the player will reproduce the video summary. If the user clicks on the ▶ **Full video** button**,** the player will reproduce the whole video. **Figure 12 – Playing video summary in Details page**
Target user groups	The Use Case fits anyone who would prefer to view a summary of a long video prior to actually viewing the whole or downloading the whole video. An informative video summary should enable a user to quickly get an overview of the contents of a video and thereafter decide whether the whole video is relevant. The users of this service may take three things into account when searching and judging for relevance: the audio (e.g. quality), visual aspects (e.g. single or multiple key-frames or sections in order to decrease uncertainty) and linguistic aspects. If the user is not knowledgeable in the language spoken on the video, the relevance may be judged from visual and audio features.
Service Responsible	Fernando López (UAM)

2.1.8 Semantic cross-linking browsing

Search by Similar Class service	
Objectives	The objective of semantic cross-linking is the establishment of semantic

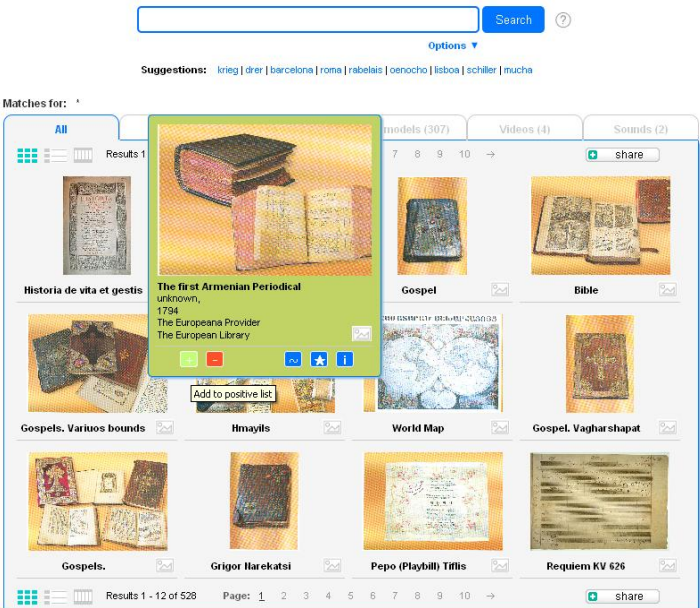|	
	links between pieces of information presented in different media. Semantic labels are assigned to the media objects, so that clusters of semantically similar objects are created. This way, the user can navigate faster through large amounts of media content.
	The user accesses the functionality by clicking on the category of a selected media file. Results of the same category are returned.
Use Case	An art historian is searching for images of a special type of column. First the user searches for "nelson column" as text query, then selects one of the 3D results (e.g. Nelson's Column, London) and opens the Details page. After having looked at the individual object she would like to find similar columns. The link "Memorial columns" exists under the Semantic categories, by clicking on this link the system returns several results, all of them belong to the "Memorial columns" category. **Figure 13 – Semantic browsing search in Details page**
Target user groups	Some advanced search skills may be needed here, since concepts like *class* or *semantic categories* could not be familiar to common users). This service is best suited for users with good digital and/or domain knowledge.
Service Responsible	Michalis Lazaridis (CERTH)

Relevance Feedback service	
Objectives	Relevance feedback techniques allow the user to search more interactively. After a search request, the user can mark the degree of relevance of the retrieved results, and perform a new search based on this criteria.
Use Case	The user performs a search and inspects the results. The user marks some of the results as more relevant (or less relevant) than the rest to the objects that s/he has in mind by clicking on the **+** and **–** toolbar buttons in the Zoom tool window.
	This feedback is used to refine the search and return more similar results. Refined results are presented to the user. The refinement can be repeated until the user is satisfied.

Figure 14 – Relevance Feedback

Target user groups	In terms of user behaviour this service implies both advanced search skills (knowledge of concepts like similarity and relevancy) as well as expert knowledge in a specific domain (to evaluate relevancy of results). However, relevance feedback may also allow the general interest user to interactively explore websites for a more serendipitous search experience.
Service Responsible	Michalis Lazaridis (CERTH)

2.2 ASSETS Community services developed in Stream 3

Below is a brief description of the end user services developed in Stream 3.

2.2.1 *Taxonomy Based notification*

Taxonomy-based notification service	
Objectives	The purpose of this service is to allow users to be notified when items of interest to them are added to the system. This is done by allowing users to define subscriptions by choosing terms from the ASSETS taxonomy. When new item is added, the taxonomy-based notification service compares the description (a set of terms from the ASSETS taxonomy) of the item with registered subscriptions by taking account of the hierarchy of terms in the ASSETS taxonomy. If a subscription matches the description, the service notifies the subscriber of the matching subscription.
Use Case	The user wants to subscribe to the ASSETS notification service for a specific category of items. After the subscription completes, he is told to check if any notification has arrived, and to check the suitability of them with the

subscription criteria he selected previously.

The user interface of the taxonomy-based notification service has been splitted into the "Subscriptions" tab and the "Notifications" tab.

In the "Subscriptions" tab, the user creates new subscriptions, modifies subscriptions and deletes subscriptions. To create and modify subscriptions, the user chooses terms from the ASSETS taxonomy though interactive term suggestion.
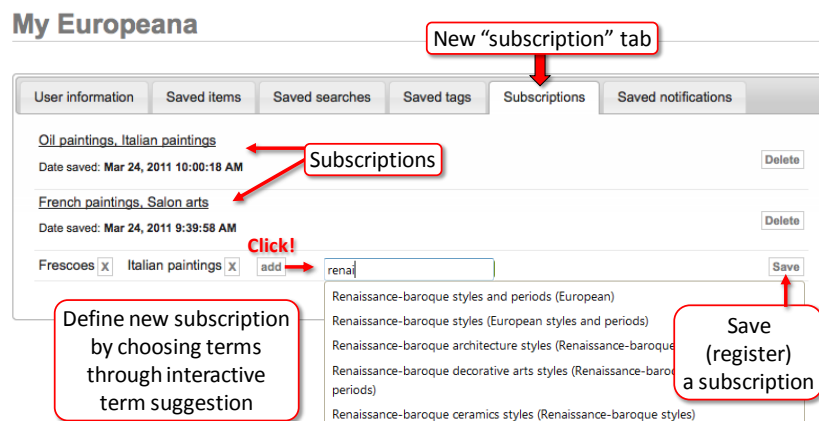
Figure 15 – "Subscriptions" tab

The "Notifications" tab shows the user the list of new items matching one of his subscriptions by taking into account the term hierarchy of the ASSETS taxonomy.
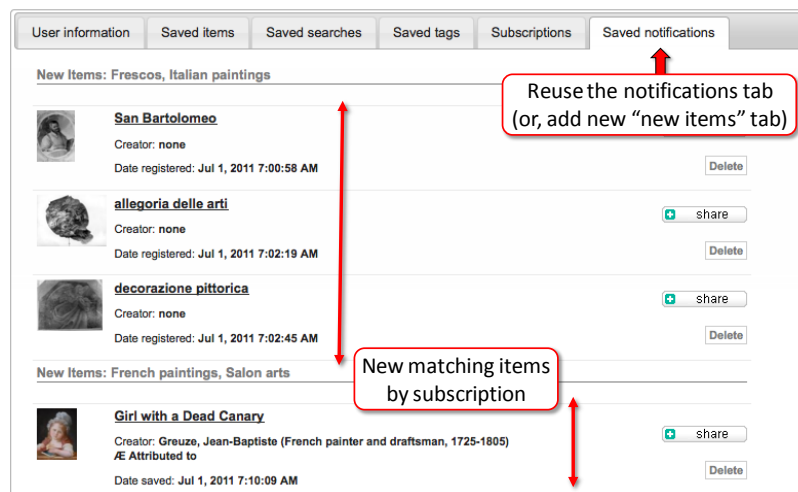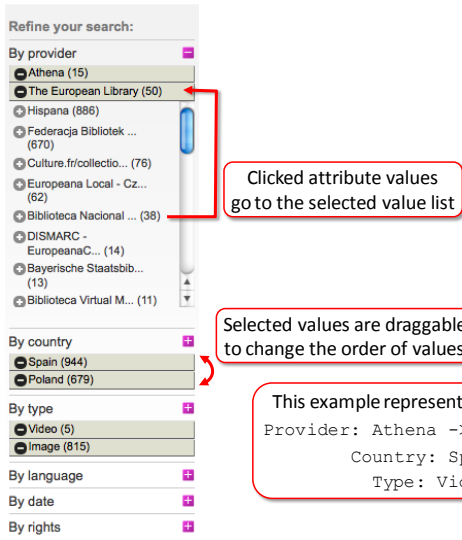
Figure 16 – "Notifications" tab

Target user groups	This function works for users who have the ability to express their interests by choosing appropriate sets of terms from the ASSETS taxonomy .
Service Responsible	Tsuyoshi Sugibuchi (UPS)

2.2.2 Preference query (Personalization)

Preference query service

Objectives	The purpose of the preference query service is allowing users to personalise search results by expressing their preferences. The users express their preferences by defining the ordering of the values for each attribute (Europeana facet).
Use Case	The user interface of the preference query service is implemented as an extension of the faceted query menu in the ASSETS portal. The Use Case of the service consists of the following tree steps. 1. The user starts a new search with an initial query as in an ordinary search task. For example, images of Van Gogh paintings. 2. The user defines preferences through the extended faceted query menu in the left column of the search result page. This extended menu allows users to (1) select values in each facet to filter search results (2) make "preferences" over selected values by dragging values to change the order of them. For example, select a specific year and a specific provider. 3. The user navigates in result blocks of ranked queries. The preference query service generates a list of ranked queries from preferences defined in the previous step. The search result page shows the first non-empty block of a ranked query at first. Then user moves forward / back to a different block by clicking navigation links. Figure 17 – **Outline of the preference query Use Case**

Figure 18 – **Close-up of the preference query user interface**

Target user groups	This function works for advanced users who frequently use the "advanced search" facilities of web search services.
Service Responsible	Tsuyoshi Sugibuchi (UPS)

3.   User Evaluation Plan

ASSETS aims to improve the usability of Europeana by developing, implementing and deploying services focused on search, browsing and interfaces. The objective of the user evaluation is to test the usability of (prototypes of) services developed by the ASSETS partners, which impacts the user experience and likelihood to use the service.

The usability study will look at two areas in particular, namely the user interface and the user's behaviour when interacting with the service/functions.

This section describes how the end user testing was planned and why. The ASSETS description of work specified some guidelines and a budget. The project team also defined the plan based on technical and financial feasibility and which method gave the most effective feedback on users' experiences, providing the ASSETS project with actionable user insights.

Based on the team's requirements and the Description of Work , a tender was submitted in Q2 2011 (see Appendix 1), resulting in 6 proposals. Important factors or considerations were:

- complying to the Description of Work (3 countries, different user groups, budget, timing)

- international experience

- experience with public sector, preferably cultural insitutions, library services and/or publishing

- responsiveness and flexibility in planning

After assessing the 6 proposals, it was concluded that User Vision's plan was most fit for purpose.

Founded in 2000, User Vision has a wealth of experience conducting usability testing with 15 employees purely focused on user experience, and an experienced management team that has led many high profile projects for private and public sector bodies around the world.

User Vision is especially familiar with the requirements of the public information sector, having worked on numerous Information retrieval (IR) projects including The English Heritage website, The National Library of Scotland, Historic Scotland, the British Library and the British Medical Journal. They have also undertaken many large international usability projects including projects for the European Union-Occupational Safety and Health Administration (EU-OSHA), Emirates Airline and HSBC.

The consultants assigned to the project had extensive experience testing in several European countries.

The web usability methods proposed by User Vision are well established and were suited for the purposes of this projects. User Vision also demonstrated a proactive and flexible approach to project management, ensuring that they could accommodate the project timescales at any moment.

These considerations made the team decide to work with User Vision to carry out this

project.

3.1 Location, types of end users and how many

The ASSETS DOW specifies, among other things, testing services with users from different backgrounds and countries in a laboratory setting environment.

The types of end users should broadly match the Europeana 'Personas' (see Appendix 1 - 8.9 Appendix I – Summary of Europeana Personas):

- Group 1: Learners (secondary education, students and teachers).

- Group 2: Academic researchers in the humanities.

- Group 3: Citizens with personal interest in cultural heritage, such as tourists.

Testing should take place in 3 different countries and each location will follow the same recruitment specifications. The locations Spain, Sweden and Cyprus were selected, to provide a good spread over European countries. Furthermore, In Sweden and Cyprus most citizens are proficient in English.

All users must be able to speak and read English fluently (or at least, be able to read English without any issues). The tool will only be presented in English (this was for cost reasons). All users must also be comfortable using the Internet and be familiar with using web search engines like Google or Yahoo.

A total of 54 users will participate in the tests, with 18 in each of the 3 countries. In each country 6 people will be tested from each of the 3 target user groups. This number came out as optimal as it allowed for extensive interviewing (1 hour per user) but still having sufficient quantity to base conclusions on. Next table shows the composition of the user group in each test session (quantity refers to the number of users).

Quantity	Participant type	Description
6	Learners (secondary education, students and teachers)	Users should have some relation to Culture/Heritage search. (i.e. Arts or History students or teachers)
6	Researchers	Researchers on Heritage or History monuments, graphics or documents.
6	General Public	Citizens with personal interest in cultural heritage, such as tourists

3.2 Evaluation plan and timeline

With User Vision, an evaluation process was designed that follows the steps defined in the following figure.

Stage	Activity	Output
Project Initiation Meeting (Physical meeting to discuss set up process)	Define and agree approach to entire project. This meeting to happen at one of the Europeana offices.	• Project Initiation Document **(PID)**
Expert review	We will start the project with an expert review, looking at the user journeys and user interface to pick up major usability issues	• Report highlighting usability and accessibility issues • Expertise advice on how to make the service more user friendly (based on extensive experience) • Annotated screen shots and prioritised recommendations based on the findings from the expert review
Subject Recruitment	Recruiting subjects for lab usability tests in three European countries. We will recruit all of the users based on the profiles provided by Europeana.	• Confirmed list of participants to take part in usability testing and online questionnaire.
Test preparation	Creating a test plan with pre/post test interview questions for review, refinement and approval.	• Full test plan
Usability Testing (User testing sessions to be held)	**Usability Testing** in France, Spain and Sweden (or whichever 3 countries you prefer – we have partners in several) to capture direct feedback from a broad range of representative users based on the actual usage of the prototype. Details of our proposed method are detailed in **section 2.**	• Ensure the prototype meets the aims and objectives outlined in the RFP with as diverse coverage as possible – see Recruitment Matrix in **section 2.3** • Detailed **findings and recommendations** highlighting the main user experience and usability barriers **illustrated by screenshots** • **DVD** of test sessions containing picture-in-picture recordings.
Preliminary Report	The **Preliminary Report** will give the Europeana team a good overview of how the project is progressing, and will be an opportunity to review the way the report and recommendations are displayed. Along with this, User Vision will provide Europeana with a high-level update report **after testing in each country.**	• Annotated screen shots from website • Recommendations • Status of project update
Final Report	The **Final Report** will give the Europeana team an overall project report summary and recommendations.	• Annotated screen shots from website • Recommendations • Detailed **findings and recommendations** highlighting the main user experience and usability barriers **illustrated by screenshots** • Excel spreadsheet of issues to track resolution post-delivery
Workshop and Presentation	A detailed high quality **workshop** and **presentation** covering the findings and recommendations.	• Workshop and chance to ask consultant **questions face-to-face** • Presentation to Europeana team and stakeholders, with the opportunity to ask questions about the findings.

Figure 19 – Evaluation steps

The timeline of the testing plan is shown in the next figure. The final plan has been definitely agreed at the Evaluation Kick-off meeting (Pisa, 14th September 2011) with a one-month shift, taking into account the final dates of services deployment defined by the ASSETS Service developers and the Stream 2 leader.
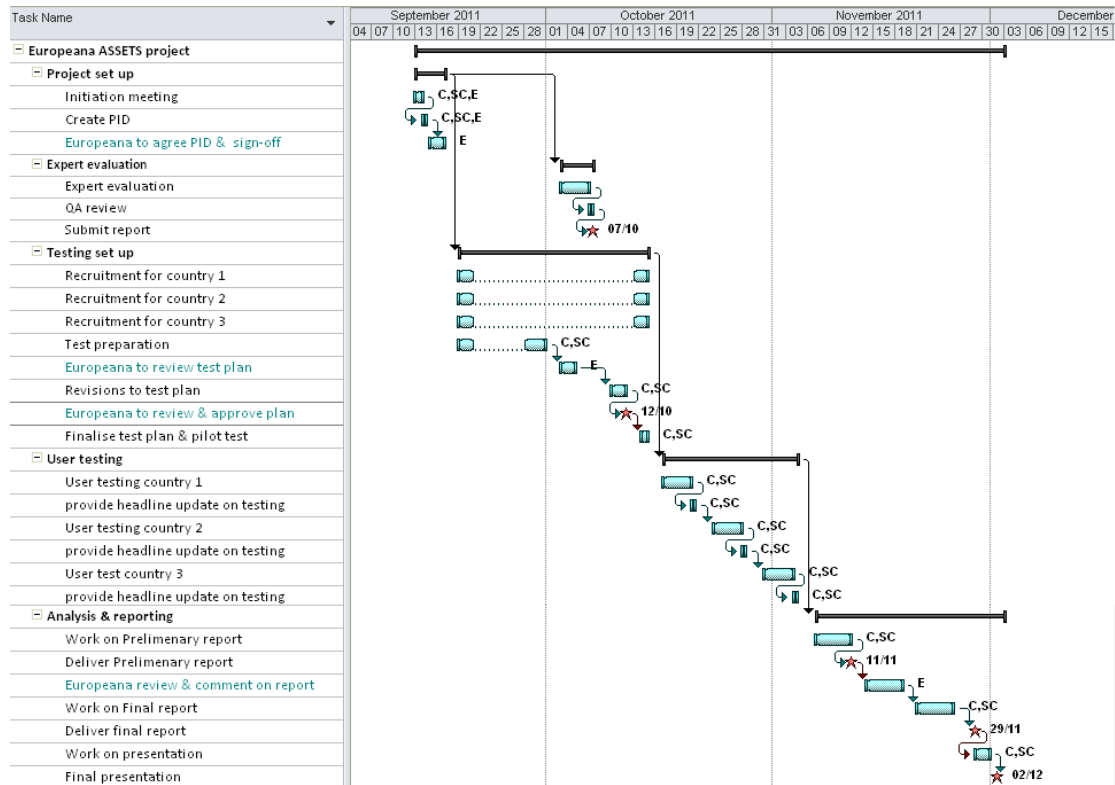
Figure 20 – Timeline of the testing

Abbreviations:

- C - User Vision Consultant
- SC - User Vision Senior Consultant
- E – ASSETS project team

3.3 Deliverables

The agreed way of project reporting is detailed below:

- **Expert review report** delivered in Power Point format with the main findings and recommendations found during the activity.

- **3 Headlines reports** delivered in Power Point format with the headlines of the testing in each country. These reports also include possible quick wins, to ensure that relevant points would be fed back to the ASSETS developers as quickly as possible.

- **Usability testing report** delivered in Power Point format with the main findings and recommendations found during the activity.

- **Power Point presentation** with the main findings, recommendations and wireframes with a summary of all the activities.

3.4   Task scenarios

User Vision developed some task scenarios based on the Storyboards provided by the Assets Evaluation project team (see Appendix 3: Storyboards), who also reviewed them to ensure they were complete and accurate before testing commenced.

Tasks for the General Public and Learners:

Task	Instructions	Service
T1	Find content related to Versailles. Once you get the results, find information about someone famous who lived in there.	Query Suggestions
T2	Find images for "Leopold II" of Belgium. Once you get the results, find similar images to the image you like the most. Upload the lighthouse picture that you have on the Pictures folder of the computer and search for similar images	Search Similar Images Upload and Search Similar Images
T3	Listen to a song named "Give my love to Rose". Find other songs with the same musical style (without using the similar button)	Find tracks by music descriptors
T4	You are interested in viewing one of the Eurovision song videos. Could you find as well a shorter version of the same video? You want to jump to half way through the video. How would you do this? Can you find other videos that contain images similar to one of those in this video?	Video Summarization Search Similar Videos
T5	Hand draw a shape and find a similar 3D model	Sketch query and search

Tasks for the Experts:

Task	Instructions	Service
T6	Google "Sir Charles D'Oyly" and get an image you like. Get the URL of that image and search similar images in the Europeana site.	Upload and Search Similar Images
T7	You want to find 3D models of memorial columns. Could you find other columns similar to the one you like the most? Search for other 3d models that correspond with this type of monument.	Search Similar 3D models Search by Similar Class
T8	You would like to find a similar audio file to the Johnny Cash song you have in the Desktop folder of the computer.	Upload and Search Similar Audios

	How would you do it?	
T9	You are interested on finding documents and monuments related to "Chopin". Refine your search results by editing your preferences.	Preference query
T10	You are interested in portraits from the Italian Renaissance-Baroque and want to know when future content of this type is added to the site. Create a search alert to notify you when new content is added to the database. You want to limit your results to only watercolour portraits. Edit your search query accordingly.	Taxonomy based notification
T11	You are interested in column-shaped 3D monuments. Once you get the results, refine your search by marking which results are relevant and irrelevant in your opinion	Relevance Feedback

The tasks were created according to the skills of each one of the user type groups. Each participant had time to attempt approximately 4 tasks during the hour. In order to ensure adequate coverage of all the tasks, each participant would attempt 4 randomly selected tasks as shown below. When time allowed, they could attempt further tasks after the initial four had been completed.

	Learners						**Experts**						**General Public**					
	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**	**P14**	**P15**	**P16**	**P17**	**P18**
T1	X	X	X	X											X	X	X	X
T2	X	X	X	X	X	X							X	X				
T3	X	X	X	X	X	X							X	X	X	X	X	X
T4	X	X	X	X	X	X							X	X	X	X	X	X
T5				X	X								X	X	X	X	X	X
T6							X			X	X	X						
T7							X	X			X	X						
T8							X	X	X			X						
T9							X	X	X	X								
T10								X	X	X	X							
T11								X	X	X	X							

3.5   Data Collection

Data will be collected through observation and the collection of objective test metrics. The collected metrics are described below.

3.5.1   Quantitative information

After each task, we collect how successful the task has been on a scale between 0% and 100%, based on the following criteria:

- 100%: Participant completed the task to the agreed success point on the first attempt with little or no moderator intervention
- 75% - 25%: Participant completed the task with increasing levels of moderator intervention and/or increasing attempts
- 0%: Participants failed to complete the task successfully, had to be guided to the end point or expressed that they would have given up if carrying out the task on their own.

3.5.2   Qualitative and observational information

Qualitative information is collected through the users providing verbal commentary whilst using the site. This provides a very good insight as to why users are performing various steps, what they are thinking at each step, and the issues they encounter. Any issue of interest or concern arising in the evaluation will be probed further for additional detail and clearer understanding. In addition observational data is collected n users' performance by recording, for example, any errors and likely causes, and indicative signals from their body language.

3.5.3   Subjective ratings

After each task, the user will be asked to provide a rating on a 7-point Likert scale to judge the difficulty in using the site. This scale is illustrated below:

Using this site, how easy or difficult was it to find the information you were looking for?

Very easy	Easy	Slightly easy	Neutral	Slightly difficult	Difficult	Very difficult
1	2	3	4	5	6	7

Overall, how relevant or irrelevant do you think the results given by this task were?

Very relevant	Relevant	Slightly relevant	Neutral	Slightly irrelevant	Irrelevant	Very irrelevant
1	2	3	4	5	6	7

How likely or unlikely would you be to use this feature?

Very likely	Likely	Slightly likely	Neutral	Slightly unlikely	Unlikely	Very unlikely

1	2	3	4	5	6	7

This will also be an opportunity for the user to summarise verbally their thoughts and to discuss any difficulties experienced, and to explain why a particular rating was given.

3.5.4   Pre-test Interview

The pre-test interview assesses the expectations and any previous experiences in using the site. It provides an excellent opportunity to gauge the user's previous experience/knowledge of similar services, any concerns they have had and the types of information they would be likely to need. This helps to create a more rounded picture of the participant and can highlight potential issues before commencing the test. *The pre-test questionnaire can be found in **Appendix 2**.*

3.5.5   Post-test Interview

After completing all the task scenarios in the usability test (or as many as time allows if they cannot complete all tasks), users will be asked to complete a short interview about their experiences with the ASSETS prototype. This will further help to identify and clarify likes, dislikes, any difficulties encountered and the overall experience in relation to their expectations and any previous experience.

They will also be asked to complete an "Emotional Response Questionnaire" to establish how they felt about their experience using the site. *The post-test interviews can be found in **Appendix 2**.*

4.   Heuristic evaluation

An expert (or heuristic) evaluation of the ASSETS for Europeana website has been carried out well before the actual user testing started, ensuring that some major bugs or quick wins were identified in time. The focus was on the usability of the site and the quality of the user journeys. In addition, some accessibility issues were also noted for consideration.

Expert evaluations do not involve testing with real users, but comprise a structured evaluation of the site with reference to a set of recognised 'best practice' principles (or 'heuristics') for user interface design. These principles are based on extensive research and the experience of leading usability practitioners to determine what constitutes a usable system. Applying them in a structured analysis quickly reveals the usability defects.

For each usability issue identified, advice is provided on how best to minimise or eliminate the issue. Also provided is an indication of the severity of impact for site visitors and hence the urgency of the need to implement this advice. Usability comments are classified in one of four categories shown below. The distinctions between these are not strict, and should be interpreted as such.

Category	Definition
Positive (P)	**Positive finding** - This approach is recommendable and should be preserved.
High (H)	**High Severity problem** - Likely to cause a problem that frequently prevents users from completing their task or will greatly annoy them. It is Important to fix these issues soon.
Medium (M)	**Medium severity problem** - Likely to cause some temporary confusion that delays users from reaching their goal with the site for some considerable time (30 seconds – 1 minute), and causing some degree of frustration/irritation, but unlikely to prevent them from reaching their goal.
Low (L)	**Low severity problem** - Cosmetic or minor problem that may cause users to hesitate or be distracted for a few seconds.

Accessibility issues were not specifically addressed during this evaluation. However, any that were immediately obvious during the course of the evaluation have been included as part of the report.

Due to the nature of expert evaluations, generally many more usability problems than positive issues are reported. This does not mean there are few good points within the site itself; rather it is a reflection of the focus of the evaluation, which is to find problem areas in order to propose solutions. Good points, by their nature tend to be less noticeable, as is the nature of good design, and require no solution.
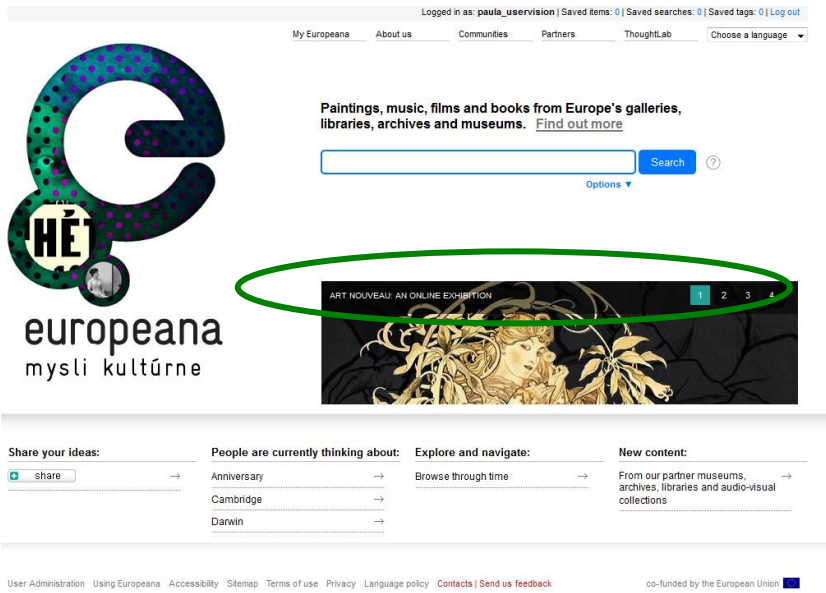
4.1   Usability Issues and Recommendations

The following tables outline all the findings of the usability expert evaluation performed over the ASSETS portal (http://assetsdemo.atc.gr/portal/).

Since the Heuristic evaluation was performed just before the first user testing, it was not feasible to implement any changes to the portal following the recommendations of the Expert. However, most of usability issues identified in the Heuristic evaluation were addressed after the first evaluation session that took place at Spain:

- " ~" button has been replaced by 'similar' in Zoom tool and Details page
- 'Suggestions' has been replaced by 'Related Searches'. The terms have been highlighted with a background.
- 'Options' below the search box has been replaced by 'More search tools'. 'Enter URL' has been added to the menu. 'Refine search' and 'Select skin' have been removed.
- The tooltips of the relevance feedback buttons have been replaced by 'More like this' and 'Less like this'.
- 'Audio descriptors' has been changed to 'Audio categories' in the Audio Details page
- 'Video keyframes' literal has been inserted above the video keyframes of the details video page. The '~' button on each keyframe has replaced by a "similar" button, that is displayed only when the mouse is over the keyframe.
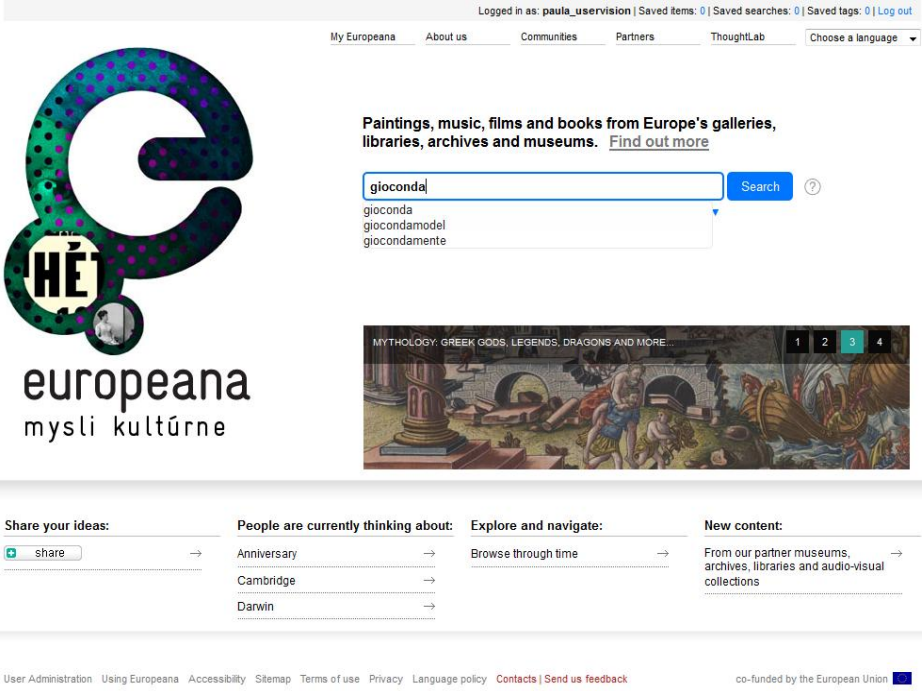
As result of these changes, an improvement of the usability of the corresponding services was observed in the next evaluation sessions, at Sweden and Cyprus, specially the usability of the 'Similar' functionality was greatly improved (see Figure 22 - Average success per task - General public and learners).
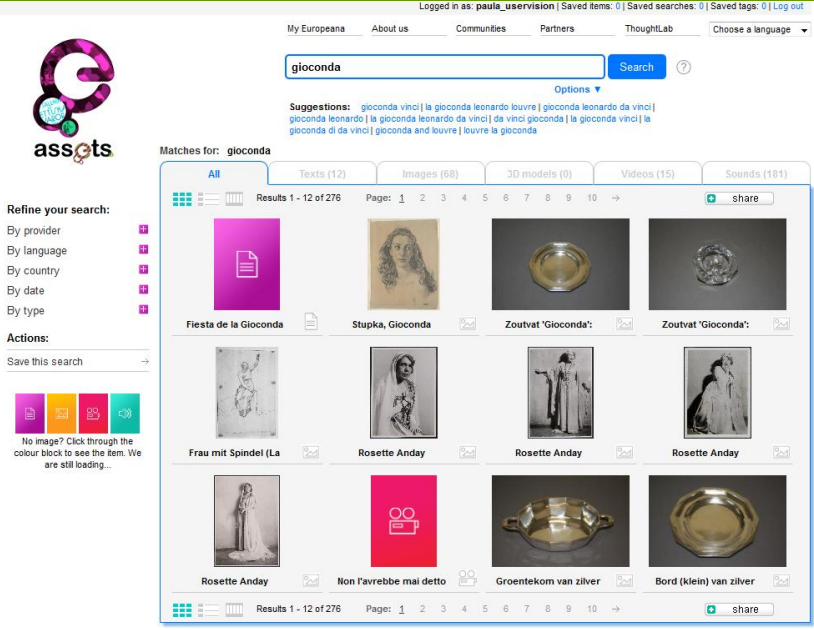
4.1.1   Homepage

| Usability Issue	Severity
Good introduction for content (P) The site includes a very good introduction that allows to manage the user expectations.	P
Slightly distracting carousel (L): The carousel on the main page could cause issues as the headers are too small and the colour contrast makes it difficult to read. This could raise confusion about its purpose and operation.	L

Usability Issue	Severity
Recommendation: Consider increasing the size and colour contrast of the carousel items' headers to ensure its visibility and avoid any confusion.	

4.1.2 General Search

Usability Issue	Severity
Good use of suggestions for searching (P) The dropdown of suggestions is good practice and recommended for accessibility reasons and is particularly useful for visitors with dyslexia. 	**P**
Different display options for the results (P): Giving the users the possibility of changing the display is very useful and consistent with other technologies. Furthermore, the icons are self-explanatory and easy to understand.	**P**
Recursive link on Pagination (M): In the pagination links at the foot of the page, the current page appears as an active link, while the other pages appear to be disabled. This is likely to create a lot of confusion. Additionally, this is an issue of accessibility as users using assistive technologies could understand that they are not on the correct page.	**M**

Recommendation: Swap the styles adopted here so that the current page appears grey and is not clickable, while the other pages are a higher contrast colour and underlined (i.e. looking more like hyperlinks).

Ideally the other pages should have the same look and feel as the current page currently.

Use of tabs for distinguishing different types of content (P): The use of tabs for displaying different types of content is considered very good practice.	**P**
"Sounds" is not the best descriptions for content type (L): The use of "Sounds" is not as broad as "Audio". **Recommendation:** Consider changing the wording to "Audio" to avoid any potential confusion.	**L**
Poor colour contrast (H): Currently the site has various areas with very poor colour contrast. Apart from being an important accessibility issue, it may appear to users as thought the content is disabled. **Recommendation:** Increase the colour contrast to improve its visibility.	**H**
Inconsistency on breadcrumbs (M): Although the use of breadcrumbs is very positive, the wording used needs to be consistent and meaningful for the users. The 'metadata' style wording used here is unlikely to be meaningful to most users.	**M**

Recommendation: Change the "TYPE:IMAGE" for the tab header to "Images".

Close-up is shown when mouse is over but not on focus (M): Having a close-up of the item when the mouse is over is a very useful feature, but it needs to be made accessible by replicating this behaviour when using the keyboard only. **Recommendation:** Replicate the functionality for on focus event.	**M**
Icons are not self-explanatory and sometimes missing (H): The "Save to my Europeana" and "Find similar" icons are not self-explanatory and will rely on users checking the tooltip to understand its purpose. Furthermore, when the icons are not available, they are not shown as disabled to the users. This creates inconsistency which makes it more difficult for the user to learn how to interact with the system. **Recommendation:** Consider changing the Similar icon for a link with the "similar" word on it replicating Google system or find a more suitable option. Change the wording on the "Save to my Europeana" tooltip for "Add to my Europeana Favourites". All icons need to be visible, although disabled (or greyed out) when not available to click.	**H**
Redundant function (M): The "i" icon duplicates the behaviour of clicking on the main image. As clicking on the images is an intuitive action in this context, the "i" button is redundant and therefore only generates visual clutter on the page. **Recommendation:** Consider removing the "i" icon/button.	**M**
Related content and actions do not appear related to the main content (H): The proximity of two types of content indicates association. Therefore, the further away two pieces of content are, the less they will be associated together. In this example, the current location of "related content" and "actions" do not look related to the main content as they are visually separated and outside of the visual context of the item detail.	**H**
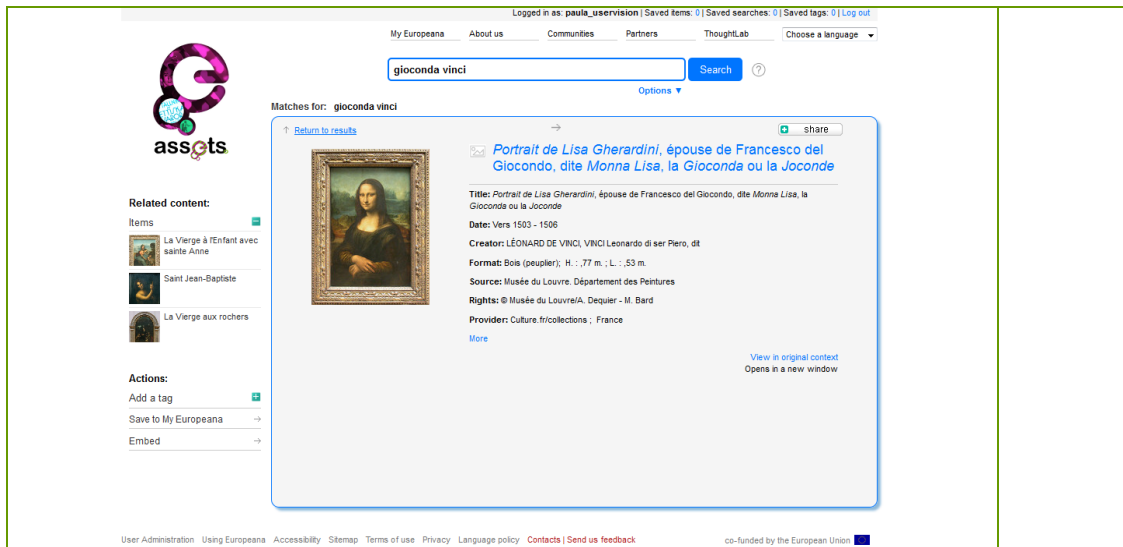
Recommendation: Move the related content and actions to inside the blue box for better association.
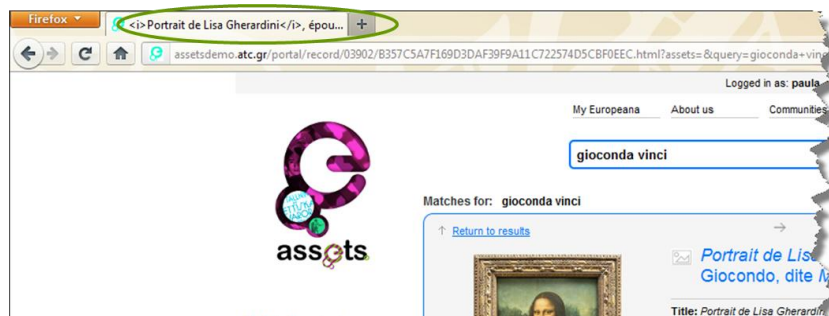
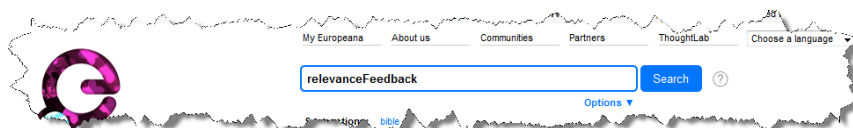|	
Left hand side menu inconsistency (M): On the content details, the left hand side menu changes from the default and is therefore not consistent with the rest of the site. It is important to maintain consistency when using the site to understand where the user is. **Recommendation:** Maintain the same functionality of the left hand side menu as the rest of the site to improve its consistency.	**M**
Title includes HTML tags (H): Some content titles includes HTML tags as part of the title reducing its readability and creating an accessibility issues for users with screen readers. **Recommendation:** Ensure that the site removed all HTML tags before displaying the text for titles, etc.	**H**
"relevanceFeedback" is displayed on the search text field when relevance rating is provided (H): Once the user has provided the relevance ratings, the search field shows "relevanceFeedback", having this query visible can affect the users trust on the site as they have not introduced it. **Recommendation:** Ensure that the search box is only ever populated with terms that have been entered by the user. Any system commands such as this should	**H**

be hidden from the user to avoid any confusion and to increase the users trust on the site.

4.1.3 Image search similarity

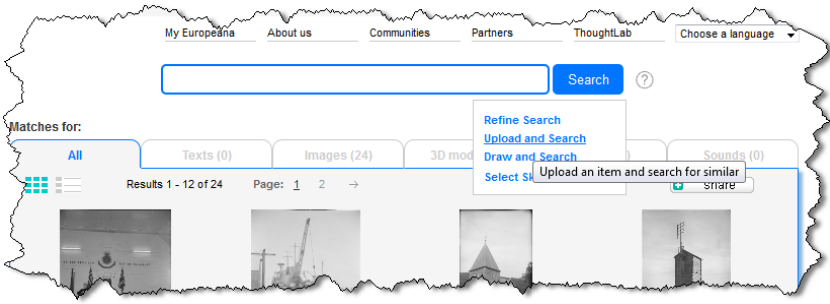| Usability Issue	Severity
Upload and search does not relate to search for url (M): The method for performing an image search is currently unintuitive. When users are searching by url, they need to click to "Upload and Search" and "Upload a URL". The wording used here is not intuitive for this task. **Recommendation:** "Upload a URL" should be changed to "Insert a URL" and be on the first level menu. See the example from www.tineye.com on the next page as a good example of clear distinction and intuitive language.	M
Meaningless query is shown when searching for similar images (H): Finding similar images command should not be visible for users to avoid confusion. **Recommendation:** Ensure that the search box is only ever populated with terms that have been entered by the user. Any system commands such as this should be hidden from the user to avoid any confusion and to increase the users trust on the site.	H
Missing "similar to" image (M): When users get the results for the similar images, the actual (source) image is missing. This is inconsistent with other ASSETS services. **Recommendation:** Display the image the user has searched for to help finding similar images and to ensure consistency across the site.	M
Results are not similar to the image in the URL provided (H) In this example, a URL to an image of the Eiffel Tower was provided (below). However, it is not clear why the results provided have been returned for this image.	H

Recommendation: Provide a clear explanation as to why particular results have been returned.

4.1.4 Similar/Upload search results

Usability Issue	Severity
Missing similar image or uploaded image (L): Users were not sure if the images shown on the similarity by relevance feedback or by uploading were taken into account. Furthermore, they missed the original images for comparison purposes.	**L**

Recommendation: Ensure that the image that is used to generate similar results is always shown in the "matches for:" area for comparison purposes.

4.1.5 Audio Search

Usability Issue	Severity
Playing the audio file as part of the website is very good (P): Having the embedded functionality for playing the audio as part of the site is good practice and creates a smooth user journey.	P
"Audio descriptions" is inconsistent with "semantic categories"(H): The site uses two different names for the functionality that allows users to find similar content by clicking on descriptors. This inconsistency is likely to prove confusing for users. This is particularly true when it has the same look and feel, as this does. Elsewhere on the web, both "audio descriptors" and "semantic categories" use a more generic term: "tags." **Recommendation:** Consider joining both functionalities under the same generic name (i.e. tags, descriptors) to maintain consistency and to avoid any confusion.	H

4.1.6 Video search and summarisations

Usability Issue	Severity
Video keyframes look very much like related videos (H): The incorporation of selected keyframes at the bottom of the video could be confused with related videos. The principle of adding related videos as thumbnails overlaying the video screen is common practice on video sharing sites, such as YouTube (shown).	H

Recommendation: The keyframes require clear description as this is not a common concept. To remedy this we suggest three changes:	
1. Move the keyframes from out of the video panel (to avoid associations with related videos 2. Give the keyframes section a clear title 3. Consider using a visual metaphor to describe the frames in the video, e.g. a film strip.	

4.1.7   3D search

Usability Issue	Severity
Drawing space is too small for detailed drawings (H): The drawing space is too small to allow the user to draw in any detail, thus reducing the chances of accurate results being returned. There is also no possibility of increasing it. **Recommendation:** Increase the size of the drawing canvas to allow a more detailed drawing. Furthermore, add the ability to increase the size of the canvas if needed.	H
"Reset" option should be located close to the other drawing related calls to action (M): All drawing related actions are located to the left of the canvas, meanwhile "Reset" is located under "Search" on the right. This location does not build association. **Recommendation:** Move the "Reset" option to the left of the canvas to improve the natural grouping.	M

4.1.8   Relevancy identification

Usability Issue	Severity
Accessibility issue for relaying only on colour identification (H): The meaning of a particular element should never be communicated by colour alone as this excludes users with visual colour impairments from understanding and interpreting the results. **Recommendation:** Add some text or icon that identifies positive and negative relevance.	H
Missing instructions or context for relevance buttons (M): The current relevance buttons do not have any headers or instructions. As this feature is not very common it is likely that this functionality will not be well understood by many of the users. This problem is compounded by the fact that the tooltip is not very explanatory - "Add to positive list" or "Add to negative list" may not have meaning for the users. **Recommendation:** • "Relevant" icon is removed to simplify process for user • Star icon is replaced with more commonly used yellow "favourites" icon	M

• Similar image icon is replaced by clearer text link (see page 15) • Redundant info icon is removed (see page 15)	
Unnecessary step (H): The relevance feature could be significantly simplified by removing irrelevant steps right away. This would avoid the need to colour the images, and then pressing search again – which is not an intuitive action for users. **Recommendation:** Simplify the process by adopting immediate removal of irrelevant results, i.e. clicking on the 'irrelevant' icon removes that entry from the list (see example overleaf).	H

4.1.9   Taxonomy based notification service

Usability Issue	Severity
Query suggestions (P): Including a list of suggestions for subscriptions is very good practice for usability and accessibility reasons. Users don't have to spell the words and a more limited choice for topics.	P
Look and feel inconsistency (M): The style of the tabs and buttons of the subscription service does not follow the style of the rest of the site. This could create trust issues of the services. **Recommendation:** Ensure that all the functionalities offered by the site maintain the same look and feel.	M
Good selection of filters (P) The selection of filters for the faceted search is very good and offers the user the flexibility needed for this type of search engine.	P
Easy to miss function (M): The "Save this search" link looks so similar to the rest of options that is easy to be missed. **Recommendation:** Change this link to a button in order to highlight it. As a general rule, actions should be represented as buttons, while hypertext represents navigation links.	M

5.   User Evaluation

Usability evaluations took place between the dates of the 29th of November and the 15th of December in three different European locations: Barcelona (Spain), Stockholm (Sweden) and Nicosia (Cyprus).

- Barcelona (La Salle University, UserLab laboratory[2]): 29th November- 1st December.

- Stockholm (Usability Partners offices, Usability Lab[3]): 5th - 7th December

- Nicosia (Pulso Lab): 13th -15th December

Finally, a total of 51 participants took part in them, giving equal representation to market segments identified by the ASSETS project. All participants were offered an incentive of 50€/500SEK to participate in the testing.

The breakdown of actual participants is shown in the next table. The most relevant details for each of them can be found in Appendix 2.

No. of participants	User Group	Group description
20	Experts	Researchers on Heritage or History monuments, graphics or documents.
15	Learners	Users who are fluent in English (or if they are Spanish, able to understand written English) that have some relation to Culture/Heritage search. (i.e. Arts or History students or teachers)
16	General Public	Citizens with personal interest in cultural heritage, such as tourists

The usability tests have been conducted with one subject at a time, lasting 60 minutes in total. Although focus groups can produce highly rated user insights, one on one testing usually means users are not influenced by the group, which allows every user to give their own option. It also means that every user will get time to speak. One on one usability testing is also far more realistic since the manner in which people will encounter the website once it is launched is more likely to be on an individual basis than as a group of people to collectively navigate.

The usability test compromised three steps:

- Pre-task interview, to assess the expectations and any previous experiences in using the site.

---

[2] http://www.salleurl.edu/portal/departaments/home-depts-DTM-infraestructura-userlab

[3] http://www.usabilitypartners.se/about-us/usability-lab

- Site testing with a series of pre-agreed tasks that have been were provided for all participants depending on their group type.

- Post task interview/emotional response questionnaire, a short interview about their experiences using the site.

During the usability tests empirical information has been collected on the task performance and each subject's opinion of ease of use, as they 'think aloud' to perform the task.

The test facilitator has recorded their actions with screen capture and Picture-in-Picture recordings showing both the user and the screen they are using, and taking notes throughout.
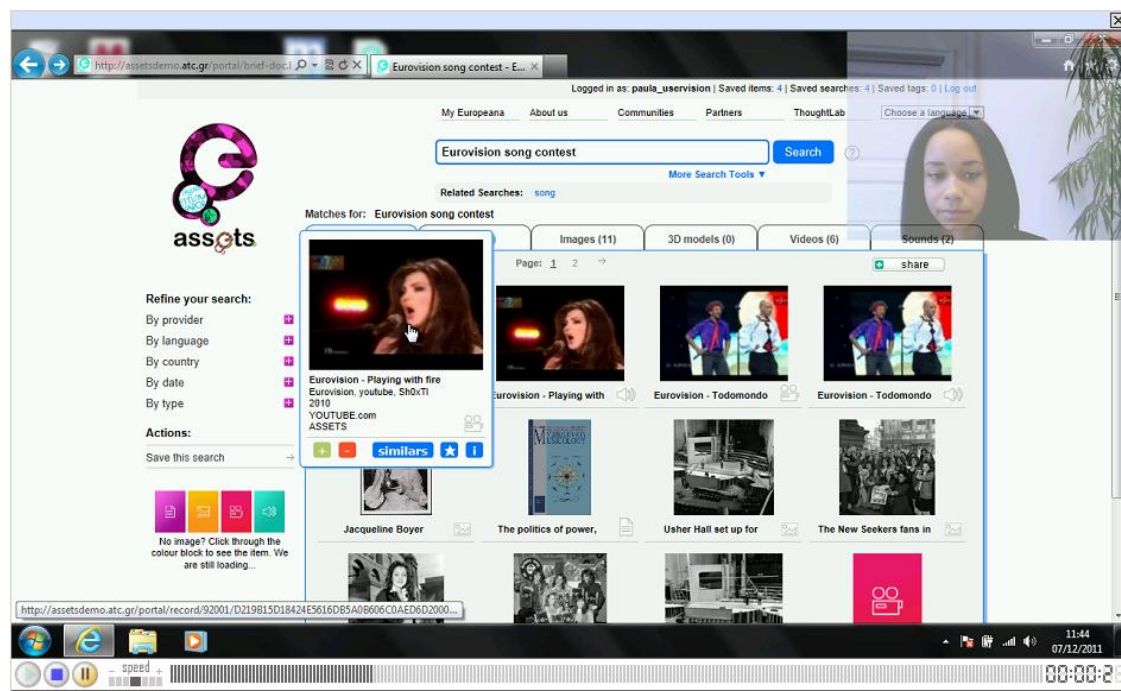

Figure 21 – Picture-in-Picture recording

The ASSETS evaluation project team has attended as observers the test sessions at Barcelona and Stockholm. This has allowed us to discuss some usability findings directly with the User Vision consultants and to raise awareness of any technical or other constraint in relation to the recommending solutions.

5.1 Tasks performed by General Public and Learners

5.1.1 Task Completion for General Public and Learners

Most participants completed all 4 tasks in the given time. The overall task success rate was 35% (total tasks completed to the agreed end-point for all participants with no moderator intervention)

- 177 tasks attempted over all 20 participants
- 62 tasks (35% of total) completed successfully (100% success)
- 28 tasks (16% of total) completed with partial success (25/50/75% success)

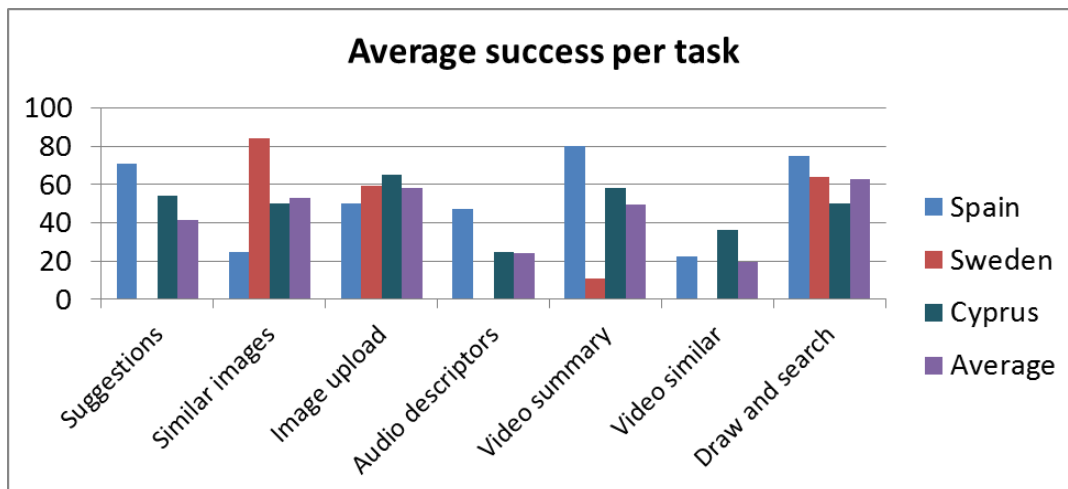- 87 tasks (49% of total) were not successfully completed (0% success)

Figure 22 – Average success per task (General public and learners)

This graph shows that the most successful tasks were "Similar images", "Image Upload" and "Draw and Search", having all of them above 50% success rate.

The least successful tasks were "Audio descriptors" and "Video similar" having less than 25% task success.

Some tasks were more successful in other countries than others. For example the "Similar images" and "Image upload" noticed a big improvement with the new designs, wording and locations.

The detailed description of the issues encountered by the users and actionable recommendations for resolving these issues have been gathered in Section 5.3 Usability Issues and Recommendations.

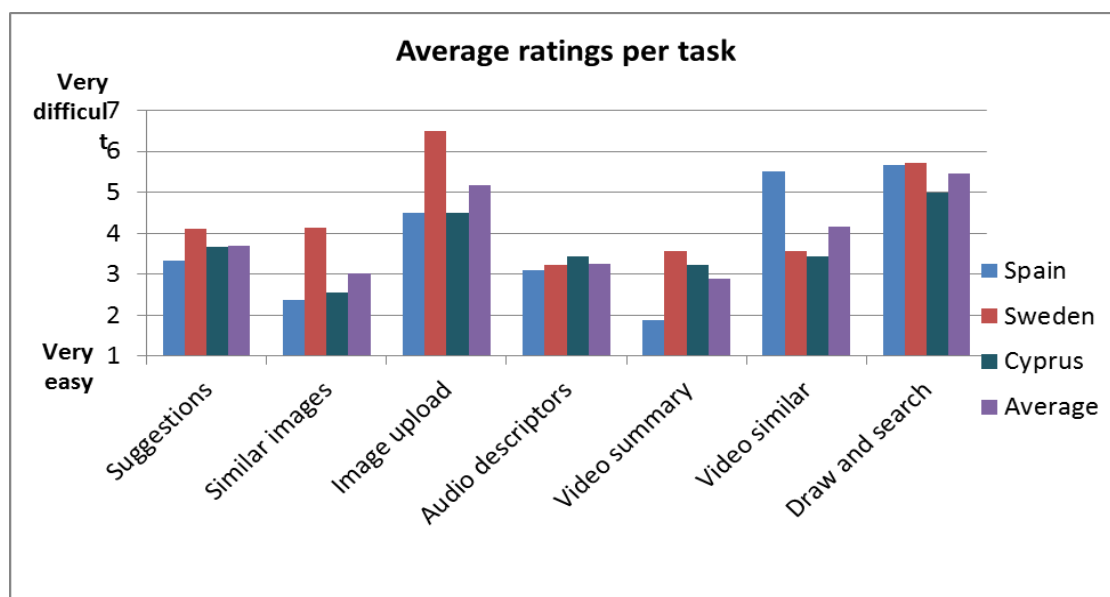### 5.1.2 Subjective Ratings for General Public and Learners

Figure 23 – Average ratings per task (General Public and Learners)

This graph shows that "Image upload" and "Draw and Search" had the worst ease of use ratings, due mainly to it not being evident on the site how to get to that functionality and being unfamiliar functionality.

"Similar images" and "Video summary" had the best ratings although users mentioned that they did not expect that functionality on the site.

The next figure shows a comparison between how the users perceived the difficulty of a task once formulated (expected rating per task) and how they assessed the difficulty of that task after its execution (rating per task).
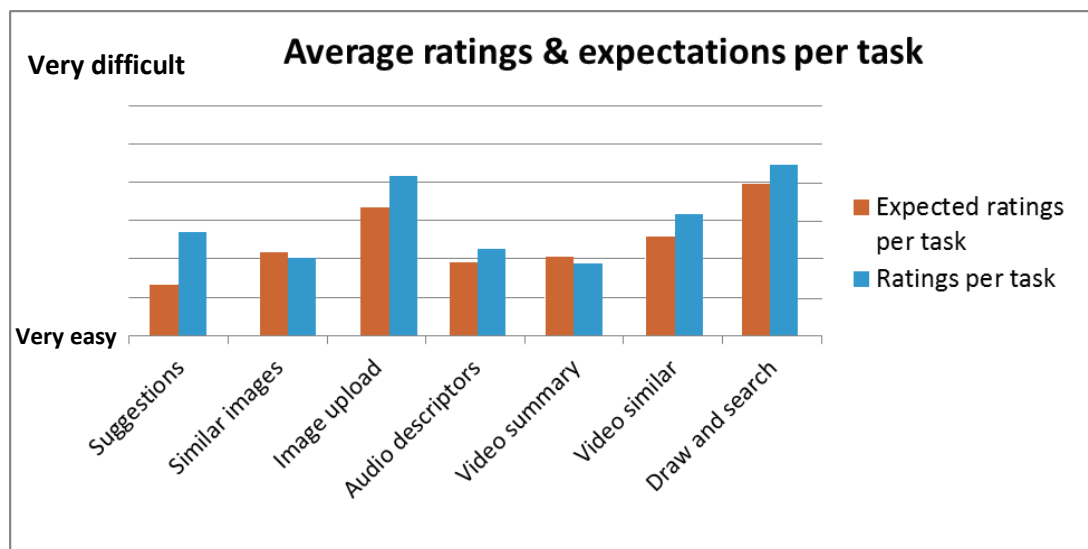

Figure 24 – Average ratings & expectations per task (General Public and Learners)

Most tasks were considered more difficult than expected, with only "Similar Images" and "Video summary" slightly easier than expected.

The biggest difference with expectations was on "Suggestions" and "Image upload".
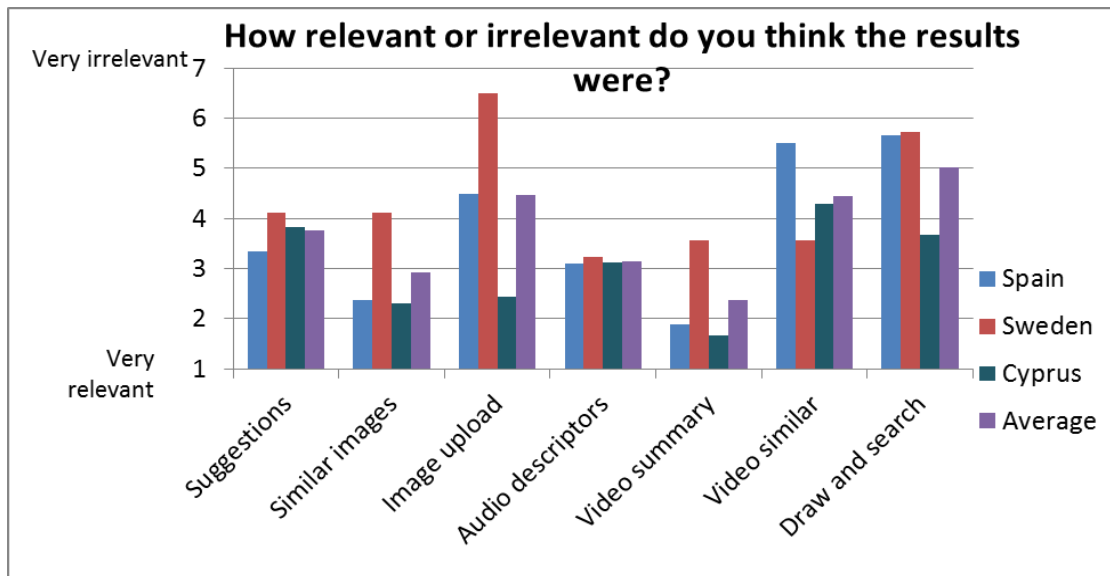
Figure 25 – Relevancy of the task results (General Public and Learners)

On average the worst results for relevancy were for "Image Upload", "Video Similar" and "Draw and Search". With regard to the 'similar' function; users mentioned several times that they were not sure what criteria was used to define the similarity and that affected their opinion of the relevancy of the results.

The most relevant results, according to the users, were for "Video summary" and "Similar images".

Figure 26 – Likeliness to use the service in the future (General Public and Learners)

Users commented that they did not expect to find most of this functionality on the site but if they were aware of it, they would probably use them.

The worst rating for likeliness of use were for "Video similar" and "Draw and Search". Users mentioned that those services were not something that they could see themselves using in the future.

5.2 Tasks performed by Expert users

5.2.1 Task Completion for Experts

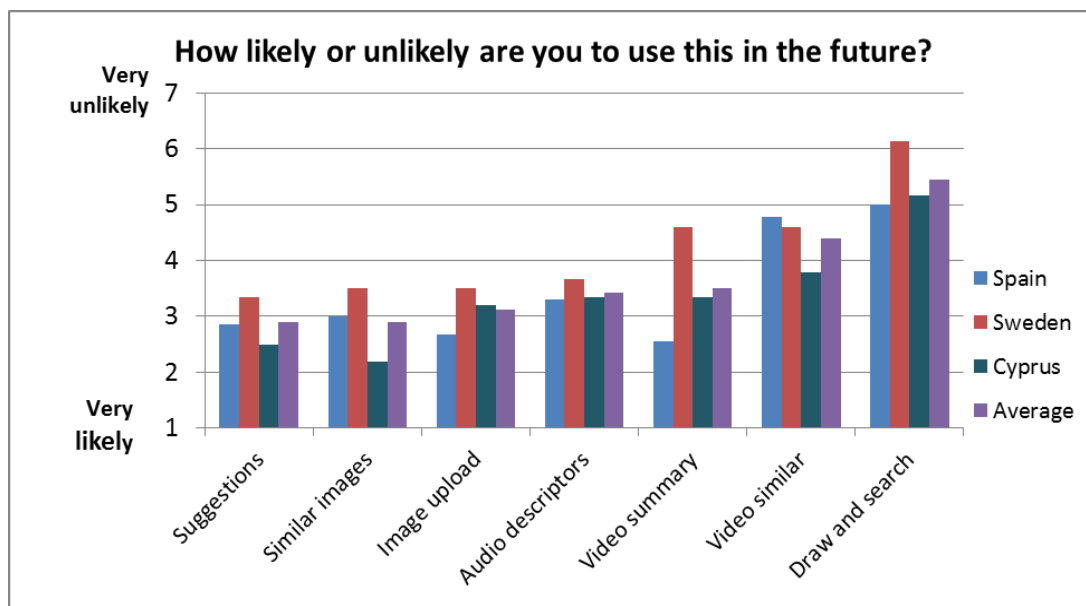Most participants completed 4 tasks in the given time. The overall task success rate was 26% (total tasks completed to the agreed end-point for all participants with no moderator intervention)

- 84 tasks attempted over all 20 participants

- 22 tasks (26% of total) completed successfully (100% success)

- 19 tasks (23% of total) completed with partial success (25/50/75% success)

- 43 tasks (51% of total) were not successfully completed (0% success)
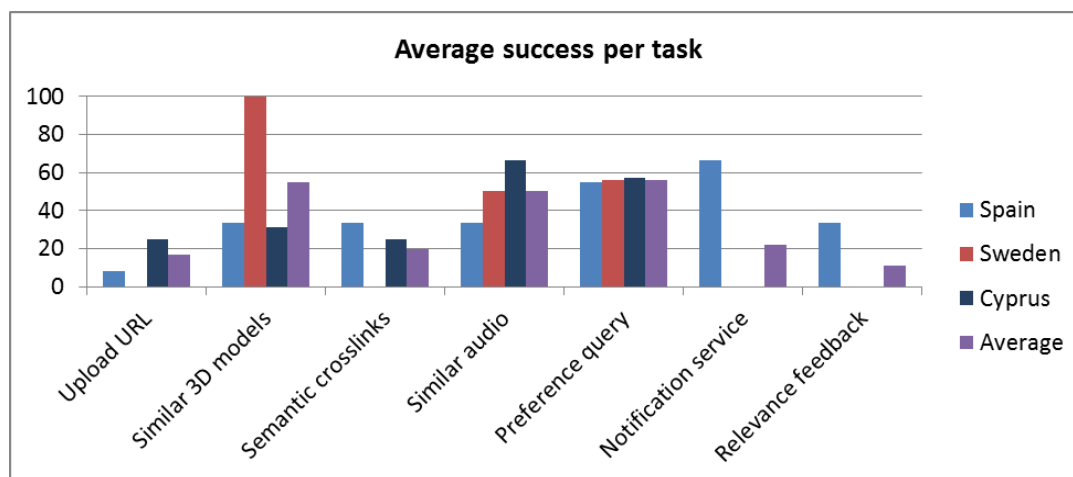
Figure 27 – Average success per task (experts)

This graph shows that the most successful tasks were "Similar 3D models", "Similar audio" and "Preference query", all of them having 50% or greater task success and the least successful tasks were "Upload URL", "Semantic Crosslinks", "Notification service" and "Relevance feedback" having less than 25% task success.

"Notification service" and "Relevance feedback" were particularly unsuccessful with some countries not having any success at all.

The detailed description of the issues encountered by the users and the actionable recommendations for resolving these issues have been gathered in Section 5.3 Usability Issues and Recommendations.

5.2.2 Subjective Ratings for Experts
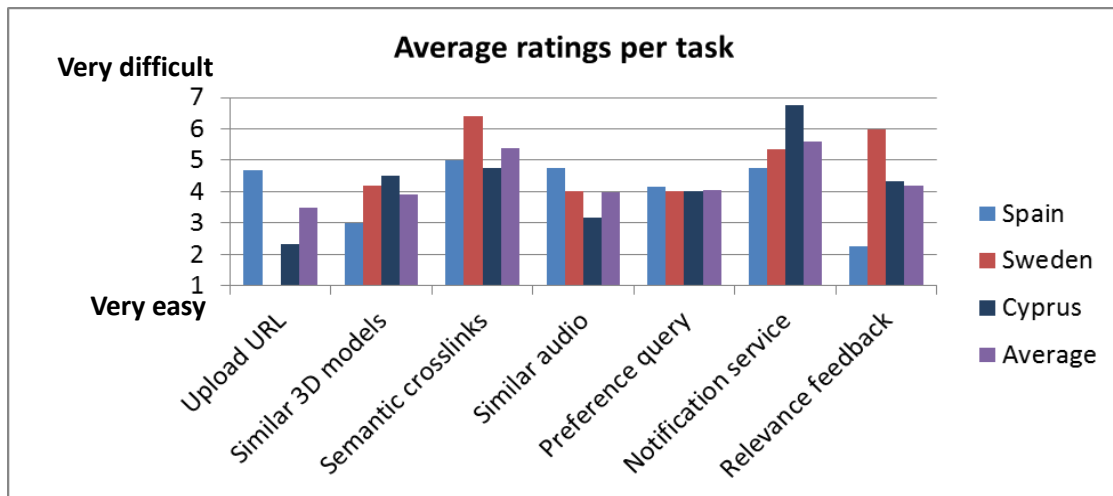
Figure 28 – Average ratings per task (experts)

This graph shows that "Semantic crosslink" and "Notification service" had the worst ease of use ratings, due mainly to not being perceived on the site or not behaving as they expected in the case of the "Notification service".

"Upload URL" and "Similar 3D models" had the best ratings although users mentioned that they were not familiar with some of its terms (URL, 3D model, etc.)

The next figure shows a comparison between how the users perceived the difficulty of a task once formulated (expected rate) and how they assessed the difficulty of that task after its execution (task rate).
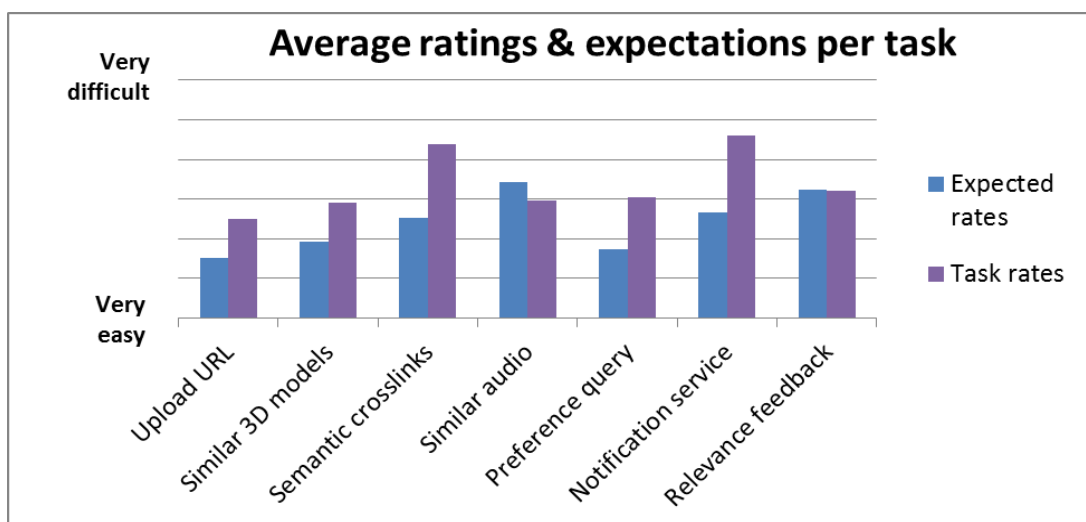
Figure 29 – Average ratings & expectations per task (experts)

All tasks except "Similar Audio" and "Relevance feedback" were rated as more difficult than expected.

The biggest difference between expectations and actual ratings were for "Semantic crosslinks" and "Notification service", as users considered themselves to be familiar with

those functions but they were not behaving as expected.
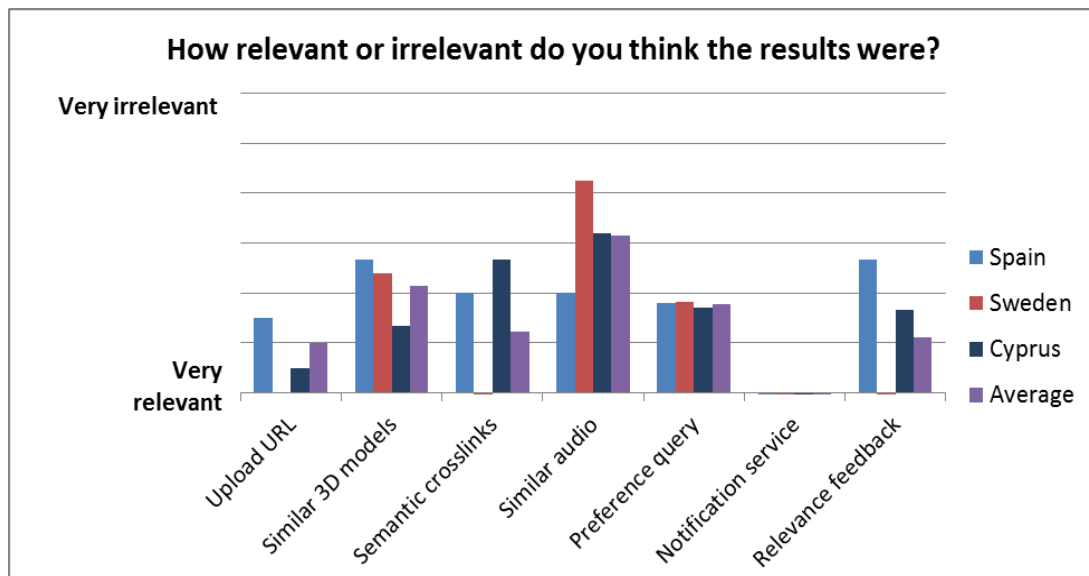
Figure 30 – Relevancy of task results (experts)

Users rated the results provided by "Similar audio" as being the least relevant, mainly because they were unsure what criteria for used to define the similarity.

Note that "Notification service" was not rated by the users, as the results provided were dummy data and not actual results.
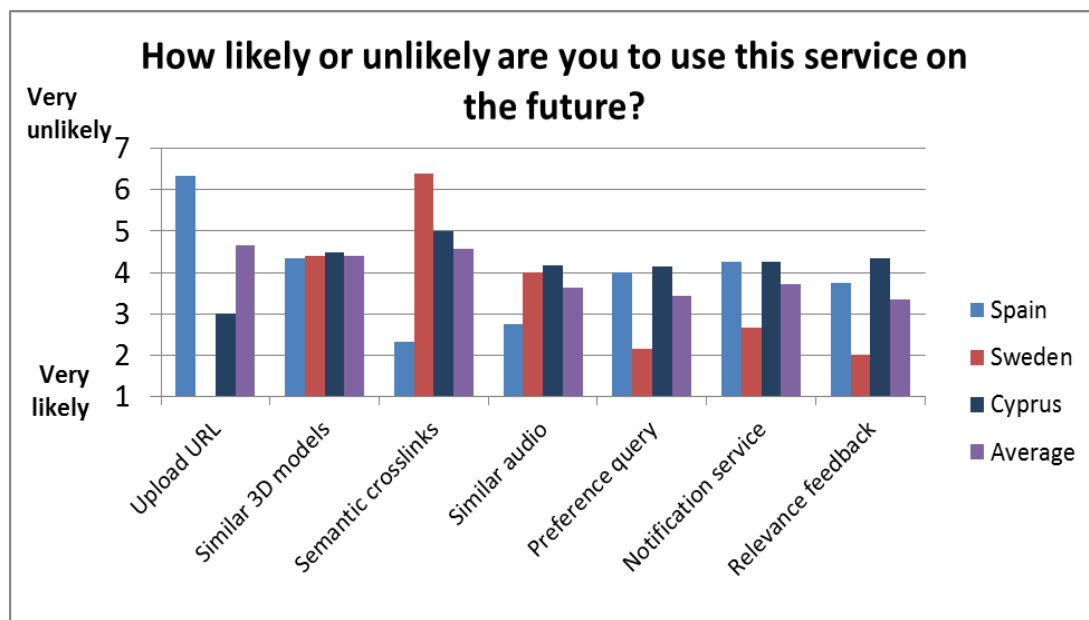
Figure 31 – Likeliness to use the service on the future (experts)

Users said that they would be unlikely to use "Upload URL", "Similar 3D models" and "Semantic crosslinks" in the future. This was mainly due to the small 3D dataset used in the tests.

In general, the services tested by the Experts were not rated highly with regard to being likely to be used in the future.

5.3 Usability Issues and Recommendations
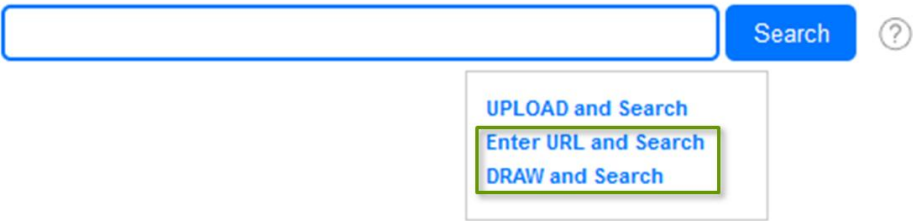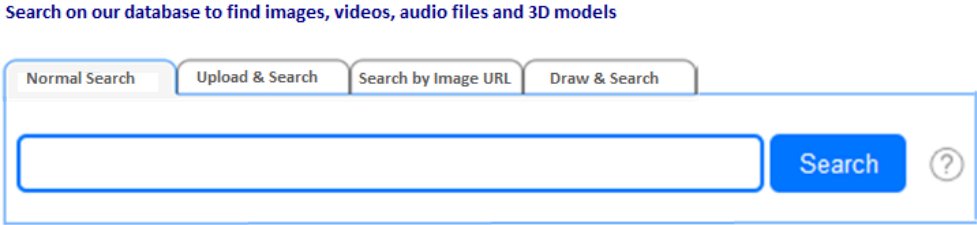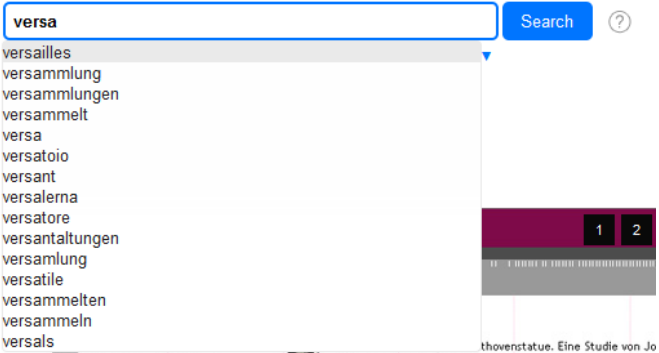
For each usability issue identified, advice is provided on how to best minimise or eliminate the issue. Also provided is an indication of the severity of impact for site visitors and hence the urgency of the need to implement this advice. Usability comments are classified in one of four categories shown overleaf. The distinctions between these are not strict, and should be interpreted as such.

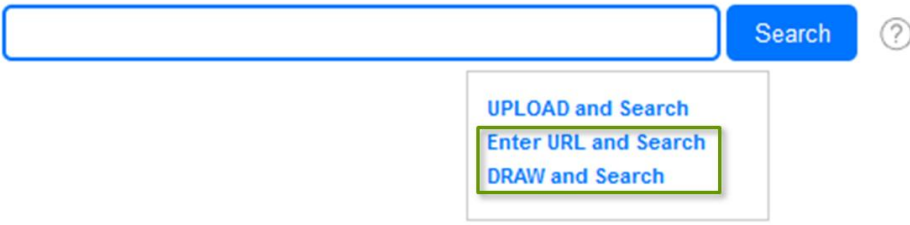| Category	Definition
P	**Positive finding -** This approach is recommendable and should be preserved.
H	**High Severity problem** - Likely to cause a problem that frequently prevents participants from completing their task or will greatly annoy them. Important to fix this soon.
M	**Medium severity problem** - Likely to cause some temporary confusion that delays participants from reaching their goal with the site for some considerable time (30 secs – 1 minute), and causing some degree of frustration/irritation, but unlikely to prevent them from reaching their goal.
L	**Low severity problem -** Cosmetic or minor problem that may cause participants to hesitate or be distracted for a few seconds.

The following tables outline the usability issues associated with the website.
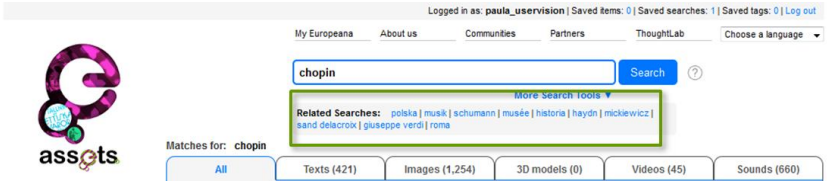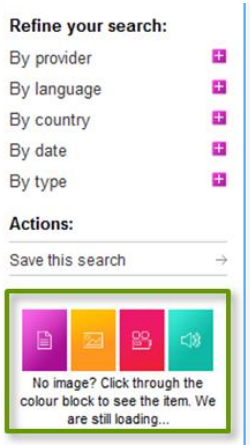
5.3.1 Search box
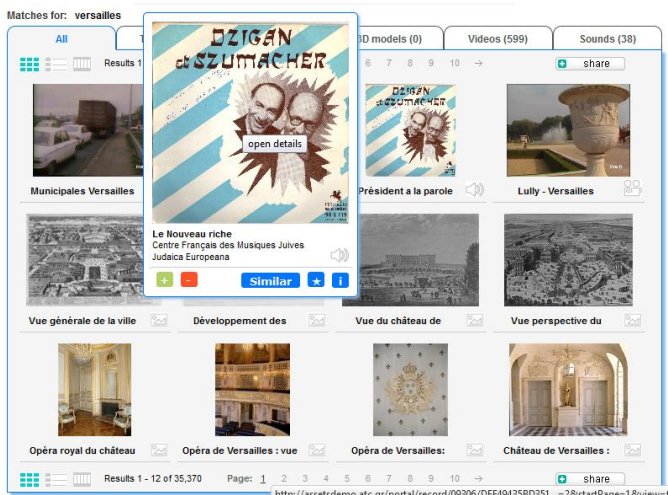
Usability Issue	Severity
Search options were not understood (M): Many users did not understand what Enter URL or Draw meant in this context. **Recommendation:** For those users who are not used to this type of search, extra explanation will be needed when these options are selected. A good example of explaining Image URL is http://www.tineye.com/imageurl. Having a video for explaining how to search by those different types could be helpful as well.	M
Other search options were very difficult to find (H): Users did not expect to find these functions on this page, so they were unsure what the "Other search options" would provide.	H

Recommendation: The new search options should be more visible and explicit, particularly on the homepage, allowing the users to learn all the different features the website can offer. Having them as tabs on the main homepage and hidden on the rest of the pages could be an option for making them more visible:

Tabs instead of Search Options: Having the different options more visible allows the user to understand what the site is offering.

Furthermore, each option can have their own "Help" to explain to users how to use this functionality better.

Missing spell corrector (H): When users misspelled any of the words in the query, the system did not provide any spelling corrections. **Recommendation:** It is considered best practice to include a spelling corrector for accessibility and usability reasons.	**H**
Suggestions dropdown rows are not fully selectable (M): Users often tried to click on the whole row to selecting one of the suggestions and were left confused when it was not selectable. Only clicking on the word itself allowed the row to be selected. 	**M**

| Recommendation: Change the clickable area to the whole highlighted row, not only the word to avoid any confusion. | |

5.3.2 Results Page
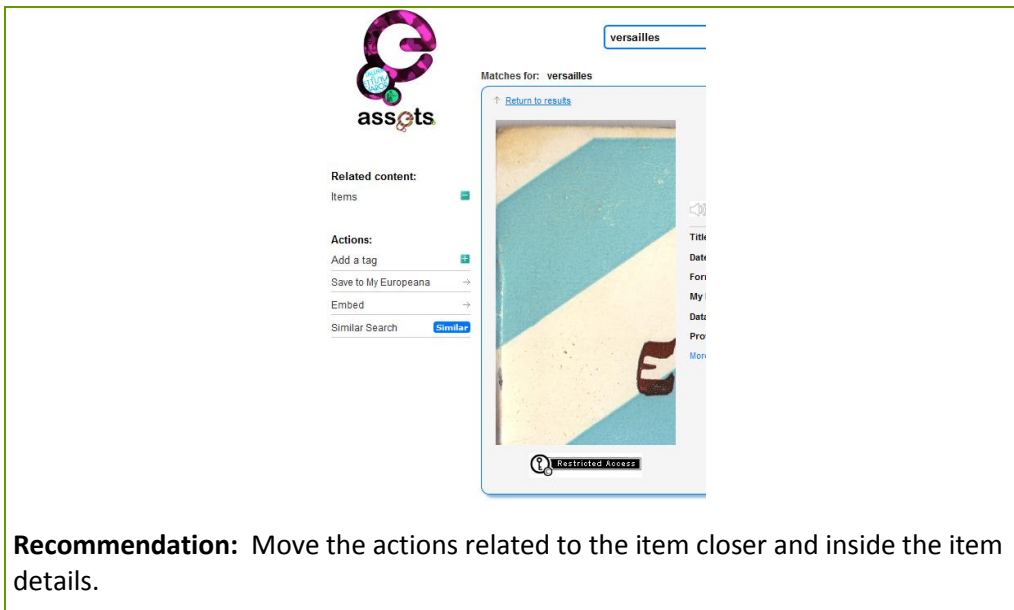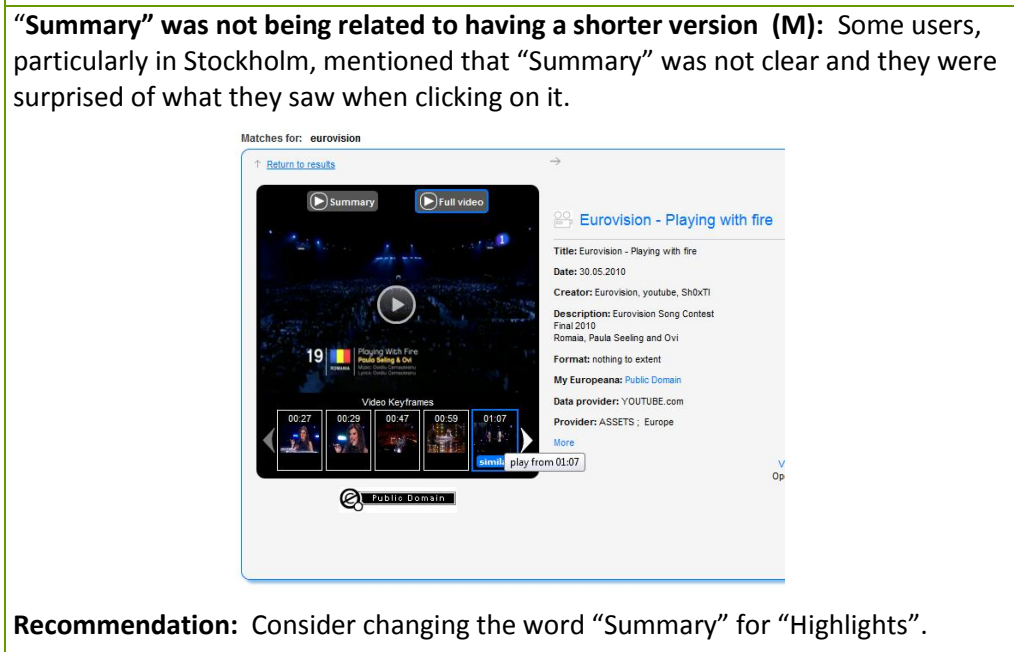
Usability Issue	Severity
	M
Related searches are easily missed (M): Users did not look at Related Searches for more information or for other options. Most of the time they didn't see them, confused them with spelling correctors or went to the bottom of the page for something similar. **Recommendation:** Consider moving the Related Searches to the bottom of the page following Google practice. Additionally, and in order to avoid any confusion, the heading for the suggestions should indicate how they have been selected by using the words: "Other users searched for" or "Other users who searched for <query> also searched for:"	
Legend is not clear (L): Although many users noticed the legend on the right of the page, they did not understand what it meant for them. This could be caused because it was calling their attention before noticing the images on the results or because it was not related to the results page. **Recommendation:** If the icons are intuitive enough, there should be no need for a legend, consider using a single colour with different icons and "No thumbnail" or "Thumbnail not available" message at each item.	**L**
"My Europeana" section is expected to be on top of the page (L): Users did not find "My Europeana" link easily or confused it with other information. They mentioned that it was on the wrong location, as the rest of the personal information was on the top of the page line. **Recommendation:** Move My Europeana to the top of the page, next to the Login	**L**
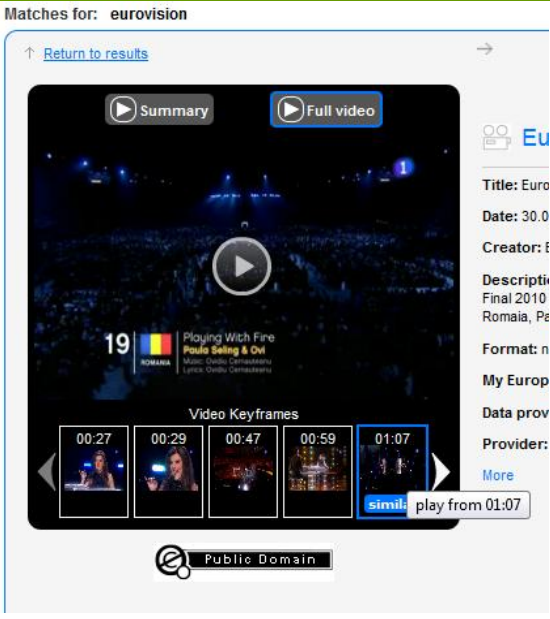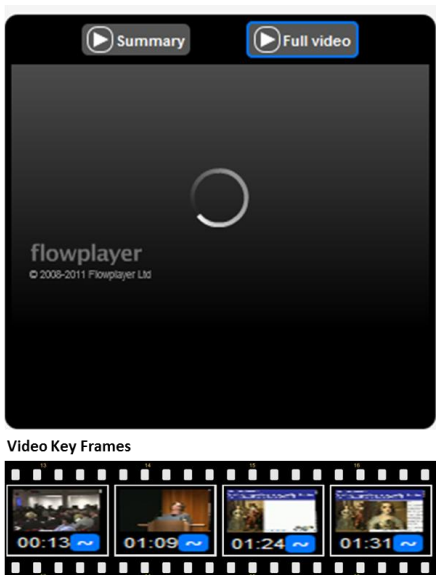
details.

	M
Expanded image makes it difficult to click some sections (M): The lack of delay for popping up the expanded image was making it difficult to click on certain buttons, like the information icon or to move around the different results. **Recommendation:** Include a delay for showing and fading the close-ups. The standard for this delay is usually around 500-1000 milliseconds.	
Missing home link (L): Some users tried to find a home link and did not recognise the logo as a link. **Recommendation:** Include a home link as part of the top main navigation to avoid confusion.	L
Tabs need to maintain selection and be automatically selected (L): Users found irritating that the tabs were not maintained when searching again or that they were not automatically selected when included as part of the query. Furthermore, some users missed the tabs even when including 3d models or videos as part of the query. **Recommendation:** Ensure that the selected tab is maintained when searching again and that it is pre-selected when the type of document is part of the query.	L

5.3.3   Result details

Usability Issue	Severity
Icons for file types are confused with buttons (M): Particularly for audio and video files, users tend to click them for playing the content. Users were confused due to not having any action when they were clicking on the icon.	M

Recommendation: Include the icon as the "Open details" link so the user get some action when clicking.

Furthermore, change the icons for the audio files to something different that is not easily confused with the "Play" icon, for example:

Similar button is now very clear (P): Users found the new design for similar button very easily and were happily surprised by this functionality. Some users were still unsure about the criteria used for finding similar items. A further explanation of this functionality or being able to select that criteria may be needed for better understanding.	**P**
Relevance feedback buttons are confusing (H): Users were confusing the relevance feedback button with "Zoom in/out". Additionally, users did not understand the difference between "More like this" and "Similar". The whole concept of filtering by relevance was very difficult to understand by the users. **Recommendation:** Change the icons for something more meaningful for the users that represents the functionality. For example: 	**H**
Actions appear unrelated to the content details (M): Users did not see the Actions list at the left of the file details. Only when probed around were able to notice them but mentioned that they expected them to be below the details, not on the left.	**M**

Recommendation: Move the actions related to the item closer and inside the item details.

"**Summary" was not being related to having a shorter version (M):** Some users, particularly in Stockholm, mentioned that "Summary" was not clear and they were surprised of what they saw when clicking on it. **Recommendation:** Consider changing the word "Summary" for "Highlights".	**M**
"**Full video" was confused with "Full screen" (L):** Some users confused the "Full video" button with the "Full screen" button. **Recommendation:** Consider changing the two buttons for a selection element type (i.e. a radio button), so users understand that they are opposite concepts and that they are related.	**L**
Users confused the whole frame with the Similar button (H): When asked about finding videos that contained similar images, the users tried to click on the frame itself instead of the icon at the bottom. This was probably caused due to having the same colour when the mouse was hovering.	**H**

Recommendation: Change the highlight colour or the button colour to something different so it does not confuse users.

"Similar key frame videos" concept was difficult to understand (H): Users did not understand the concept of searching for similar frames. The whole idea was alien and regarded as pointless. Having a functionality that is not user-friendly, adds noise to the site and prevents users to use other useful services. **Recommendation:** Consider removing this service and allow only for searching by similar videos.	**H**
Frames were not completely clear (M): Users were unsure of what the keyframes were; on the first glance they thought they were related videos or shorter versions of the same video. **Recommendation:** Consider redesigning the frames as proposed by the expert evaluation. 	**M**

The keyframes require clear description as this is not a common concept. To aid in this we suggest three changes:

- Move the keyframes from out of the video panel (to avoid associations with related videos)
- Give the keyframes section a clear title
- Consider using a visual metaphor to describe the frames in the video, e.g. a film strip.

"Paid access" created trust issues (L): When users were trying to play certain videos or audio files, they noticed the Paid Access icon and thought that they had to pay for accessing the content. This created trust issues with the page as users thought that it was a marketing strategy instead of free content. **Recommendation:** Consider moving the "Paid Access" icon a less visible location or remove it completely for ensuring the users' trust. Also allow the users to get more information about what the "Paid Access" is by clicking on the icon.	L
Audio categories were misunderstood (M): Users did not understand what audio categories meant and what was it for. Most of them could not understand the descriptors like "relaxed, happy, blue, tonal" and were confused by them. **Recommendation:** Use the Audio descriptors for words that are meaningful for the users like "country", "1960s", "blues", "upbeat", etc.	M

5.3.4 Similar/Upload search results

Usability Issue	Severity
Missing similar image or uploaded image (L): Users were not sure if the images shown on the similarity by relevance feedback or by uploading were taken into account. Furthermore, they missed the original images for comparison purposes.	L

Recommendation: Ensure that the image that is used to generate similar results is always shown in the "matches for:" area for comparison purposes.

5.3.5 Preference query

Usability Issue	Severity
Dates were not sorted chronologically (M): When users tried to filter by specific dates, they found out that the dates were sorted by the number of results, instead of chronologically. This confused users who are used to search chronologically when referring to dates. 	M

Recommendation: Use chronological order for dates or similar types of filters.	
The refine button was not visible (H): Most users expected the filter to be applied just by clicking the filter options, instead of having to press "Refine". Additionally when some filter categories are open the button falls under the fold and is not visible. **Recommendation:** Remove the "Refine" button step and apply filters directly when selecting filter options. If this is not possible, move the button to the right of the menu, change the wording to "Refine my results" and duplicate to the top as well so it becomes more visible.	H
Use of language short code instead of the actual language name (H): When filtering by language, the language short code is shown, most users were not familiar with them or were not sure which language they were referring to (i.e. mul, pl, etc.) **Recommendation:** Use the full name of the language when filtering by language in the filters and on the breadcrumbs.	H
"Show block by block" was confusing (H): Users did not understand what "Block-by-block" meant and assumed it was to show each filter separately. **Recommendation:** Change the words "block-by-block" to "Show combination of filters separately". Furthermore, only enable this option when there is more than one combination selected, otherwise this should stay disabled.	H
"Next" & "Back" links are not meaningful or visible (H): The buttons for moving between "blocks" are not visible at all for users. When probed about them , users are not sure about what they are or what they will get.	H
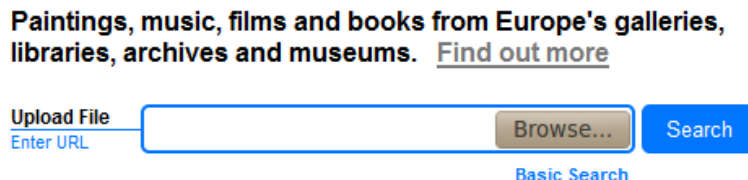
Recommendation: Change the location of the buttons to the next line and change the wording of them to something more meaningful like "Next combination" or use the actual filter combination selected (i.e. "Go to Language: French; Year 1901"

Not evident how to sort the filters (L): Users were not sure how to sort the different filters selected. **Recommendation:** Consider changing the mouse pointer to an icon that indicates movement. For example: 	**L**

5.3.6   Notifications

Usability Issue	Severity
"Subscriptions" was not related to "Notifications" (M): Users did not easily find where to create new alerts or notifications for new content. When looking at "My Europeana" tabs, they thought that Subscriptions was something different. **Recommendation:** Change the word "Subscriptions" to "My notifications" or "My alerts" and join "Saved Notifications" under the same tab.	**M**
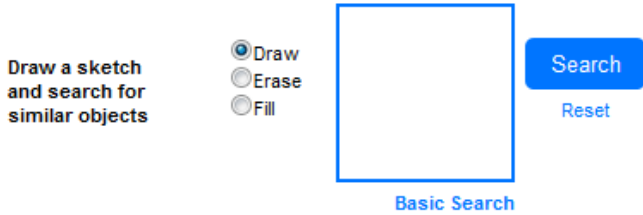
Furthermore, some users expected to be able to create new notifications at the search results page, having an option under "Actions" that allows to create a new notification for that particular query.	
Lack of instructions (H): Users did not understand that they have to choose a term instead of free writing their own search query. Many were left feeling lost when they were typing something that was not on the suggestions list. **Recommendation:** Include some instructions for creating new notifications. Consider as well, the option of adding new suggestions or terms to the list if the user cannot find any satisfactory ones.	H
Buttons were not visible enough (L): Due to the low colour contrast and not being placed closer to the notification text, the buttons were easily missed by the users. **Recommendation:** Consider changing the style and colours to make them more visible. Ensure that they are easily related to the notification text by placing them closer.	L

5.3.7 Upload & Search

Usability Issue	Severity
Missing "Upload" button (M): Some users expected to find an Upload button instead of Search. They expected a two-step action. **Recommendation:** Include some feedback after the file is selected that indicated that it has been uploaded (i.e. a sentence acknowledging the name of the file or a preview of the file).	M
Missing terms & conditions for the upload (L): Some users mentioned that they were not sure about what was going to happen with their uploads and it created trust issues. **Recommendation:** Ensure that further explanation is provided on the page for those users. Include information about how the image will be used or if it will be stored on any of the Europeana servers.	L
Confusing option (M): Users were not sure which option was selected: Upload File or Enter URL. Using the colour and font weight change was not enough for clarifying which action was selected.	M

Recommendation: Remove the other option when one is selected and keep it under "Other search options" as previously.

Consider the recommendation made for the homepage search on 5.2.2.

5.3.8 Draw & Search

Usability Issue	Severity
Using the mouse is not a natural way for sketching (M): Users mentioned that the mouse was not accurate enough for sketching and it felt unnatural. **Recommendation:** Consider searching by other options like choosing shapes or monument types instead of drawing them.	**M**
"Fill" was not understood by the users (M): Users did not use the "Fill" option most of the times and therefore the system misunderstood what the user was trying to draw. This meant that models were returned that were not at all relevant to what had been sketched. E.g. a filled triangle returned a pyramid model, but an unfilled triangle did not. **Recommendation:** Provide results that relate to empty shapes as well as filled shapes.	**M**
Very small canvas (L): Users complained that the canvas was too small for sketching and that it was difficult to define the details, making the drawing very inaccurate. **Recommendation:** Make the drawing area much bigger and resizable.	**L**

5.4   Questionnaires

5.4.1   Pre-test questionnaire results

Before starting the tasks, participants were asked a series of questions to understand their expectations and their online search experience and skill.

The word cloud provides a visualization of the responses of the participants. The larger words are the most frequently mentioned by the participants.

Q1: What do you expect from the Europeana site?
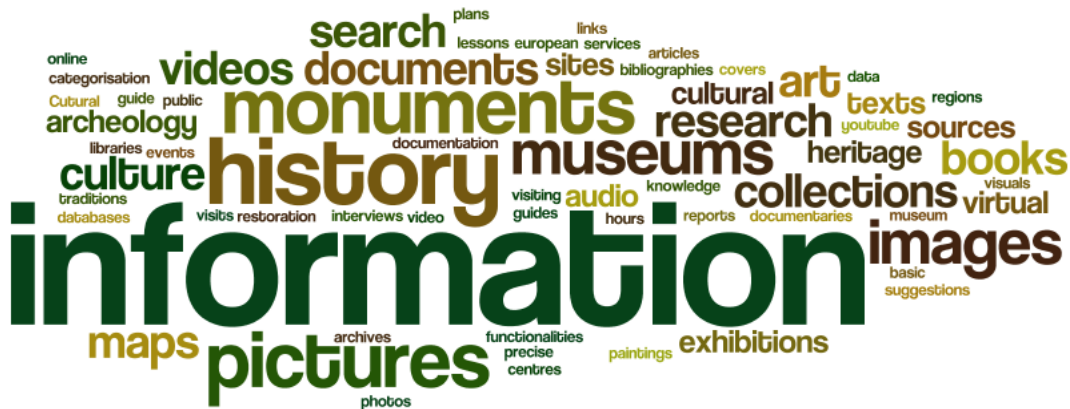
Figure 32 – What do you expect from the Europeana site?

Only 4 participants had heard of the Europeana site before, their experience was not described as positive.

When asked about what the participants expected from Europeana, most of them mentioned the word "information" about Europe and its cultures.

A small number of participants mentioned as well the possibility of having interactive guides to museums.

Q2: Searching information online

o   **How often do you search for information on-line?**
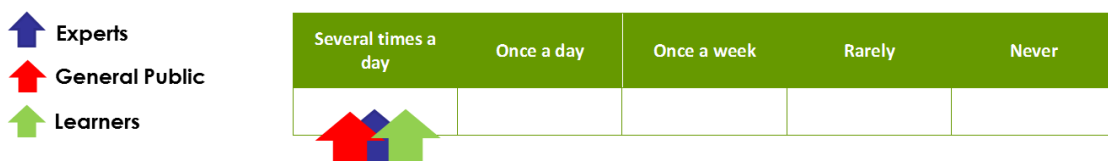
Figure 33 – How often do you search for information on-line?

o   **Which other sites do you visit regularly related for finding information about**

historic monuments,   paintings or cultural heritage?

Figure 34 – What other sites do you visit regularly?

o   **How often do you use advanced search features offered by search engines?**

	Usually	Often	Once a week	Rarely	Never
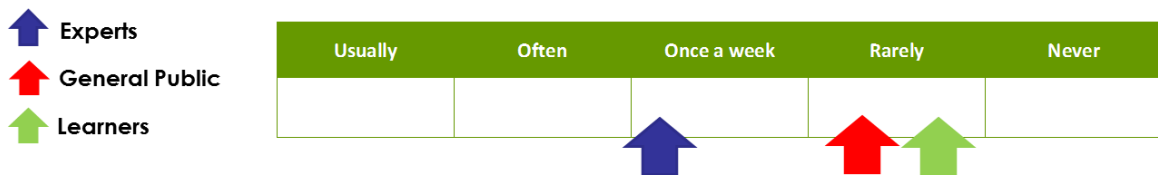
Figure 35 – How often do you use advanced search features?

o   **What type of content do you search more often?**
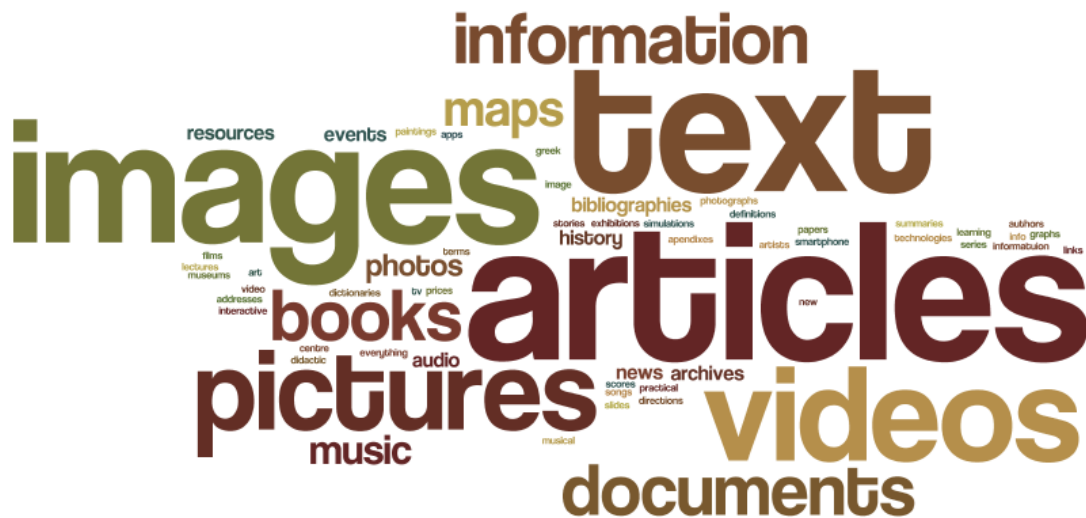
Figure 36 – What type of content do you search more often?

o **Do you only look for information, or do you also add information or content to sites?**
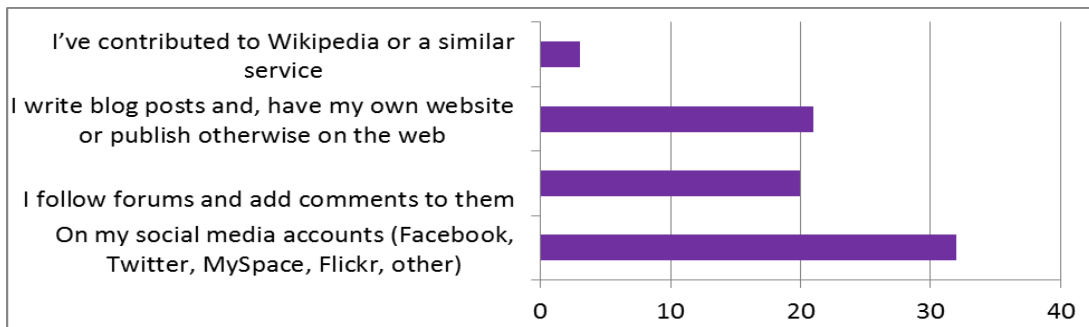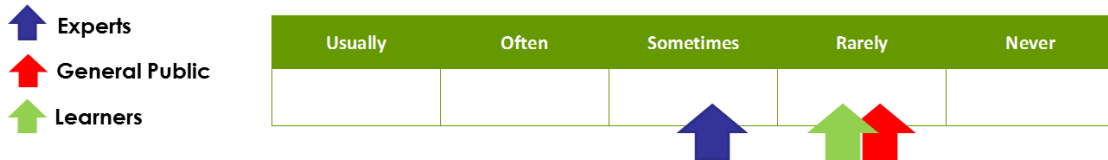
Figure 37 –Do you only look for information or also add information or content to sites?

Regarding the bar diagram above, the results happened to be very similar for all user types, with most of the users using social media and commenting on forums, etc.

5.4.2 Post-test questionnaire results

Once the tasks were finished, participants were asked a series of questions to understand what they thought of their experience.

Each graph shows the average of the ratings for each of the questions.

The word cloud provides a visualization of the responses of the participants. The larger words are the most frequently mentioned by the participants.

Q1: Looking at both sets of results, which version of the results is more relevant for you? (ASSETS metadata base ranking service vs Europeana ranking service)
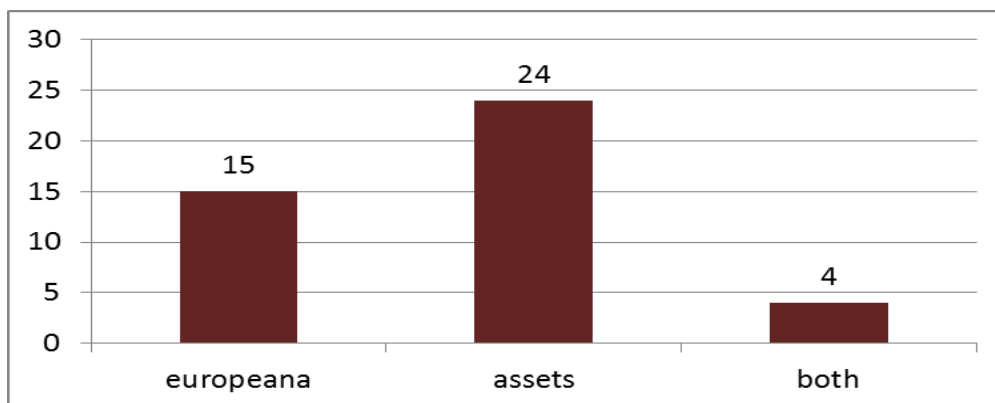
Figure 38 – Preference of ASSETS metadata based ranking vs Europeana ranking

During the sessions on the different countries, we used different queries (berlin wall, Italy, etc.) at the Metadata base ranking ThoughtLab for having a general view of users opinions. For those users who answered the question regarding which version they preferred, 24 of them chose the new version but 15 of them went for the older version and 4 of them mentioned that both of them were as good in their opinion.

The reasoning behind varied greatly, some users mentioned that the language of the results and the quantity of information provided on the first results was a good indicative. Meanwhile other preferred to have pictures of the content.

Q2: Was there any particular section you liked?

Figure 39 – Was there any particular section you liked?

When users were asked about the functions or sections that they liked, they mainly mentioned the possibility for uploading pictures and finding similar images.

In general, the content was considered good and the tabs a good structure for the information.

Users used words like "useful", "easy" and "interesting" for describing the whole site.

Q3: Was any particular section that you did not like or you would like to change?
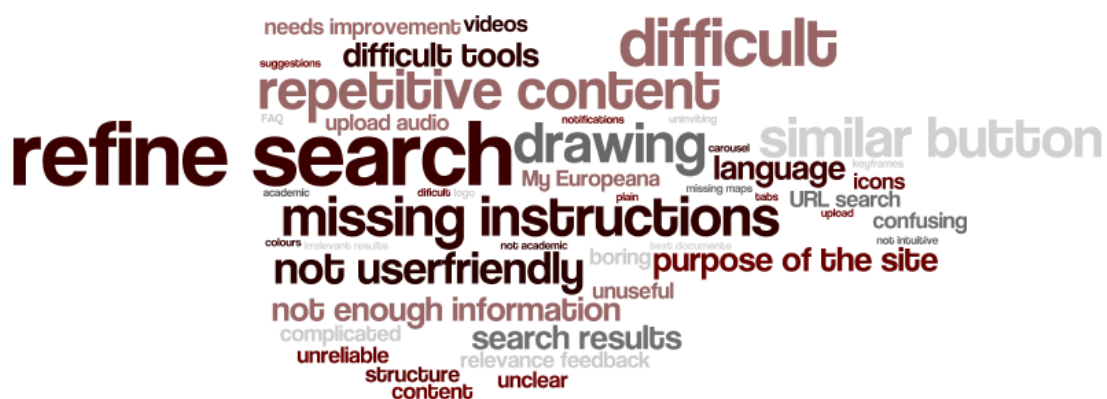
Figure 40 – Was any particular section that you did not like or you would like to change?

Users mentioned the refine search section (both the current version and the preference query version) as one of the negative aspects of the site. Mainly because it was not behaving as they were expecting or they were not sure how to interact with it.

They mentioned as well that the site needed extra instructions, a clearly stated purpose and

was missing information regarding the items.

Some users said that most of the content was already provided by other sites like Google, Youtube, Flickr and that they would be unlikely to change to the Europeana site as it did not offer anything extra.

Q4: Did your experience meet the expectations at the start?

Figure 41 – Did your experience meet the expectations at the start?

According to the users the site was neutral/slightly above their expectations, although it still needed significant improvement.

Q5: How likely would you be to use this site in the future?

Figure 42 – How likely would you be to use this site in the future?

Users mentioned they would be slightly likely/likely to use the site again because the content provided was interesting and they were curious about some of the services included. Although some of them said that some of the services were not something they could see themselves using or needing in the future.

5.4.3   Emotional Response questionnaire results

Participants were asked to fill in a questionnaire to measure their emotional response to using the website.

They were presented with a series of pairs of words and asked to circle the number for each part which they thought best represented the ASSETS website.
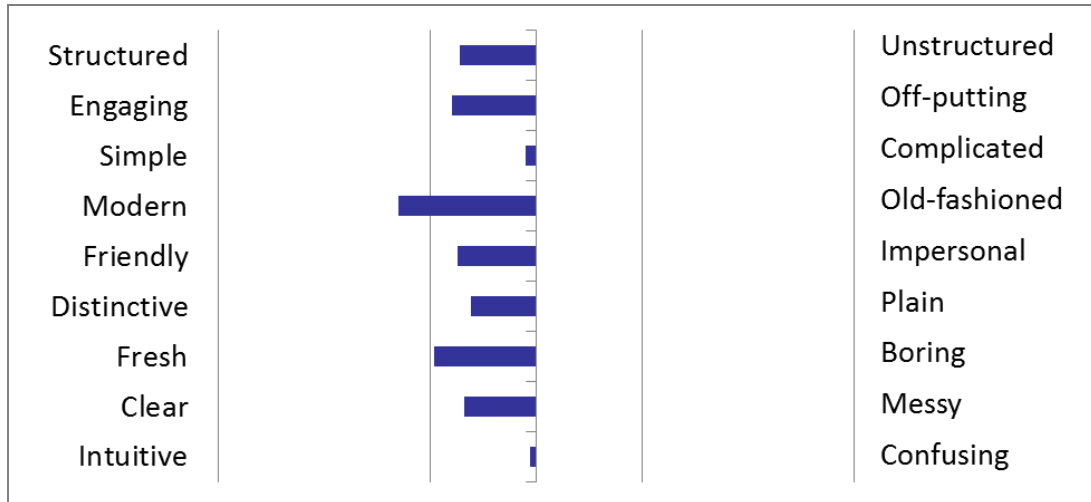
Figure 43 – Emotional Response questionnaire results (I)

Although the responses to the Emotional questionnarire were in general of moderate intensity, we can stand out that all of them were positive, as can be seen in the previous graph.

Users considered the site as 'Modern' and 'Fresh' (as opposite of 'Boring') for the new functionalities offered. The users also defined the site as not very 'Simple' or 'Intuitive', which could be closely related to the unfamiliarity with the novelty of some features.
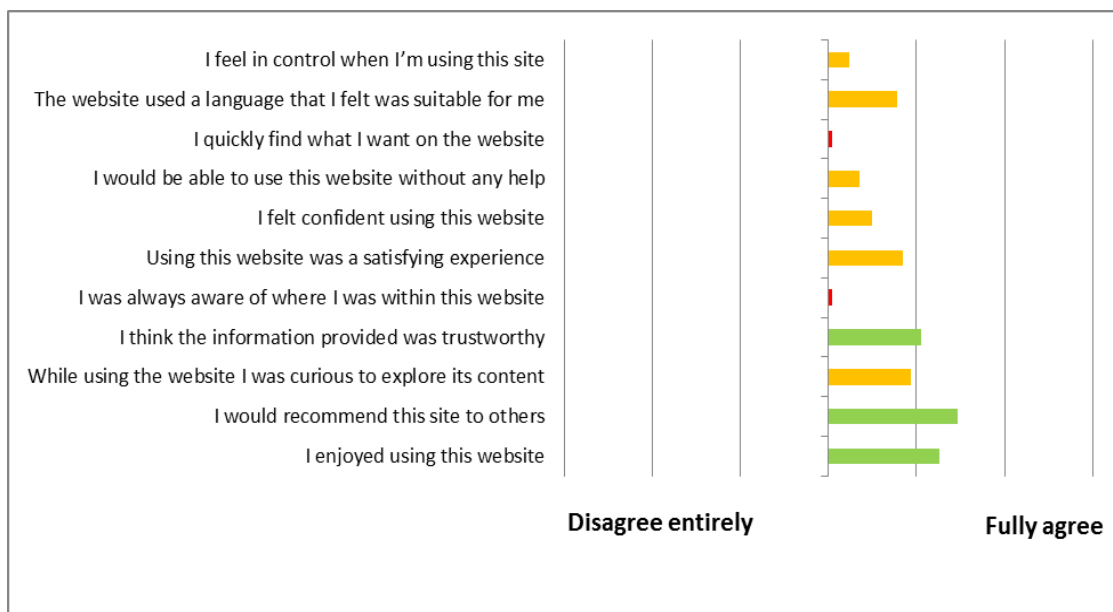
Figure 44 – Emotional Response questionnaire results (II)

Figure 45 – Emotional Response questionnaire results (III)

In the two figures above the colour is used as an approximate indicator of the significance of the result. Scores at, or below, the neutral point are painted red (the highest priority to take action towards), followed by amber for medium priority and green being the lowest priority.

6. Conclusions and next steps

6.1.1 Main findings of the Heuristic evaluation

The Heuristic Evaluation has identified several positive findings as well as some usability issues. The most important of these are listed below:

- The icons used in the zoom tool might be more self-explanatory. Also, for consistency reasons, they should be visible (disabled) even when their functionality is not available.
- Internal system queries are sometimes shown in the Search box. This is likely to confuse users and reduce trust on the site.
- In the Details page, the "Related content" section should be grouped together with the main content, to make its association more obvious.
- Guidelines for some functionality might be required, particularly when aiming at a generic public user group.

6.1.2 Main findings of the Usability testing

In general, users were able to use the site for simple search queries and were able to interact with most of the functionalities. Other search options like 'search by url' or 'uploading a file' were not easy to find but users considered them useful and innovative.

Most expert users struggled to use the most advanced functions like 'notification Service', 'relevance feedback', 'search by URL' or 'preference query'.

Some functionality, like 'draw and search' or 'similar keyframes' was not regarded as particularly useful.

The changes made between the tests in Barcelona and Stockholm improved greatly the usability of the 'Similar' functionality (see Figure 22 - Average success per task - General public and learners).

The next paragraphs give a more detailed view of the usability issues grouped by end-user service:

Metadata based ranking

- 24 of the users chose the new version but 15 of them went for the older ranking version and 4 of them mentioned that both rankings were equally good in their opinion.

Image based similarity search

- Users liked the similarity search functionality and found it easily with the new design that was implemented after the Barcelona testing (replacing '~' button by "similar" button). There were some questions about the criteria used to implement the similarity search (similar colour, similar shape, etc.).
- Regarding the image upload, some users expected it to be a two-step process (uploading & searching) instead of a single-step process and had questions about how their uploads will be managed by the site.

Query Suggestions:

- o Users did not easily find the suggestions, and they often confused them with spelling corrections. Instead of relying on the suggestions provided by the site, users mainly tried more complex queries in the Search box.

Audio based similarity search:

- o Regarding the audio descriptors, users did not understand what they meant and did not click on them.

Video summarisation:

- o Although most users were able to find the summary button, some of them could not relate it to a shorter version of the video.
- o Some users questioned the usefulness of this service.

Video similarity search:

- o Users did not understand the concept of searching for videos containing similar key frames.
- o The video key frames were sometimes confused with related videos.

Relevance Feedback:

- o Users commented that the "More like this"/ "Less like this" buttons were the same as the Similar button, so they expected to be able to apply the filter by clicking on Similar.
- o Some users related it to the Like/Dislike option in Facebook and thought the site will keep this criteria into account for the future so it would influence in future queries.

3D based similarity search:

- o The "Draw and search" service was not considered particularly useful mainly because the results were not always similar to the provided sketch.

Semantic cross-linking:

- o Users did not easily find the semantic cross-linking and tended to change the search query for a more specific one.
- o Those users familiar with the functionality named the semantic cross-links as "tags".

Taxonomy based notification service:

- o Users did not find the notifications under the "Subscriptions" tab and struggle to create new notifications.
- o Some users expected to have the possibility of free writing subscription items instead of selecting from the dropdown.
- o Users expected to have an option to subscribe to a specific query from the results page directly.

Personalisation service:

- o Although users liked the idea of being able to filter the results, they expected the filters to be applied automatically instead of clicking the "Refine" button.

o Most users did not understand the "block-by-block" option.

6.1.3 Next steps

The next steps in the Evaluation task will be reported in the D3.1.3 deliverable:

o **Analyze the usability issues** identified in the Heuristic evaluation and User tests, classify them by responsible and define a priority for their implementation based on their severity and their relevance, taking into account the new and future Europeana web site versions.

o **Address the issues in priority order** starting with the High Priority issues, then as many of the Medium and Low Priority issues as time and resources allow.

o **Report the final changes and decisions** in the D3.1.3 *Final Report of User Evaluations* deliverable.

7. Appendix 1: Tender for testing ASSETS services

Request for Proposal – User Testing ASSETS services

7.1 Background

Europeana

Europeana.eu was launched in 2008, with the goal of making Europe's cultural and scientific heritage accessible to the public. Europeana is funded by the European Commission. Its office is based in the National Library of the Netherlands, the Koninklijke Bibliotheek.

Overseeing the project is the Europeana Foundation. Its members are the presidents and chairs of European associations for cultural heritage and information associations.

Work to enhance Europeana with more functions and features for users is being done by the project Europeana v1.0. Support comes from related Commission-funded projects and over 180 heritage and knowledge organisations and IT experts across Europe. They are helping to solve technical and usability issues.

The Europeana v1.0 project is funded under the Commission's eContentplus programme. It's a Thematic Network that brings over 100 partners from libraries, museums, archives and audio-visual collections together. We're building open source solutions and reusing technologies developed in related Commission-funded projects. Universities, ministries and heritage strategy bodies are also part of our knowledge-sharing consortium.

ASSETS

ASSETS (Advanced Search Services and Enhanced Technological Solutions for the European Digital Library, site: www.assets4europeana.eu) is one of these related projects.

ASSETS is a two-year Best Practice Network co-funded by the CIP PSP Programme, It aims to improve the accessibility and usability of Europeana by developing, implementing and deploying software services focused on search, browsing and interfaces.

These services include searching for multimedia objects using metadata, content similarity and ranking algorithms for improved results; rapid navigation of multimedia objects through semantic cross-links and better interfaces designed for interacting with multimedia objects.

ASSETS run until March 2012. The ASSETS consortium includes 24 partners from 10 European countries and one partner from Japan. It is led by Engineering Ingegneria Informatica in Italy.

7.2 Objective

The objective is to test the usability of (prototypes of) services developed under the ASSETS project that have an impact on the user interface and end user experience of Europeana.eu, specifically:

- The user interface

- The user behaviour when interacting with the services/functions

This will be performed in Work Package 3.1 Service user evaluation and assessment, in task 3.1.2 User evaluation activities [M3-M22]:

"The methodology for user-centered design and user evaluation is based on both laboratory involvement and experimentation, and field results i.e. post integration of the functionalities in Europeana. They will therefore include lab based interactive sessions with users from different age groups and backgrounds. These labs will allow the user to drive the user interface design; once the user interface is designed, the sessions will allow us to follow the user as they use the services, recording their click movements on screen. These will be held in at least 3 different Europeana countries."

The end result will be a public report in English (deliverable 3.1.2 Report on user evaluation of functionalities) and a presentation of this report during an ASSETS project meeting. A preliminary internal report is expected as an intermediate step.

User comments and difficulties observed during the testing of the prototypes will feed into functional design changes.

7.3 Requirements

- The user testing methods can be (for example) user observation, focus groups or eyetracking. The appointed company is free to choose the methods they consider most appropriate and the proposal should elaborate on this.

- The base of test users should at least correspond to the following target groups:

  - learners (secondary education, students and teachers)

  - researchers

  - citizens with personal interest in cultural heritage, such as tourists

  As extra information, summaries of Europeana Personas and their search skills and behaviors are relevant (Appendix I).

- The testing has to be executed in three different EU countries, excluding the United Kingdom, the Netherlands and Germany. Other countries are optional, but the countries of the ASSETS project consortium members have the preference:

  - Italy

  - Greece

  - Austria

  - Spain

  - Sweden

  - France

  - Luxembourg

  - Cyprus

  Test users can partially be drawn from an existing registry of Europeana test users and members of the ASSETS user community.

  Users should be tested within their own country. It is therefore advised to form a consortium with organisations in two other countries, similar to the work done for the Focus group and media lab research for Europeana.

- The services to be tested are (elaborate descriptions available in Appendix II).

- Metadata based ranking (T2.2.1)

- Image based similarity search (T2.2.4)

- Search recommendations/Post-query processing (T2.2.1)

- Audio based similarity search (T2.2.6)

- Video summarisations (T2.2.7)

- Video similarity search (T2.2.7)

- 3D based similarity search (T2.2.5)

- Semantic cross-linking (T2.5.1)

- Content creation by re-use service (T3.2.2)

- Taxonomy based notification service (T3.2.3)

- Personalisation service (T3.2.4)

- To be included in the proposal:

  - Proposed approach and methods

  - Overview of timings and proposed deliverables

  - Background and expertise of your company

  - Short biographies of participating team members

  - Budget

7.4   Planning

Deadline for submission proposal:	27th of May
Selection of company:	Week 22 / 23
Physical meeting to discuss set up and process:	Week 24
Workplan to be agreed by:	Week 25 or 26
User testing sessions to be held:	September & October
Preliminary report ready:	15th of November
Final report ready:	1st of December
Final presentation: TBD	

7.5   Indicative Budget

The indicative budget is €75,000 Euro in total, which should include:

- VAT

- Incentives for the participants of tests

- Travel costs to be made by the contractor;

- Translation costs

- Use of lab facilities, venues, etc.

- Producing the reports in English

7.6 Criteria for assessment of contractors

The criteria on which the proposals will be judged are:

- Demonstrable experience in user-oriented evaluation of interactive IR systems, both laboratory-based evaluations and evaluations with users performing work tasks in "real life" situations

- Knowledge of the field of cultural heritage online (libraries, museums, archives & audiovisual collections)

- Demonstrable understanding of the issues relating to web access of cultural heritage and to online search

- Demonstrable understanding to issues relating to multilingualism

- Best value for money – price to work undertaken

- Track record in the field(s), and past work in these areas

- Ability to write a succinct report and give presentations on the research results

7.7 Submission date for proposal

Friday 27th of May, 2011 COB

7.8 Contacts

The submission with reference ASSETS user testing Proposal should be sent to:

Jobs.edlf@kb.nl with cc adeline.vandenberg@kb.nl

If you have any questions please contact Adeline.vandenberg@kb.nl or annemarie.vangerwen@kb.nl or +31 (0)70 314 0391

Documents for information

- Results Interfaces workshop (Cyprus, 21th Sep 2010) / D3.2.1 community service use cases (to be made available after appointment)

- 'Plan for the design and evaluation of User Interfaces' June 2010 (to be made available after appointment)

- Focus group and media lab research for Europeana.eu

7.9 Appendix I – Summary of Europeana Personas

Description of 4 main Europeana Personas, taken from the Personas Catalogue, developed within the Europeana Connect Project by The Royal Library, National Library of Denmark and Copenhagen University Library. This is the version as it was available at the time ; meanwhile the catalogue has been updated by adding a cultural professional. The Personas catalogue can be found through the EuropeanaPro website.

MARIA

Personal information:

Maria is 25 years old, lives in the suburbs of a large city and works as a teacher at a public school. Her major subject is the local language and history. She teaches secondary school. Maria is very enthusiastic about her teaching and doesn't mind spending a lot of time preparing lessons; she aims to get her pupils responding and engaged in class.

Interests:

Maria is very social and often visits the theatre, a museum or an exhibition. She frequently travels in Europe with friends who share her interest in European culture and history.

Media use:

Maria spends a lot of time on the internet preparing lessons, buying books, or networking on Facebook, Myspace and other social network places. She has set up several Facebook groups together with her pupils and her friends, but keeps the two networks separate.

Maria doesn't go anywhere without her mobile phone, and she only turns it off when she is teaching.

She mainly uses the mobile for talking, texting and for checking her e-mail.

She uses her iPod to listen to music and shares her playlists with her friends. They often recommend music to each other.

She will watch the news and historical programmes on TV, and when she can she will find a local angle to bring to class.

Search strategy:

Maria feels very comfortable searching the internet and using all sorts of social networks as she grew up with computers. She is aware that different search strategies are suitable for different situations and needs. Personally, she prefers to do a "quick and dirty" search in Google. She will spend a long time on a site if it has useful or entertaining content. She finds it interesting to see what other people have done ("People who have bought this, have also bought …") and often checks out recommendations and context relevant links.

WILLIAM

Personal information:

William is 53 years old and lives in a small village with his family. He is not really proficient in other languages, but has learnt enough terminology to manage within the context of his interests.

Interest:

William is interested in local history. The last few years his interest for genealogy has increased and now he spends a lot of his free time travelling around checking church books and archives.

Media use:

Every morning William reads the national newspaper while he eats his breakfast, and listens to the morning radio while he drives to work. At work William only uses his computer for work related tasks, emails, spreadsheets and word processing. At home he uses the internet almost every day after dinner to check out his usual websites – most of them related to local history and genealogy. From one of his friends, another genealogy enthusiast, he heard about

different internet groups and forums, and now he often enters into online discussions with other enthusiasts. He has his own website where he writes guidelines to genealogy, and links to relevant sites and databases.

Search strategy:

William is not easy to impress with new fancy web stuff and websites that claim to offer new "indispensable" facilities. He prefers to visit websites he trusts, and only visits new sites if they are recommended in a discussion forum or a trusted network. Williams's field of interest is focused on local history and genealogy, and he is not really interested in other content on the internet. His children advice him to use Google, and he does try now and then, but he much prefers to know what sort of results he will get, and that they will be relevant for his purpose.

JULIA

Personal information:

Julia is 22 years old and a university student majoring in humanities and the arts. She is quite proficient in foreign languages.

Media use:

Julia spends most of the day on campus and in the library where she meets her friends and class mates. At the library they can read and work together, and use the computers or their own laptops to prepare for their lessons or do assignments. They use the student platform for e-learning. They often arrange to meet by sending e-mails or texting each other. Julia uses Facebook to communicate with her friends, also those from her home village as it is a good way to keep in touch. She reads newspapers from a social point of view and to check out reviews of concerts.

Search strategy:

Julia always starts her search on Google, unless her teacher has asked her to use a specific website or database. When she finds a page she finds relevant, she skims from page to page to find any

relevant information. She will often follow links she thinks look interesting and frequently goes on detours during her searches. Now that she is writing her first major assignment, she knows she needs to be more focused.

JUKKA

Personal:

Jukka is 43 and works at a university as an ethnomusicologist. Frequent traveller. He is very good at all sorts of electronic devices and heavily into online social networking. In his current job he works with automatic music analysis of Hindustani classical music.

Interests:

Jukka's main interest is Hindustani classical music (which was the topic of his PhD thesis) and he moderates a forum on this topic. At least once a year, he travels to participate in music festivals and conferences related to his job. He always looks for concerts in the towns and countries he visits. He is part of both professional and social networks related to his interests and job, with a lot of friends and contacts.

Media use:

Jukka is technically literate, working with specialist databases and he will write a new search

application or macro if the existing possibilities don't work well enough for his purposes. He is an early adopter of mobile technologies, and is always on his iPhone, talking, texting, browsing the web and listening to mp3s downloaded from P2P-sites. Jukka downloads music from his favourite websites and whenever he gets a new lead on something interesting. He gets a lot of information from magazines specialised in internet technology or electronic media. He is heavily into social networking (500+contacts on LinkedIn) and shares the newest information with his professional social network – if possible, in real time.

Search strategy:

Jukka is confident in his ability to use all kinds of search. He likes to explore search results and to pull in any new information he can find on his way. If a website looks promising he will explore all its corners, just in case something might prove interesting or entertaining. He shares his findings with his students and although in his classes he is strict about relevance and credibility of sources, he encourages them to go out and explore the internet for themselves.

7.10   Appendix II – Descriptions of services and user scenarios

1. IMAGE BASED SIMILARITY SEARCH

Scenario 1: Similarity search by upload

A Bulgarian amateur historian is looking for images of medieval church statues in Europeana. The work is part of his interest in the influence of Byzantine art on the West. He's fluent in Bulgarian and is fairly proficient in Russian, German and English. He already has a number of photos from Wikipedia and some that he's taken himself on his hard-drive. He wants to find images in Europeana showing similar statues.

So he goes to Europeana to upload example images of his own in order to see whether Europeana has images of medieval church statues matching his interest. He's aware that Europeana only has limited metadata and query translation capabilities and believes he'll get better quality search results by comparing images rather than typing in queries.

Target user groups:

First, at a basic level, a user wants to make a casual ad-hoc image search or a group of users, such as school children doing an assignment. For example, a person wants to see similar images of artist "X", "type of landscape". The user may be using an image taken by her/his camera (e.g. from a vacation 6 months ago), downloaded from the web (e.g. workplace or home environment), or a picture taken in a distributed environment (e.g. currently visiting a museum).

Secondly, at a more advanced level, a similarity image search could be performed involving more pre-knowledge on a certain item/context, such as an expert/specialist user (e.g. amateur historian) that is very familiar with the object and the context. This user group may look for a certain detail or characteristics in similar images. This user also knows how to judge the relevance in a larger set of image result list, while in the first user group; they are satisfied with the first "best" image. The user may be using an image taken by her/his camera (e.g. from a vacation 6 months ago), downloaded from the web (e.g. workplace or home environment), or a picture taken in a distributed environment (e.g. currently visiting a museum).

Scenario 2: Similarity search by prior search result

A Europeana user searches for "vindmølle" (Danish for windmill) in Europeana and finds an

interesting drawing of a windmill. She wants to find other windmills like the one she's found and clicks on "Find similar image" in the object description.

More generalised the scenario can be described like this: A user makes a search using image or text as input. The result list of that search contains a list of images. The user inspects the result list and identifies one image in order to make a similarity search.

Target user groups:

The search pattern isn't as specialised as scenario 1, but finding similar images by similarity might still not be mainstream search behaviour. However, a similar feature does exist e.g. in Google's image search. Many younger users are using Google's image search and so to those users this is a rather well established search.

2. SEARCH RECOMMENDATIONS/POST-QUERY PROCESSING

Scenario:

A Europeana user searches for "Leonardo da Vinci" and receives the standard search result as well as the following search recommendations:

Leonardo da Vinci paintings

Leonardo da Vinci inventions

Leonardo da Vinci sculptures

Leonardo da Vinci journals

These recommendations are presented to the user as clickable links. The user was actually mostly interested in paintings by Leonardo da Vinci and so clicks that link and receives the matching search result.

Target user groups:

This function works for users who are less likely to make detailed precision searches and prefer to search first and then filter/narrow down search results step by step.

3. AUDIO BASED SIMILARITY SEARCH

Scenario 1: Similarity search by prior search result

A Europeana user has found a piece of music she likes through a standard keyword search. She can't identify the genre and the metadata doesn't state one or at least not one in a language she understands. So to find more pieces of music similar to the one she's found she clicks "Find similar audio".

Target user groups:

This search pattern is most common on popular on-line music sites, but is uncommon in general search engines or specialised academic/research. The user group using this search pattern may usually be listening to music often, either professionally, at home or and often via a mobile device (smartphone).

Comment:

The related items calculated by metadata and the similar audio calculated by audio indexing similarities need to be presented in a manner that's distinct and understandable to the user.

4. VIDEO SUMMARISATIONS (T2.2.7)

Scenario:

A teacher finds a number of early 20th century films that are of interest to him. In order to evaluate which films to include in his lessons he would first like to look through "trailers" or summaries of the films in order to assess their suitability.

Target user group:

The scenario fits anyone who would prefer to view a summary of a long video prior to actually viewing the whole or downloading the whole video. So it's universally intuitive to most users due to the popularity of the trailer concept.

An informative video summary should enable a user to quickly get an overview of the contents of a video and thereafter decide whether the whole video is relevant. The users of this service needs to take three things into account when searching and judging for relevance: the audio (e.g. quality), visual aspects (e.g. single or multiple key-frames or sections in order to decrease uncertainty) and linguistic aspects

5. VIDEO SIMILARITY SEARCH

Scenario: Search by similarity

A film historian is searching for early colour films and has performed a search in Europeana where he's added facets and filters limiting the search results to video only and video created before 1945. As a last step, he activates the filter "Colour film" only to limit the results.

Target user group:

In specialised image search services like Google Images results can be filtered by colour and black and white. Within popular video sites this is still a rare option though. Presented as simple filters the function should be no more difficult to use than using any filter or facet. The service can be considered as not demanding very strong search skills.

6. CONTENT CREATION BY RE-USE SERVICE

Scenario 1

"As an historian, I'm very interested in creating an enhanced pre-publication object where I can aggregate my research article with a selection of sources related to my work. I would also appreciate if my object could integrate some open comments from my peers and be notified when my article or similar ones (approaching keywords) are annotated."

Scenario 2

An amateur artist interested in Gaudi, visiting several places and taking photos of their buildings.

He wants to complement his photos with some video contents extracted from Europeana, to create a comprehensible collection of the artist's work that will be uploaded and shared through Europeana.

Scenario 3

"I am a news addict, a traveller and a hobby video maker. I regularly browse national and international news sites online, also from my mobile. I wish I could use video footage from broadcasters such as DW in order to create new videos of my own that are published somewhere.

I could include my own video material that I bring back from my travels. It would be ideal to create this new video with a tool on the site where I actually find the content that I may re-use, for example on Europeana. Even better if I could do it on the road while I travel via my mobile or in collaboration with people from my Facebook Group."

Target user groups:

These scenarios fit professional researchers, learners and amateurs alike, exploring multiple online sources and engaged in online publishing and peer consultation. It supposes an above average level of internet/technical skills.

7. TAXONOMY BASED NOTIFICATION SERVICE

Scenario 1

A university teacher is a common user of Europeana that often uses the search system to gather material (videos, images and texts) to prepare his classes. He needs to be aware of new objects related to his teaching subjects (Medieval history), so he looks at the available profiles to identify those that satisfy his needs. He makes the subscription and assigns it a significant name.

Scenario 2

A journalist has a profile in Europeana that is no longer valid because it's not aligned with his current interests, e.g. she has moved to a different section in the magazine editorial office. So, she decides to unsubscribe to that profile.

Target user groups:

This scenario is suitable for a type of user that is interested in new information/content about a certain topic and is able to clearly define this topic or match it to existing scholarly subject definitions.

8. PERSONALISATION SERVICE

Scenario 1

A user asks for all works of Renaissance artists with a preference on "Sculpture" over "Painting".

Then, the results from all works of those artists the system should present the works on "Sculpture" first, followed by the works on "Painting", and finally all other works.

Target user groups / UI implications:

In the UI, it must be assured that people can easily retrieve and change their settings. If they use the service infrequently they might not remember what their settings were which can create confusion about the rankings of search results. This can be done by an icon/text informing the user that the results are "ordered by personalization".

8.   Appendix 2: User Tests data

8.1   Participant details

Place	Age	Group	Profession
Barcelona	42	Experts	Art, Architecture and Digital Society researcher
Barcelona	27	Experts	Cultural events organiser
Barcelona	30	Experts	History researcher
Barcelona	42	Experts	History Researcher
Barcelona	39	Experts	Restaurateur
Barcelona	45	Experts	Journalist
Stockholm	39	Experts	Researcher
Stockholm	50	Experts	Researcher
Stockholm	34	Experts	Researcher
Stockholm	32	Experts	Researcher
Stockholm	52	Experts	Palaeontologist
Nicosia	30	Experts	Teacher in American College. This period she is doing research in Ancient History in Cyprus (Phinikes)
Nicosia	30	Experts	Clerical in Nicosia Municipality. Search through her job principality and her academic experience
Nicosia	35	Experts	Researcher in history. She works at University of Nicosia
Nicosia	40	Experts	He is teaching sociology. Research about culture of the societies
Nicosia	29	Experts	she is doing research program
Nicosia	30	Experts	she is teaching and doing research
Barcelona	24	General Public	English Teacher
Barcelona	36	General Public	English Teacher
Stockholm	52	General Public	Product Manager
Stockholm	44	General Public	Deacon
Stockholm	43	General Public	IT technician

Stockholm	20	General Public	Warehouse worker
Stockholm	20	General Public	Youth worker
Stockholm	36	General Public	Actor
Stockholm	64	General Public	Nurse
Nicosia	31	General Public	clerical in Minister of education and culture
Nicosia	34	General Public	she is interesting in history of theatre
Nicosia	22	General Public	she is interesting in history of art
Nicosia	24	General Public	he started his master in history and he has personal interesting in this subject
Nicosia	27	General Public	Responsible in museum of pop art
Nicosia	67	General Public	personal interesting in museums and ancient monuments
Barcelona	34	Learners	Teacher. PHD in Culture and History.
Barcelona	19	Learners	Student
Barcelona	17	Learners	Student
Barcelona	18	Learners	Student
Barcelona	30	Learners	Teacher
Barcelona	17	Learners	Student
Barcelona	24	Learners	Student
Barcelona	29	Learners	PHD History of art
Stockholm	48	Learners	Teacher
Stockholm	45	Learners	Author / Teacher
Stockholm	22	Learners	Student
Stockholm	22	Learners	Art student
Stockholm	57	Learners	Teacher
Nicosia	19	Learners	student
Nicosia	25	Learners	Teacher in the American Academy. He is interesting in classic music, traditional music
Nicosia	32	Learners	she is teaching history and Greek

Nicosia	28	Learners	She is teaching history in primary school. She is a participant in group of "History Dialogue" and work in relationships between Turkish Cypriots and Greek Cypriots
Nicosia	28	Learners	he is teaching in primary school
Nicosia	37	Learners	she is teaching history
Stockholm	28	Learners	PHD Student

8.2 Test Interviews and Questionnaires

8.2.1 Pre-test interview

1. Have you ever used the Europeana site? Do you know it?

    a. If yes, for what purpose(s)? What has been your experience?

Very positive	Positive	Slightly positive	Neutral	Slightly negative	Negative	Very negative
1	2	3	4	5	6	7

    b. If not, what would you expect from a site like the Europeana site?

2. How often do you search for information on-line?

Several times a day	Once a day	Once a week	Rarely	Never

3. Which other sites do you visit regularly related for finding information about historic monuments,   paintings or cultural heritage? (I.e. Google, Wikipedia, etc.)

    a. What kind of information are you looking for?

    b. What is your opinion of these sites?

    c. What do you like the most about them?

    d. What do you like least about them?

4. How often do you use advanced search features offered by search engines?

Usually	Often	Once a week	Rarely	Never

5. What type of content do you search more often?

6. Do you only look for information, or do you also add information or content to sites?

Usually	Often	Sometimes	Rarely	Never

a) If yes, on which type of websites do you add content?

☐ On my social media accounts (Facebook, Twitter, MySpace, Flickr, other)

☐ I follow forums and add comments to them

☐ I write blog posts and, have my own website or publish otherwise on the web

☐ I've contributed to Wikipedia or a similar service

8.2.2 Post-test interview

1. As I mentioned to you earlier, the ASSETS for Europeana project objective is to improve the current Europeana website. I will like to show how the current website results and the new version results.

[Show the participant both versions with the different pages with the results for Berlin Wall]

http://assetsdemo.atc.gr/portal/assets-rankingDemo-brief-doc.html?start=1&view=table&assets=&query=berlin+wall

a) Could you please let me know which version is more relevant for you?

b) Why?

2. You have now used the Europeana new website for around 40 minutes.

a. Was there any particular service you liked?

b. If yes, why did you like it?

c. Was there any particular service you did not like?

d. If yes, why did you not like it?

3. What, if anything, would you change or improve about the site?

4. How did your experience of using the site compare with your expectations before you started?

Far Above Expectations	Above Expectations	Slightly Above Expectations	As Expected	Slightly Below Expectations	Below Expectations	Far Below Expectations
1	2	3	4	5	6	7

+ve comments	-ve comments

5. On the basis of your experience today, how likely, or unlikely, would you be to use this site in the future?

Very Likely	Likely	Slightly Likely	Neutral	Slightly Unlikely	Unlikely	Very Unlikely
1	2	3	4	5	6	7

Why?

assets

Please answer the following questions about the tasks you just completed. Please circle the figure on the scale from fully agree to disagree entirely. Please add your comments on the space at the bottom of the page.

	Fully agree			Neutral			Disagree entirely
1. I feel in control when I'm using this site	1	2	3	4	5	6	7
2. The website used language that I felt was suitable for me	1	2	3	4	5	6	7
3. I quickly find what I want on the website	1	2	3	4	5	6	7
4. I would be able to use this website without any help	1	2	3	4	5	6	7
5. I felt confident using this website	1	2	3	4	5	6	7
6. Using this website was a satisfying experience	1	2	3	4	5	6	7
7. I was always aware of where I was within the website	1	2	3	4	5	6	7
8. I think the information provided was trustworthy	1	2	3	4	5	6	7
9. I enjoyed using the website	1	2	3	4	5	6	7
10. While using the website, I was curious to explore its content	1	2	3	4	5	6	7
11. I would recommend this site to others	1	2	3	4	5	6	7

How would you rate the following visual elements?

	Very appealing			Neutral			Very unappealing
12. Colours used	1	2	3	4	5	6	7
13. Graphics	1	2	3	4	5	6	7
14. Layout of information	1	2	3	4	5	6	7

Any final comments you would like to add?

8.2.3 Emotional Response Questionnaire

Look at the following pairs of words. Please circle the number for each pair which you think best represents the ASSETS for Europeana new website which you have experienced today.

Intuitive	1	2	3	4	5	6	7	Confusing
Clear	1	2	3	4	5	6	7	Messy
Fresh	1	2	3	4	5	6	7	Boring
Distinctive	1	2	3	4	5	6	7	Plain
Friendly	1	2	3	4	5	6	7	Impersonal
Modern	1	2	3	4	5	6	7	Old-fashioned
Simple	1	2	3	4	5	6	7	Complicated
Engaging	1	2	3	4	5	6	7	Off-putting
Structured	1	2	3	4	5	6	7	Unstructured

9. Appendix 3: Storyboards

See attachment:
D3.1.2 Appendix 3 - Storyboards.pdf