



## DELIVERABLE

**Project Acronym:** Europeana Cloud  
**Grant Agreement number:** 325091  
**Project Title:** Europeana Cloud: Unlocking Europe's Research via The Cloud

### Deliverable 1.5: Reports on all 4 Expert Fora

**Revision: Final**

#### Authors:

**Dr. Agiatis Benardou** (Digital Curation Unit – Athens)  
**Eliza Papaki , M.Phil.** (Digital Curation Unit – Athens)  
**Prof. Costis Dallas** (Digital Curation Unit – Athens)

**Kees Waterman** (KNAW - DANS)

**Björn Sjögren** (Swedish National Data Service/University of Gothenburg)  
**Dr. Stefan Ekman** (Swedish National Data Service/University of Gothenburg)

**Prof. Susan Schreibman** (Trinity College Dublin)  
**Dr. Jennifer Edmond** (Trinity College Dublin)  
**Vicky Garnett, M.Phil.** (Trinity College Dublin)

Project co-funded by the European Commission within the ICT Policy Support Programme	
Dissemination Level	
P	Public
C	Confidential, only for members of the consortium and the Commission Services

## Revision History

Revision	Date	Author	Organisation	Description
DRAFT	23 June 2015	Vicky Garnett	Trinity College Dublin	First draft with all reports to date completed
	1 <sup>st</sup> July	Vicky Garnett	Trinity College Dublin	Re-formatting
	6 <sup>th</sup> October 2015	Vicky Garnett	Trinity College Dublin	Addition of API Workshop report, and formatting.
	14 <sup>th</sup> October 2015	Vicky Garnett	Trinity College Dublin	Formatting, addition of appendices relating to final expert forum, and updates to the Executive Summary

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise.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



## Deliverable 1.5. Expert Forum Reports

### Executive summary

This report deals with four Expert Forums that took place as part of research undertaken in Work Package 1 of the Europeana Cloud (eCloud) project (2013-16). The first three of these forums dealt with content and tools and the requirements within the research communities for scholarly activity. The fourth looked at API use among the Humanities and Social Science Community, and whether or not they constitute a useful tool for researchers. This was forum held in conjunction with NeDiMAH.

The first Expert Forum took place in Dublin in June 2013 and was hosted by Trinity College Dublin, The University of Dublin. The forum aimed to investigate the uses for Europeana by researchers in both the Humanities and the Social Sciences by developing case studies that reflect the typical needs of researchers at whom Europeana Cloud will be aimed. The outcomes of this Expert Forum fed into Deliverable 3.1 in Work Package 3. The second Expert Forum<sup>1</sup> was more focused, looking at the content and tools that would be of use to Humanities researchers. This took place in Amsterdam in November 2013, and was hosted by KNAWS-DANS. The Expert Forum aimed to generate ideas for how Europeana could be developed into a useful resource for humanities research. The third Expert Forum focusing on the content and tools for Social Science research took place in Gothenburg October 2013 to assess social scientists needs and requirements for using Europeana as a research source.

The discussions at the first three Forums provided the following key recommendations for Europeana:

- functionality for adding annotations, comments, and user-enhanced metadata to records
- robust, user-friendly functionality for export to non-proprietary, ubiquitous, and/or third-party software
- refined search functionality and additional filters
- an approach to additional content that focuses on quality in few areas/subjects/topics rather than quantity in many
- improved metadata quality
- clear, easy-to-find information on Europeana, its providers, collection strategies, and inclusion (and exclusion) criteria.

Finally, the points of discussion at the last Forum meeting on API-use by researchers, which took place in The Hague in December 2014, indicated that while APIs are indeed useful for research, the barriers to using them among the wider Humanities and Social Science community remain

---

<sup>1</sup> The second and third expert forums were originally planned in the Europeana Cloud Description of Work to take place in that order. However, due to unforeseen circumstances, it became necessary to change the order in which these two Expert Forums took place. Despite this, in order to remain consistent with the DOW, we have retained their original numbering within this report.

insurmountable for some. These included a lack of knowledge about what APIs can do, and a lack of time to be able to find out. The outputs from projects that can harvest data using APIs, as shown by some of the participants, can be impressive, however. Other providers of APIs also discuss the responsibility and difficulties that arise when applying API access to digital collections.

# Table of Contents

REVISION HISTORY .....	II
EXECUTIVE SUMMARY .....	IV
TABLE OF CONTENTS .....	VI
LIST OF FIGURES .....	IX
EXPERT FORUM 1 - 'CASE STUDIES', TRINITY COLLEGE DUBLIN, JUNE 2013 .....	1
1. THE PURPOSE OF THE EXPERT FORUM .....	1
2. IDENTIFYING DISCIPLINARY AREAS TO DEVELOP CASE STUDIES.....	1
3. IDENTIFYING THE EXPERTS .....	1
4. DESIGNING THE EXPERT FORUM.....	2
4.1 Session 1 - Europeana Treasure Hunt.....	2
<b>THE THREE TASKS WERE DEVELOPED TO HIGHLIGHT SEARCH FUNCTIONALITY AND METADATA PRESENT IN THE CURRENT INSTANTIATION OF EUROPEANA. THE EXERCISE CAN BE FOUND IN .....</b>	<b>2</b>
4.1.1 Multiple Search Terms for the Same Named Entity.....	3
4.1.2 Minimum Metadata .....	3
4.1.3 Europeana-Whacks .....	3
4.2 Session 2 - What Research Can Europeana Support in its CURRENT form?.....	3
4.3 Session 3 - Future Possibilities of Europeana as a Research Platform .....	4
5. EXPERT FORUM FEEDBACK .....	4
5.1 Insufficient Metadata .....	4
5.2 User Ranking.....	4
5.3 Multi-lingual Resources.....	4
5.4 Date Ranges not Recognised .....	5
5.5 Spatial and Temporal Mapping of Results.....	5
5.6 More Transparent Citation Methods .....	5
5.7 Tool Development.....	5

5.8	Europeana as a Teaching Resource .....	5
<b>6.</b>	<b>CASE STUDIES .....</b>	<b>5</b>
	Case Study 1: Early Stage Historian using Europeana in its Current Form (History) .....	6
	Case Study 2 – Experienced Computer Scientist using Europeana in its current form (Ontologies).....	6
	Case Study 3 – An Experienced Lecturer in Sociology using an enriched Europeana (Social Sciences) .....	7
	Case Study 4 – An early-stage career Archaeologist using an enriched Europeana (Archaeology).....	8
<b>7.</b>	<b>CONCLUSIONS FROM EXPERT FORUM 1 .....</b>	<b>8</b>
7.1	Recommendations for Europeana.....	8
7.2	Recommendations for Future Expert Forums.....	8
	<b>EXPERT FORUM 2 – TOOLS AND CONTENT FOR HUMANITIES RESEARCH, KNAWS-DANS, AMSTERDAM, NOVEMBER 2013.....</b>	<b>10</b>
<b>8.</b>	<b>INTRODUCTION TO THE TASK .....</b>	<b>10</b>
8.1	Expert Forum 3 in the context of Work Package 1.....	10
8.2	The purpose of the forum .....	10
8.3	How the results will be used .....	11
<b>9.</b>	<b>SELECTING THE RESEARCH AREAS AND THE PARTICIPANTS .....</b>	<b>11</b>
9.1	Research areas.....	11
9.2	Selecting the experts.....	11
9.3	Participants .....	11
<b>10.</b>	<b>DESIGNING THE FORUM .....</b>	<b>12</b>
10.1	Europeana Treasure Hunt.....	12
10.2	Session 1 – Tools.....	13
10.3	Session 2 – Content .....	13
<b>11.</b>	<b>RESULTS.....</b>	<b>13</b>
11.1	Tools .....	14
11.1.1	Summing up the Tools session discussions.....	17
11.2	Content.....	18
11.2.1	Summing up the Contents session discussions.....	20

<b>12. CONCLUSIONS FROM EXPERT FORUM 2 .....</b>	<b>21</b>
<b>EXPERT FORUM 3 - TOOLS AND CONTENT FOR SOCIAL SCIENCE RESEARCH, UNIVERSITY OF GOTHENBURG, OCTOBER 2013.....</b>	<b>23</b>
<b>13. INTRODUCTION TO THE TASK .....</b>	<b>23</b>
13.1 Expert Forum 3 in the context of Work Package 1 .....	23
13.2 The purpose of the forum.....	23
13.3 How the results will be used.....	23
<b>14. SELECTING THE RESEARCH AREAS AND THE PARTICIPANTS .....</b>	<b>24</b>
14.1 Research areas.....	24
14.2 Selecting the experts.....	24
14.3 Participants.....	24
<b>15. DESIGNING THE FORUM .....</b>	<b>25</b>
15.1 Europeana Treasure Hunt.....	25
15.2 Session 1 - Tools.....	25
15.3 Session 2 - Content .....	26
<b>16. RESULTS.....</b>	<b>27</b>
<b>16.1 Tools .....</b>	<b>27</b>
16.1.1 Current use of tools .....	27
16.1.2 Tools for working with Europeana now .....	28
16.1.3 Tools for working with Europeana in the future .....	28
16.1.4 Summing up the Tools session discussions.....	29
<b>16.2 Content.....</b>	<b>30</b>
16.2.1 Useful Europeana content now .....	30
16.2.2 Europeana as a source of research material .....	30
16.2.3 Europeana as a transparent research source .....	33
16.2.4 Summing up the Contents session discussions.....	34
<b>17. CONCLUSIONS FROM EXPERT FORUM 3 .....</b>	<b>34</b>
<b>EXPERT FORUM 4 – “SO WE’VE BUILT IT, BUT HAVE THEY COME? INVESTIGATING BARRIERS AND OPPORTUNITIES FOR API USAGE AMONG THE AHSS COMMUNITY” .....</b>	<b>36</b>
<b>18. INTRODUCTION .....</b>	<b>36</b>
<b>19. BACKGROUND .....</b>	<b>36</b>



<b>20. STORYBOARDING .....</b>	<b>37</b>
20.1 Newspaper API Builder .....	37
20.2 Image Clusters.....	37
20.3 The Allusionator.....	38
<b>21. RECOGNISING THE BARRIERS: ENABLING THE PARTICIPANTS .....</b>	<b>39</b>
21.1 Unknown Unknowns and Very Clever Things.....	40
<b>22. ACKNOWLEDGEMENTS FOR THE E-CLOUD/NEDIMAH API WORKSHOP.....</b>	<b>40</b>
<b>APPENDIX I – LIST OF PARTICIPANTS - DUBLIN .....</b>	<b>41</b>
<b>APPENDIX II – EXPERT FORUM RUNNING ORDER - DUBLIN .....</b>	<b>42</b>
<b>APPENDIX III – EUROPEANA TREASURE TRAIL – DUBLIN.....</b>	<b>43</b>
<b>APPENDIX IV – LIST OF PARTICIPANTS - AMSTERDAM.....</b>	<b>45</b>
<b>APPENDIX V – AGENDA OF THE EXPERT FORUM - AMSTERDAM .....</b>	<b>46</b>
<b>APPENDIX VI – THE EUROPEANA TREASURE HUNT - AMSTERDAM.....</b>	<b>47</b>
<b>APPENDIX VII – LIST OF PARTICIPANTS EXPERT FORUM 3 – GOTHENBURG .....</b>	<b>48</b>
<b>APPENDIX VIII – AGENDA OF THE EXPERT FORUM - GOTHENBURG.....</b>	<b>49</b>
<b>APPENDIX IX – THE EUROPEANA TREASURE HUNT - GOTHENBURG .....</b>	<b>50</b>
<b>APPENDIX X – RUNNING ORDER OF E-CLOUD/ NEDIMAH API WORKSHOP – THE HAGUE.....</b>	<b>51</b>
<b>APPENDIX XI – LIST OF PARTICIPANTS AT E-CLOUD / NEDIMAH API WORKSHOP – THE HAGUE ....</b>	<b>52</b>

## List of figures

Figure 1 - "Newspaper API Request Builder" storyboard from the Linguistics group.....	37
Figure 2 - Image clustering storyboard from the History group.....	38
Figure 3 - Images of the Allusionator.....	39

# Expert Forum 1 - 'Case Studies', Trinity College Dublin, June 2013

## **1. The Purpose of the Expert Forum**

The Expert Forums take place as part of Work Package 1 of Europeana Cloud. There are four Expert Forums in total, each addressing a particular area of research.

Expert Forum 1 took place in Month 5 of the project (June 2013) and was co-ordinated by Trinity College Dublin. The forum aimed to investigate the uses for Europeana by researchers in both the Humanities and the Social Sciences by developing case studies that reflect the typical needs of researchers at whom Europeana Cloud will be aimed. The case studies developed through the outcomes of this Forum examined both Europeana in its current form, and at Europeana as it could be as it moves into developing cloud services. The outcomes of this Expert Forum feed into Deliverable 3.1 in Work Package 3.

Expert Forums 2 and 3 will take place in Months 9 and 10 of the Project (October and November 2013) and will look specifically at tools that could be developed within Europeana Cloud for the Humanities (Expert Forum 2) and the Social Sciences (Expert Forum 3).

Expert Forum 4 is scheduled to take place by Month 30 (July 2015) of the project, and will provide a broad review of the tools and content access and use services provided by Europeana Cloud, and develop recommendations for future work, including how the engagement of researchers will continue beyond the lifetime of the project to ensure their future use and uptake of the Europeana Research platform.

This report focuses on the outcomes of Expert Forum 1 "Case Studies".

## **2. Identifying Disciplinary Areas to Develop Case Studies**

In identifying research areas targeted in this Expert Forum, it was decided not to invite experts from many different disciplines as it was felt researchers may have too little in common to develop useful case studies. Therefore four key research areas were selected that would between them use a variety of data types.

The academic areas selected were Humanities disciplines History (focusing on context-based scholarship) and Archaeology (focusing on images, 3-D, and images), with broader areas of Social Sciences (focusing on datasets) and Ontologies (focusing on the metadata itself). All the focus groups, in addition to specific datatypes, dealt with metadata in Europeana.

## **3. Identifying the Experts**

The invitees to the forum were identified to ensure the broadest range of expertise and experience. They were also selected to create a European-wide perspective. Participants came from institutions in Denmark, Greece, Ireland, Lithuania, The Netherlands, Norway, Portugal, Sweden, and the UK.

## Europeana Cloud

The group comprised of experts from eCloud-related institutions and those from institutions who were not directly involved with the Europeana Cloud project. In addition a member of the eCloud Research Communities Advisory Board, Professor Rob Kitchin, was invited to participate.

### **4. Designing the Expert Forum**

The intention of the day was to establish an overview of the research that can be conducted using Europeana currently, and develop this to include future possibilities that might be realised as Europeana moves into a Cloud-based service.

The results of Expert Forum 1 will feed directly into the work being conducted in Work Package 3. In particular, they will be analysed for use in Deliverable 3.1:

#### Deliverable 3.1

In close alignment with the tasks in WP1, KUL will lead on developing personas (descriptions of typical researchers that we address with this project), scenarios and use cases that describe in details what kind of tool a researcher would typically use, and how that use would fit in his/her typical workflow. This will result in Deliverable 3.1 (M6).

The aim of this Expert Forum, therefore, was to task the experts from within their groups to develop scenarios for specific researchers at key stages in their careers and within certain disciplines, focusing on the resources and tools they use, and the methodologies they employ.

In order to meet this objective, the day was broken up into three main sessions to create a structure in which a healthy discussion could be achieved that would result in useful scenarios that could be developed into case studies.

#### **4.1 Session 1 - Europeana Treasure Hunt**

The first session served as both an ice-breaker and gamified the task of introducing researchers to Europeana, both in terms of Europeana's content, functionality, and special exhibitions. As with the eCloud kick-off meeting, participants were divided into groups with cloud icons. These icons divided participants into the domain groups for Archaeology (Stratus), History (Cumulus), Ontologies (Stratocumulus), and Social Sciences (Altocumulus). But for this ice-breaker, sub-teams formed which had one cloud per group so that participants had the opportunity to meet participants outside their subject area.

The task itself was entitled 'Europeana Treasure Hunt'. Teams were asked to complete three Europeana-based challenges within 15 minutes. All teams were asked to create a Europeana Profile. This allowed those who had not used Europeana before to familiarise themselves with the user profile function for collecting and saving searches and results.

***The three tasks were developed to highlight search functionality and metadata present in the current instantiation of Europeana. The exercise can be found in***

Appendix III – Europeana Treasure Trail.

### 4.1.1 Multiple Search Terms for the Same Named Entity

The purpose of this task was to highlight Europeana search functionality and associated issues when metadata for key word terms is not controlled, such as in the case of the Great War. This is an issue not simply within a single language, and is only compounded when searching across multiple languages. One group found 27 cognate terms for this event, each resulting in a different data set being returned

### 4.1.2 Minimum Metadata

Many of the items within Europeana do not contain the full complement of metadata. In some cases, only the name of the item and the source is provided. The purpose of this task was to enable Expert Forum members to familiarise themselves with the kinds of metadata typically provided by contributors, and to get a feel for the shortage or completeness of metadata typically provided.

### 4.1.3 Europeana-Whacks

The ‘Europeana-Whacks’ task was based loosely on the practice of trying to find a ‘Google-whack’ whereby a search term produces only one result in Google. The purpose behind the ‘Europeana-Whacks’ task was to allow the Expert Forum members the opportunity to see if they could find something very specific within the database. Often a researcher doesn’t want to manually filter through the results to find the most relevant item. This task was to establish if that can indeed be done, and what difficulties the task may present.

## 4.2 ***Session 2 - What Research Can Europeana Support in its CURRENT form?***

In order to identify areas for development to make Europeana more useful for Researchers in future, it was necessary for Expert Forum members to understand the current scope of Europeana. Each group was asked to create a case study in which a fictional Researcher might engage with Europeana as it is currently implemented. To do this each group was invited to complete a Case Studies template .

The information the groups were asked to complete included:

- The research task/ goal of the hypothetical researcher
- Their career/ experience level (both disciplinary and within Digital Humanities)
- Discipline Research Area
- Resource/Data Type
- Tools to be used
- Methodologies
- Problems they might encounter

Participants were instructed in this session to *only* develop case studies that utilised the current functionality of Europeana resources; i.e. metadata.

### **4.3 Session 3 – Future Possibilities of Europeana as a Research Platform**

The last session of the day further developed the case studies developed during the second session. Only now, participants were instructed to imagine that there were no restrictions as to what Europeana might provide in terms of content types. Moreover, they were instructed to envisage the types of services that could be useful in the analysis of these data types especially within a cloud platform.

## **5. Expert Forum Feedback**

After each breakout session there was ample opportunity for the group to reconvene as a whole and discuss the observations and outcomes generated by the disciplinary-based groups. These conversations have been consolidated under the following broad subject headings.

### **5.1 Insufficient Metadata**

The overall response to Europeana in its current form is that the majority of items do not have sufficient metadata to make it a truly usable tool for research. The users in the Expert Forum commented that in the majority of cases, searches for a targeted term resulted in a vast number of unrelated hits or, when the hits were relevant, the lack of metadata of individual items made the result set almost impossible to work with. Moreover, images were frequently missing and the date provided would often be inaccurate (or at least not relevant), while lexical ambiguity would bring up ‘wildcard’ results.

The group as a whole found the lack or partial lack of metadata frustrating. The overwhelming consensus was that this was one of the major factors that prevented the usefulness of Europeana as anything more than a discovery tool.

### **5.2 User Ranking**

One method for addressing this issue was an incentive scheme, whereby a ‘star-ranking system’ could be used to rank the metadata provided. Users of Europeana could rate an item for the quality of its metadata from 1 star (poor quality) up to 5 stars (excellent quality). The ranking system would take into consideration the completeness of the data, the quality of the data (including provenance and providing institution), as well as a description of the item, having at least metadata for these fields comply with the best practices in digital curation.

While the individual items could be ranked, the contributing organisations could also then be ranked according to the quality of their metadata. This would be reviewed annually to encourage contributors to improve where needed.

### **5.3 Multi-lingual Resources**

At present there are no translational services in Europeana. As metadata can be submitted in the native language of the contributing organisation, a user must search in every language to be sure she has done a thorough search for a particular term. For example, if a researcher types in ‘World War One’, then she will receive results with the specific phrase ‘World War One’ in the metadata, but won’t get results with ‘La Première Guerre Mondiale’ (French) or ‘Pierwsza Wojna Swiatowa’ (Polish).

## Europeana Cloud

A key recommendation of the Expert Forum is that multilingual metadata should be dynamically generated by the system as it is impractical to ask content providers to provide multi-lingual search terms. It was strongly recommended that a thesaurus or translation tool be built into Europeana that can automatically link multilingual metadata to enable researchers to find all items associated with a search term.

### **5.4 Date Ranges not Recognised**

Many users, particularly in History and Archaeology, found when entering date ranges into the search field that only the specific years they entered were being returned. For example, a query on items from the '1800s' returned only items from the year 1800, or items in which the number '1800' was specifically mentioned as a quantity. Similarly, when entering a date range such as '1850 – 1950', only those two years would be specified in the results.

### **5.5 Spatial and Temporal Mapping of Results**

Participants unanimously recommended that Europeana needed better browsing functionality. Suggestions included spatial mapping and temporal mapping via maps and timelines, respectively. Simple conceptual maps, such as tag clouds or other visualisations would also be desirable.

### **5.6 More Transparent Citation Methods**

Forum participants agreed that a more transparent and reliable citation method was needed to both allow users to cite the results of their searches so that they could be replicated or validated by other Researchers, and to cite specific items found within Europeana.

### **5.7 Tool Development**

There was overwhelming consensus that Researchers from all disciplines would want to export result sets into specialised, often domain-specific tools. Therefore, it was strongly recommended that Europeana not develop specialised tools, but rather focus on

- generalised tools (as mentioned above) to aid in discovery
- tools that allow for (meta)data export into a variety of formats
- tools that provide more sophisticated and targeted filtering of results than are currently available.

The group felt strongly that Europeana should not concern itself with developing discipline-specific tools. While these would be expensive to develop and would serve only a small subset of users, it would also require that Researchers learn yet another piece of software. Moreover, it is also likely that in the time it would take to develop these tools, discipline-specific tools will have moved on, placing Europeana in a game of constant catch up.

### **5.8 Europeana as a Teaching Resource**

A strong recommendation was made for Europeana to strengthen its resources for teachers, allowing for lesson plans and interactive functions that could be used within the classroom, or be imported into an eLearning environment such as BlackBoard or Moodle.

## **6. Case Studies**

The following case studies present hypothetical researchers working in typical research scenarios across the EU as developed during the breakout sessions.

### ***Case Study 1: Early Stage Historian using Europeana in its Current Form (History)***

Mary is an early stage History researcher at MPhil level, who is reviewing the background information for her thesis proposal. She intends to discuss how the physical form of printed bibles changed during the reign of Queen Victoria. Mary typically uses search engines in browsers or in specialised databases, such as her library subscribes to, but she is not familiar with the technology behind these search engines. She has no experience of XML or programming. She is comfortable with drag-and-drop facilities, and tools such as 'Evernote'.

Much of her data is in the form of images, for example of the Bible, bindings and the metadata that would support such images. She also makes use of text taken via OCR software from Bibles in the Victoria Era, text of Bible commentaries, reading plans and other materials, Sunday school teaching materials such as books and pictures, Church inventories and information from bookseller' catalogues and advertisements.

In conducting her research, Mary plans to use referencing tools, both to save the results of searches, but also to populate her final written thesis. These include open source applications such as Evernote, Zotero and Endnote. She also plans to use geo-referencing tools in her research, such as CreateMap, and also wishes to publish any images she may find. She therefore needs to ensure that her sources for the images allow for reproduction in her thesis without breaching copyright laws.

The methodology for her research will typically involve communication with researchers in similar fields, data capture and analysis, publishing and dissemination, and data structuring through means such as cataloguing and indexing. She will employ methods that will allow her to use existing data for collation, image-segmentation and text recognition.

She therefore needs Europeana to integrate well with open-source tools such as EverNote and EndNote for citation purposes, as well as to allow use of image files in a drag and drop function. The records also need to be trusted. The problems she is encountering, however, are mainly due to a lack of insufficient records of analogue items that would typically fall into the scope of her research. While searching for items in Europeana that were produced during the Victorian era, she is also experiencing difficulty in finding items within a range of dates. Her searches for items produced during the '1800s', or from '1837-1901' produce limited results, often restricted to the specific years she has typed, and not to the wider range, as she hopes. The provenance information in the metadata of many of the items she is finding is not complete, and therefore does not provide her with sufficient background to either reliably use the items as data sources for analysis, or to use as evidence to support her arguments.

### ***Case Study 2 – Experienced Computer Scientist using Europeana in its current form (Ontologies)***

Jonathan is an experienced computer scientist working in development. He wants to be able to write SPARQL queries to achieve two key goals: 1) for discovery/info seeking, and 2) for visualisation of information e.g. geographical based information. He wants to then create a list of dataset collections, which have more than 10 numbers of items with certain rights qualification. The project aims to visualise the datasets of Europeana in order to inform Europeana content strategy but also to assure researchers as to the quality and depth and coverage of the research

## Europeana Cloud

they are performing. This should result in tools that visualise coverage, number of records e.g. related to a particular area.

Jonathan will be working mainly with RDF schema, and wants to use SPARQL for the analysis, and CARARE for the visualisation of his results.

Currently, the metadata on Europeana in its current form is almost useless for this method of analysis. Jonathan finds that his research is precisely the research needed to review the current state of the Europeana dataset, however the dataset itself is the biggest restriction to his research. He finds that he can't cite the state of the data, for example, he doesn't know what has been excluded from the metadata, as it does not comply with VoID, the Vocabulary of Interlinked Datasets, which sets a standard for metadata use. This would be especially useful in Jonathan's research using RDF.

### ***Case Study 3 – An Experienced Lecturer in Sociology using an enriched Europeana (Social Sciences)***

Nicola is a Senior Lecturer in Sociology at a University in 2017. She uses Europeana to develop lesson plans for her students on her BA Hons in Sociology course. She is teaching a module on migration, both immigration and emigration, asking her students to particularly look at the reasons people move, where they move from and to, and who they are. Parts of the course are delivered online through eLearning.

Much of the data she uses is taken from trusted Europeana contributors providing migration data on both Immigration and Emigration stories from the past 50 years. Europeana has already taken care of any data protection issues in regards to privacy choice and consent. The majority of this data is in text form, which she extracts for analysis by her students during lectures. Before she can do this, however, she likes to review the content of the data she finds on Europeana using an open source tool that is compatible with the Europeana infrastructure.

She also searches for items such as photographs, letters, diaries, travel manifests, travel documents, and other items that might be associated with migration patterns.

The tools she regularly makes use of are face-recognition searches in photos and video, such as 'imagevision', mapping of results, a timeline of results, the full-text search within the content of the items and the export function, particularly for use in spreadsheets (e.g. MS Excel) and another open source software package of her choosing.

Nicola knows that the data she is extracting is trusted, as it complies with the best practice for the storing and maintenance of personal information across the EU. She knows that Europeana regularly runs queries with the contributors to check that they are compliant, and she is also able to run this check herself at intervals.

She is able to select the most useful items on Europeana, thanks to the star-ranking system that has been introduced. She makes sure to rank any items she uses herself, in order to assist her fellow researchers in selected high-quality items. She also knows which contributors across the EU have the highest rankings for most consistent metadata, and often this reduces the amount of time she



## Europeana Cloud

spends searching for items, as she can select which contributors' content to search. She is also able to filter the results of her searches down to gather the most relevant items.

### ***Case Study 4 – An early-stage career Archaeologist using an enriched Europeana (Archaeology)***

Franz is a German postdoctoral fellow working on a project that will digitally recreate the Parthenon using 3D data, images and texts.

He recently completed his Doctoral Degree and has a reasonable knowledge of digital tools, but it still learning new technologies. He therefore relies on the basic tools available in Europeana that allow him to export the data he needs into the digital tools with which he is familiar. He enjoys the ability to preview 3D media objects before he analyses them for content, and is able to find the relevant metadata he needs by filtering the results and using the mapping and timeline visualization functions in his search.

As well as his native German, Franz speaks English, Greek, and some Ancient Greek and Latin. However, he is not familiar with other languages in Europe, and therefore is not able to directly translate search terms. Thankfully he is able to type a search term into Europeana in one of the languages he does speak, and is able to find useful metadata from any country, as Europeana has an automatic translation function that searches for multilingual terms.

## ***7. Conclusions from Expert Forum 1***

### ***7.1 Recommendations for Europeana***

It is clear from this Expert Forum that participants do not consider Europeana a truly useful or robust resource without better quality metadata and a suite of user tools to aid browsing and analysis. Moreover, it is essential that this metadata can be viewed, filtered, and manipulated in a variety of ways.

Key recommendations include:

- encouraging Europeana contributors to provide complete metadata in future
- the development of metadata enrichment tools for current incomplete metadata
- providing simple visual tools to enhance browsing (eg maps, timelines, tag clouds)
- providing export mechanisms in a variety of metadata formats for Researcher-generated metadata sets

### ***7.2 Recommendations for Future Expert Forums***

The three-session format of the day proved successful in terms of gathering feedback, encouraging discussion, and gaining wide community participation. Despite the very different disciplines and levels of expertise, the feedback from the disciplinary groups was generally in agreement.

The length of the Forum worked well, as it was long enough to promote debate, but succinct enough to keep the participants focused on the task ahead of them. This particular forum took place over one full day, and due to the focused delivery was able to keep to the allotted time and allow participants to travel home afterwards. However it is also proposed that for future Expert Forums, two half-day sessions (beginning with lunch on Day 1) could be more desirable.

## **Europeana Cloud**

Consideration will have to be given to the division of sessions and tasks across the two days in order to maintain momentum of debate and keep the participants focused. However, the evening break could also serve to refresh the participants for the session the following morning.

Gathering participants was by and large the most challenging aspect of the Expert Forum delivery. The organizers therefore strongly recommend that identification and invitation of participants in future Expert Forums should be conducted at least four months in advance. It is also recommended that complete invitation lists with the responding RSVPs should be kept for reference when coordinating subsequent Expert Forums.

## **Expert Forum 2 – Tools and Content for Humanities Research, KNAWS-DANS, Amsterdam, November 2013**

### **8. Introduction to the task**

This report focuses on the outcomes of Expert Forum #2 held in Amsterdam, November 11-12 (project month 10). It is the third of four Expert Forums within Work Package 1 of the EU project “Europeana Cloud: Unlocking Europe’s Research via The Cloud”. The event was organized by two institutes of the Royal Netherlands Academy of Arts and Sciences (KNAW), a partner in the project: DANS (Data Archiving and Networked Services) and NIOD (Institute for War, Holocaust and Genocide Studies). NIOD hosted the event.

#### **8.1 Expert Forum 3 in the context of Work Package 1**

The general purpose of Work Package 1 is to assess the researchers’ needs for future work with the Europeana content. To do this, the work package encompasses a number of tasks and subtasks employing different investigation methods such as desk research, web surveys, and case studies. In addition, four Expert Forums will be held in the course of Work Package 1, each providing input from external experts on different topics.

Expert Forum 1 took place in Dublin, June 18 (project month 5), earlier this year. The Dublin forum focused on the typical needs for researchers in the humanities and social sciences to be able to work with the Europeana content, illustrated in the three created cases.<sup>2</sup>

Expert Forums 2 and 3 (held in Amsterdam, project month 10, and Gothenburg, project month 9<sup>3</sup>) are to be seen as complementary in that they have much the same focus and agenda, although each explored more thoroughly the humanities and the social sciences, respectively.

Finally, the fourth Expert Forum is scheduled for July in 2015 (project month 30). This forum will provide a broad review of the tools and content access and use services provided by Europeana Cloud, and will develop recommendations for future work, including how the engagement of researchers will continue beyond the lifetime of the project to ensure their future use and uptake of the Europeana Research platform

#### **8.2 The purpose of the forum**

The Expert Forum aimed to generate ideas for how Europeana can be developed into a useful resource for humanities research. The discussions focused on what digital tools would benefit research on the Europeana content (metadata, text, images, sound, video, 3D) and how Europeana can be improved as a source of research material.

---

<sup>2</sup> See Deliverable 1.5 (1 of 4): *Expert Forum Case Studies Report*.

<sup>3</sup> Due to planning exigencies, Expert Forum 3 was held one month prior to Expert Forum 2. The original numbering sequence has been retained to comply with the description in the project’s DoW.

### **8.3 How the results will be used**

This report forms part of Deliverable 1.5 together with the three accompanying Expert Forum reports.

## **9. Selecting the research areas and the participants**

### **9.1 Research areas**

Research in the humanities covers a wide array of (interdisciplinary) topics, approaches and methodologies. For the purpose of this Expert Forum, it was necessary to select participants from selected disciplines of the humanities. During the work with Deliverable 1.1, six subject domains were identified as most likely to find Europeana material useful in their research.<sup>4</sup> These subject domains were also instrumental in understanding the humanities in the context of the Expert Forum. They informed the selection process, but other variables were also considered when selecting expert participants. One such variable was the variety in research areas a particular participant could address; it was felt that a researcher with experience from several disciplines in the humanities (especially through interdisciplinary research) would be in a position better to contribute to the discussions. Another variable was familiarity with methodologies in the 'Digital Humanities (DH); researchers who did not employ digital tools in their work would probably contribute less to the discussion than researchers well versed in deploying digital research.

### **9.2 Selecting the experts**

Potential external experts were identified by desk research and through recommendations from colleagues, both from within and outside the project. Once recognized as potentially relevant for the forum, prospective participants received an email, briefly describing the Europeana Cloud project and the general aims of the forum. Those who responded that they would like to take part were then provided with detailed information. Although the first round of invitations were sent out by late July, around 60 percent of the prospective participants were unable to accept the invitation, due to agenda complications. Still, they reacted positively to our invitations by showing great interest in the forum and the project in general, and quite a number of them suggested other suitable experts.

The final pool of experts, assembled for the Expert Forum was successful in meeting most of the organizational criteria. The main humanities research areas represented by the researchers were: archaeology, cultural heritage studies, history, musicology and philosophy.

### **9.3 Participants**

The participants of the Expert Forum were either people from Europeana Cloud-related institutions or external experts from institutions not directly involved with the project.<sup>5</sup> Karina van Dalen-Oskam represented the Research Community Advisory Board of Europeana Cloud. Participants hailed from institutions from several different EU countries, providing a broad European perspective. Institutions from the following countries were represented in the forum: Greece,

---

<sup>4</sup> These were, in alphabetical order, Archaeology, History, Law, Linguistics, Musicology and Philosophy.

<sup>5</sup> See Appendix IV – List of participants - Amsterdam for a complete list of participants.

## Europeana Cloud

Germany, Ireland, the Netherlands, Sweden, Switzerland and the UK. The Forum greatly benefitted from the participation by the overall project manager of Europeana Cloud, Alastair Dunning. His attendance allowed for direct interaction between external experts, Work Package participants and the project's own views on planning, progress and orientation.

## 10. *Designing the forum*

Implementing one of the practical recommendations from the report of the Dublin Expert Forum, the forum in Amsterdam, like the one in Gothenburg, took place during two half-day sessions, beginning in the afternoon of November 11 and ending at midday the following day.<sup>6</sup> The forum started with a short introduction of Europeana and eCloud by the Alastair Dunning and the leader of Work Package 1, Agiati Benardou. After that there was an icebreaker activity that was also meant to ensure that all participants had recent experience with the Europeana portal, followed by two main sessions focusing on tools (day 1) and content (day 2), respectively.

In preparing for the forum, each expert was requested to create a user account for the Europeana portal as it presently exists.

### 10.1 *Europeana Treasure Hunt*

The icebreaking kick-off activity, dubbed "Europeana Treasure Hunt", was designed primarily to provide the participants with hands-on experience with some key aspects of the Europeana portal and content. It used the same structure as in Gothenburg, as adapted from the one deployed in Dublin; the adjustments were intended to better align the activity with the overall agenda of these Expert Fora on tools and content.

The participants were divided into four groups of four or five participants, each consisting of a mixture of people from within the eCloud-project and external experts. The first assignment for each team was to create a Europeana profile. Those who had never used Europeana before were thus introduced to the user profile function, enabling a user to save previous searches. Having created a Europeana profile, they were asked to work for 15 minutes with three assigned tasks.<sup>7</sup>

The purpose of the first task was to familiarize the participants with the five main content types of Europeana by asking them to formulate searches that gave results containing hits with all five content types (text, images, sound, video, 3D). The second task introduced them to the metadata structure currently used in Europeana. Participants were asked to find as many metadata fields as possible, yielding a rough overview of existing metadata fields. The objective of the third task was to find the lowest possible number of search results. The purpose of this task was to give the participants the opportunity to experience the search tool and develop insights into its current functionalities.

---

<sup>6</sup> See Appendix V – Agenda of the Expert Forum - Amsterdam for the entire agenda.

<sup>7</sup> The complete instructions given to the participants can be found in Appendix VI – The Europeana Treasure Hunt - Amsterdam.

## 10.2 Session 1 – Tools

Session 1 focused on the kinds of tools that would be useful, and could potentially be developed, for humanists carrying out research on the Europeana material, current and prospective. Sessions 1 and 2 consisted of group discussions followed by a summing-up session with all participants. For the breakout session, the participants were divided into four groups. Each group consisted of people both from within and outside the project, and was intended to contain experts from related disciplines and communities. Participants with an expertise in deploying corpora of digital texts and those engaged in the broadly defined field of ‘digital humanities’ were spread out over the groups; the orientation of the other experts concentrated on archaeology-GIS, philosophy-metadata, history-Audiovisual and musicology, respectively.

A short introduction was given before Session 1, in which the participants also had the opportunity to ask clarifications. To provide a structure to the breakout session, participants were supplied with five ‘guiding questions’:

1. What tools do you use for your *own* scholarly activity?
2. What tools do you use when you *collaborate* with fellow researchers?
3. How would you use Europeana content in your research?
4. What tools do you think you could use *now* with Europeana content?
5. What tools need to be *developed* to enable you to use Europeana content in your research?

eCloud representatives in each group made notes of discussions in the breakout sessions, aiming to provide structured input for the compilation of this report.

## 10.3 Session 2 – Content

Session 2 dealt with the content of Europeana and how it should be developed, in part building on the results from Session 1. The groups from Session 1 the previous day were kept for Session 2. Again, there was a short introduction, and then the following main four questions were given to each group:

1. Is there content in Europeana that is useful currently? What content?
2. Does this content need to be improved on or added to?
3. What new content (genres, formats) would you like to see added to Europeana?
4. What are the biggest gaps in content in Europeana from a humanities perspective?

The following remark was supplied with these questions, providing a possible structure for exchanging ideas and organize feedback:

Current main types of content in Europeana are text, image, video and sound. The fifth content type, 3D visualizations or constructs, is still relatively scarce in Europeana.

## 11. Results

This section presents the results from the discussions on Tools (Session 1) and Content (Session 2). Findings of the four groups as well as comments made during the summing-up sessions afterwards are combined in these results.

## Europeana Cloud

Participants' discussions in the groups tended to switch between the various main topics, despite the 'guiding questions'. To illustrate suggestions or remarks, specific tools and projects are provided throughout. It should be pointed out that these are intended solely to strengthen the evidence base for Europeana Research's consultation; i.e., numerous other examples could be identified, and they should not be interpreted as exhaustive listings.

### 11.1 Tools

The first session focused on tools; what tools do the experts use for their own scholarly activity, and what tools could they use, now and in the future, with the Europeana content?

Starting with the **first** 'guiding question': besides widespread, prevalent tools that are currently available to any professional consuming and handling steadily growing information streams, the participants also mentioned various specially designed tools, for specific projects or research activities. The general tools included Google, Microsoft Office products and open source alternatives, and other mainstream tools. Honing in more closely on the professional activities of the experts, various general research tools were mentioned. These included Open Refine (formerly, Google Refine), R as an interface to query and analyze text corpora statistically, tools for authorship attribution, and GIS applications (or QGIS, an open source alternative). In addition, a number of annotation tools were mentioned.

For specific research clusters, the following were among the tailor-made tools that participants used or (had) developed: for some subjects within biblical studies, a geographic visualization tool, 'eResearch'; for oral history (interviews, audio-visual), a bilingual platform for voice and speech recognition, with full transcription and OCR functionalities, materials will be exported to XML and saved in a data archive in the Netherlands; in musicology, a self-built tool to help with OMR processes (Optical Music Recognition) applied to 16<sup>th</sup>-century music manuscripts and early printed music.

A number of the general research tools mentioned above also lend themselves as tools enabling collaboration between researchers (**second** 'guiding question'). More specifically, many participants indicated that they heavily used cloud-based services and tools, like Google Drive, Google Docs and Dropbox. Numerous participants voiced reservations regarding the proprietary issues with these services, but still found them to be best suited for their current needs. Other mainstream tools used in collaborative work with fellow researchers were Facebook and other social networking sites such as Twitter and LinkedIn.

Collaboration instruments more geared towards research requirements included the academic blogging website [www.hypotheses.org](http://www.hypotheses.org), and Pinterest for assembling image collections. The previously mentioned oral history platform will be opened up to the general public as a collective access and annotation tool, for free, and the OMR-tool for early music has been brought through the first steps as a shared instrument in selected research communities. Some argued that peer review is a collaborative tool in academic projects, be it that this is currently facing some fundamental challenges in sustainability.

In this report, discussions regarding the **third and fourth** 'guiding questions' have been brought together into one section. It was stated that ATLAS.TI is suited for transcription, annotation, and editing of film and sound recordings taken from Europeana. A possibility that carries wider

## Europeana Cloud

potential focused on the possibility to having Europeana present more clearly the options for creating queries in its API suite, and allowing the use of for instance Google Refine to import it into a spreadsheet. The question was raised if Europeana contains anything that currently can integrate with the API to help harvest the data for mapping, visualization, analysis, etc.

In general, however, the common opinion was that Europeana currently can only serve as a generic discovery service for developing an impression of “what is out there.” In its present state, Europeana findings would only serve as input for academic research and serve as research resources once they were exported from Europeana into work spaces of individual or groups of researchers, followed by data analyses and manipulation by means of tools or toolkits that researchers are already familiar with. One option that was discussed to help in overcoming this restriction and making Europeana more attractive to researchers in the (digital) humanities is to develop a 'plug-in-ability' for a range of more discipline specific tools.

Based on the observed consensus, mentioned in the previous paragraph, much of the ensuing discussion turned to suggestions for improved ‘tooling’ of Europeana; the **fifth** ‘guiding question.’ Topics dealt with the central themes of tooling for searchability, interactivity, visualizations, metadata quality assurance, and exploring possibilities for professional engagement and crowdsourcing.

In general, the search and presentation functionalities of Europeana were found to be insufficiently conducive to perform professional research. Concerns were raised about the current search options in the portal. These can be summarized in statements that the interface was felt to be “opaque,” that completely identical searches yielded “differing results, depending on when the search term was entered,” and the recurring remark that the interface “lacks possibilities for browsing [Europeana’s] holdings.” A recurring suggestion was that Europeana should offer basic visualization tools for its search results, such as or comparable to Wordle, AquaBrowser, etc.

Similar visualizations were deemed essential for mapping search results in terms of coverage of content (both internally in Europeana and set off against collections that are not yet covered), metadata ratings (what is the metadata quality, measured against a prescribed model), and results that were enriched with annotations or other added value by researchers and other users. It was also suggested as a side effect that mapped overviews of coverage and metadata quality could be beneficial to collection holders, in that they could use these ratings to boost applications for funding for digitizing and describing portions of materials in their care. If metadata quality for a repository or collection were also to be reported (‘seal of approval’/star system/triple key ranking?), this might stimulate repositories to step up their efforts in that direction.

A tool for assessing quality consistency of metadata, and possibilities for users or contributing collection holders to enhance such quality, together go to the issue of ‘digital criticism’ – and if and to what degree Europeana intends to facilitate that core activity within the digital humanities. Such a function would benefit from a system to log users’/writers’ actions on metadata (and, at a later stage, content as well), tagclouds with hits of the day, a tag log generator, logged searches and paths (how did the user end up there, discovery path) and other instruments to facilitate ‘two-way enrichment’ of (meta)data. In turn, this would strengthen Europeana’s reputation for facilitating digital capacity building for various user groups.



## Europeana Cloud

Other tools to be developed in, or for Europeana to strengthen its position in the fields of (digital) humanities research included

- the possibility of OCR for textual materials,
- a mapping service of OCRed materials and the accuracy of the OCRed content,
- a similar service to find and assess images and their resolution,
- tools for downloading large amounts of textual data (tagged, in comma separated value or in plain format), and
- storage solutions for researchers that could be shared with others.

Again, it was felt that any such service or tool ought to facilitate sufficient interactivity, allowing for sharing and improving item/collection descriptions, resource quality, connections with other resources, and storing and showing the enrichment and manipulation history of the item(s). Perhaps this could be developed as a 'Personal Work Space' idea within Europeana, although it was remarked that ample consideration should be placed on making it compatible with the requirements of the academic workplace.

Various participants ventilated the suggestion that Europeana should develop a more sustained role as catalogue. For an example in musicology, see further down this section.

Some comments focused on the need for different ways of exporting and importing data. A tool for harvesting data would impact on need for information on quality of data. It was also pointed out that an important toolkit for these and other purposes, especially geared to combining the key variables 'spatial' and 'temporal' is already being developed within Europeana's own ecosystem: Europeana4D, presented within DARIAH-DE, is a tool for mapping timelines of (combinations of) content available.<sup>8</sup> Its developers implemented their design in a prototype application in the context of the project EuropeanaConnect.

Additional tools to better equip Europeana as a research portal included calls for more adequate and agile translation tools, also because currently identical search actions performed in different languages yielded dissimilar results. Ideally speaking, the platform might benefit from tools for recognizing and mining manuscript materials. Persistent Identifiers (PIDs) for historical personal and place names and even for entire resource sets were deemed essential for authentication and differentiation purposes. Here, Europeana could potentially hook into existing projects. As an example, it was pointed out that the Huygens Institute in the Netherlands is developing a PID-system for 17th-century scientists.

Some discussion developed on the question as to whether Google should fully index Europeana. This ties into questions on positioning and trust: why would anyone turn to Europeana if Google were already in place? The first thing researchers need is access to the documents. Musicologists for instance want a specific entry point for music scores, i.e. better than Google. Whereas OMR can compare similar editions with different layout, an authoritative infrastructure for cataloguing the music is essential in answering basic questions such as "Where to find the sources?"<sup>9</sup> RISM plays a role here, building inventories to know where the sources are. Hence, developing a partnership with RISM seems a viable option.<sup>10</sup>

---

<sup>8</sup> <http://wp1187670.server-he.de/e4d/> (accessed December 1, 2013).

<sup>9</sup> Compare the IMLSP, International Music Score Library Project, <http://imslp.org> (accessed, December 1, 2013).

<sup>10</sup> RISM, Répertoire International des Sources Musicales - International Inventory of Musical Sources, <http://www.rism.info> (accessed, December 1, 2013).

## Europeana Cloud

A final main subject was discussed, one that is difficult to place under either of the headings 'Tools' (Day 1) or 'Content' (Day 2) because it may pertain to both subject matters. Earlier in this report, attention was drawn to the perceived need for intensive interaction with users – and the tools to facilitate such traffic. Engaging specialists in crowdsourcing for corrections in combination with indiscriminate, open crowdsourcing is still a proposition that fits somewhat uneasily – both in research communities, collection holders and information portals. The issue of moderation was discussed in all groups, but there is a growing awareness of the self-guiding potential of the informed volunteer. Besides that, academics participate in both types of crowdsourcing (see for instance the Perseus Library initiative vs. Pleiades,<sup>11</sup> both on ancient history), and are looking for ways to be credited for that work. Europeana might well find considerable rewards by looking into existing activity groups on a particular subject – be they academic or not. In this regard, a recent report by Stuart Dunn and Mark Hedges was flagged as particularly helpful in conceptualizing and organizing crowdsourcing the humanities; it is available in a longer and an abridged version.<sup>12</sup> In addition, the idea was discussed that Europeana organize various crowdsourcing events, to develop expertise in the matter. Finally, Europeana might consider setting up a young scientists competition.

In short, Europeana was called on to start experimenting and gain understandings on how to *make digital impact visible, and give credit* (academic and otherwise).

### 11.1.1 Summing up the Tools session discussions

In summing up the first day of the Expert Forum: several main points came through in the group discussions and their presentations. The researchers consulted in the forum emphasized the need for improved and uniform metadata mapping, with strong multilingual capacities. They advised that the Europeana portal strongly expands its level of interactivity with researchers and other user groups, facilitating exploration consisting of considerable higher levels of granularity and in multiple dimensions.

Key words in the exchanges were: enrichment, engagement, visualizations (spatial/temporal, and other means of mapping data, content and quality aspects), collaboration, and we can also add the general observation that Europeana was called on to “move from searching to browsing, presentation and to increased interaction with users and (their) findings,” and finally that it ought to develop means to connect more closely with specialized portals, collections, and academic projects in the digital humanities.<sup>13</sup> Some participants expressed the opinion that Europeana's

---

<sup>11</sup> <http://www.perseus.tufts.edu>, <http://pleiades.stoa.org> (both accessed December 3, 2013).

<sup>12</sup> Written for the UK's AHRC Crowd Sourcing Project Full, the full version (56 pp.) is 'Crowd-Sourcing Scoping Study: Engaging the Crowd with Humanities Research,' Stuart Dunn and Mark Hedges, [n.d.; {2012}], <http://crowds.cerch.kcl.ac.uk/wp-content/uploads/2012/12/Crowdsourcing-connected-communities.pdf> (accessed November 25, 2013); a shorter version (12 pp.) is 'Connecting Communities: Crowd-Sourcing in the Humanities. A Scoping Study,' Stuart Dunn and Mark Hedges, [n.d., {2012}], <http://www.ahrc.ac.uk/Funding-Opportunities/Research-funding/Connected-Communities/Scoping-studies-and-reviews/Documents/Crowd%20Sourcing%20in%20the%20Humanities.pdf> (accessed November 25, 2013).

<sup>13</sup> An integral part of the first main Deliverable in this WP, D1.1 Research Communities Identification and Definition Report, is an inventory of research communities and practices in the humanities, 'D.1.1\_Communities\_Table\_Humanities.'

## Europeana Cloud

ultimate capacity for tapping into existing (academic) communities and contributing to the creation of new knowledge would prove to be its main reason for survival as a European infrastructure.

### 11.2 Content

Day 2 of the forum focused on content; what content in Europeana is useful as it stands, what gaps exist and, most importantly, what changes in addition to Europeana content could encourage future humanities research? (This section also includes comments regarding Europeana's content made during session 1.)

On the **first** 'guiding question,' all participants agreed that Europeana is a great instrument for showing the diversity of what might be coined 'European Culture.' Already, some experts remarked, it lends itself well for use in introductory teaching activities; a quality that could be much improved on with better-quality metadata. Still, even in its basic function as an exploration tool for European heritage, most agreed that Europeana needs to improve its coverage on most subject matters. As it stands now, many would subscribe to the remark of one participant who stated that "Europeana currently is not thought of as a research tool."

This understanding informed much of the discussion surrounding the **second** and **third** questions. In general, participants would like to see many more textual collections added; humanities scholars are mainly interested in digital texts – so these rank among the most important materials to have in Europeana. There is no shortage such collections. Currently, Europeana is seen as one of the available repositories, and not necessarily a trusted one. If Europeana's basic goal is described as creating a repository, add more content and enrich metadata it was felt that this objective is not concisely communicated anywhere on the website, and many commented that they were not sure that any guidelines for achieving this mission were set. It was strongly recommended that the project actively engages various groups of specialists in the humanities and articulates and publishes short-, mid- and long-term goals for the fulfillment of its mission to develop into a significant research tool for the digital humanities.

Several expert groups already explored case studies in their discussion on what would be needed in this regard. Musicologists observed that some music prints from British, German and French (national) collections were already in Europeana, but as thumbnails only – and with inadequate levels of metadata (certainly when considering the needs of a researcher). The suggestion was discussed that eCloud organizes, in tandem with a group of specialists, transcribing projects of digitized music prints and scores. eCloud could host this type of content in an aggregating environment on a temporary basis (during the lifespan of the project). Tools would then be built on top of the content to allow a user community to make automated transcriptions, manual transcriptions and corrections. At the end of the project Europeana might remove the images, but retain the enriched metadata (that could also be shared with the original source libraries). If Europeana developed projects like this, it was felt that over time the quantity and quality of metadata would gain substantially more robust levels of trust within research communities.

In addition, it would be interesting to connect such data to other material from other aspects of musicology<sup>14</sup> or even other disciplines, i.e. to datasets of performance history, or connect it to

---

<sup>14</sup> Links to and cooperation with for instance RISM and IMLSP were among the suggestions here, see notes 8 and 9.

## Europeana Cloud

church history. Europeana might be able to enhance its metadata by making relations that *span across disciplines* in ways that other research projects are not approaching. In turn, this allows for different conceptualizations of the material.

One subgroup comprising specialists on Biblical studies found that while searches in Europeana for New Testament manuscripts and Bible manuscripts did yield numerous results, they also concluded that these findings were problematic to understand and handle. One search came up with 112 images and 175 texts, a difference that is explained by the fact that numerous links went to a record in the European Library, not directly to an image. Next, the user finds that a new search has to be conducted within the other institution's catalogue search function. In addition, nine volumes of the same work (*Nouveaux Fonds*) showed up as three different items. Examples like this underscore the need for better and more consistent metadata mapping. But it was also felt that it clearly illustrates the need for Europeana to consult with scholars, specialists and dedicated institutions in order to devise concepts and organizational solutions in searching for both overarching qualities and fine granularity, required for catering to the needs of humanists studying –in this case- biblical sources (both printed and manuscript).<sup>15</sup>

Similar projects could be set up for a range of communities and subjects, for instance all European repositories with stewardship of collections of pottery, or a project on the 19-century novel across Europe. An interesting challenge was presented, where Europeana might apply and develop its potential for contributing to 'the European mission.' Numerous broadcasting agencies are in the process of building national repositories of digitized (or digital-born) content from their aired programming. If Europeana would get involved and organize and present all these national initiatives and collections as linked data, cross-national and across languages, in a manner that overarches such national, domain-specific infrastructures (for instance developing overarching, multi-lingual ontologies and semantic web services), this might well constitute the perfect showcase for presenting Europeana as a unifying agency, free from national, institutional or even disciplinary concerns. In some countries, national digital repositories have already joined such projects, aiming to develop as one of these put it "an innovative *cross-archival semantic content discovery platform*."<sup>16</sup>

During the discussions and the group presentations at the end of the session, it was difficult at times to clearly distinguish specific contributions on the **fourth** 'guiding question.' Earlier paragraphs in this report already contain suggestions and explorations of gaps in Europeana from a digital humanities perspective. The fundamental and basic concern of the experts was that, to start with, at present it is not feasible to establish the portal's coverage of (meta)data on any given subject. This has already been identified in the preceding section on tools.

Researchers engaging in oral history remarked that Europeana contains very little materials for their studies. They call for a collection programme of oral history resources, that need not be limited to audio / visual resources; transcriptions are suitable as well. Oral history was described as

---

<sup>15</sup> Suggestions for institutional connections and collaboration included the INTF (Institut für Neutestamentliche Textforschung) in Muenster, the New Testament Virtual Manuscript Room, and the Walt Wittman Archive Room (USA).

<sup>16</sup> In the DRI-INSIGHT RTÉ project, the Digital Repository of Ireland is taking part in such a project with the Irish national broadcasting organization, see <http://dri.ie/dri-insight-rte-project> (accessed December 4, 2013).

## Europeana Cloud

booming all over Europe, and it was remarked that European funds are available for the creation and collection of interview transcriptions. Various groups called for increased coverage of maps, with the added proviso that at the minimum they ought to be geo-referenced and preferably in a manner that allows for their usage in GIS-applications (also note the open source in the previous main section, QGIS). Some confirmed that 3D-representations and models are virtually absent from Europeana. While up to a few years ago that would have been prohibitively costly, it was argued that nowadays one could easily create even mobile 3D-applications. In this regard, the Europeana4D-project was mentioned again as it combines some functions that Europeana as whole could strive for: movement, multidimensionality, projections (space/time), virtual exhibitions and virtual narratives.

A different approach for Europeana to expand its content is to tap into existing interest groups and allow for them to upload resources, combined with descriptions. A recent project by the Digital Repository of Ireland was successful in creating a mobile app on ephemera and photographs that accommodated for this function. Whether such a crowdsourcing project should also come with some moderating agency and how that might be organized is of course a different matter (but for 'best practices' in the humanities, see the recent report by Dunn and Hedges, mentioned earlier).

Also a number of significant additional comments and questions came up that are not easily classified by answering this session's four question. They merit listing in this report nonetheless:

- Currently many libraries request a sign-in access. Does Europeana intend to become a lobby group to campaign for open access for all content or data? For these purposes linking up with Centernet, an international network of digital humanities centres, might be especially beneficial (<http://digitalhumanities.org/centernet/>).
- Will Europeana provide access to scholarly/scientific journal articles? Many of the main bibliographic databases operate on a subscription base, but it would carry many advantages if a researcher/user of Europeana could get a direct link to entries (even if the content itself remains closed).
- Present circumstances and funding provide Europeana with a unique opportunity to step up efforts to move beyond current offerings of more or less isolated, rigid silos of information resources offered through a prescribed model of understanding, to a domain in which serendipity is allowed considerably more space and where it is accorded more intellectual acceptance. It is precisely in unexpected connections between nuggets of information that new forms of significance or understandings can be explored. An example was given of an important repository offering unrequested, unsolicited and non-prescribed associations between freely linked data and items that might potentially restructure researchers' approaches to queries and our understanding of their results.<sup>17</sup>

### 11.2.1 Summing up the Contents session discussions

Many of the exchanges in this session reflected observations mentioned also elsewhere in previous reports from this Work Package of Europeana. In the first Deliverable of the package, it was remarked that: "The projected *growth of Europeana's content* as presented in its DoW, deriving from both existing and new aggregators, will significantly increase eCloud's offerings to various research communities." But it also added the important proviso: "For these additions to comply

---

<sup>17</sup> DHO:discovery, <http://discovery.dho.ie>. (accessed December 4, 2013; operations ceased). The project's description of "serendipitous discovery of related knowledge" is on <http://discovery.dho.ie/discover.php>.

## Europeana Cloud

with the needs and requirements of various research communities, one of the *key challenges* for Europeana will be to develop enhanced calibration of the metadata of individual items and entire research collections with *relevant resource descriptors* and *identifications of possible deployment in humanities and social science research*.” The Deliverable concluded with stating that: “Europeana Cloud should make concerted efforts to *reach out and engage* with projects on both a larger and a smaller scale, where scholars and scientists are actively developing and reshaping their e-research practices.”<sup>18</sup>

## 12. **Conclusions from Expert Forum 2**

Europeana has identified large research collections from a wide range of content aggregators in Europe for inclusion in its portal. For Europeana Cloud to become an important research portal for researchers in the humanities, participants of this forum identified a number of critical elements that deserve attention from Europeana in its prospective uptake of these materials and in its further development as a repository.

At a basic level, the convened experts emphasize that improved metadata quality and consistency is essential to attain the earlier mentioned objective. In addition, Europeana should develop fundamental interfaces for *mapping* and *visualizing* the distribution of its holdings – and the characteristics of the results from queries: coverage in Europeana, additions from other collections; fullness and relevance of metadata; provenance information; deep links available or not; annotations available or not; various forms of *contextualization*, etc.

Europeana is called on to enhance the presentation of the key variables within the project: what is considered ‘European’, what will be offered in Europeana, who organizes this, what are the project’s next steps, what exactly can we find here (in main groupings). In a similar vein, many of the forum participants find that Europeana’s landing page is in need of reconfiguration. It should present the subjects listed in the previous remarks, but also be an attractive entrance point for various user groups (including for instance API-developers), and offer a basic presentations on the various ways that Europeana can be approached and queried.<sup>19</sup>

Within the service itself, participants called on Europeana to greatly enhance its capacities and functionalities for interactions between the service and its users, and for exploring of and reporting on interconnectivity between its resources.

Key terms in the discussions include:

- adding and logging user comments, accommodating user-enhanced metadata;
- boosting import and export possibilities;
- enhancing search functionality and filtering functionality;
- moving from exploration and discovery to in-depth descriptions and interconnectedness;
- stepping up development for interaction and connection with users / user groups.

Similar to the recommendations in the closing paragraph of the preceding section 4.2.1, another Deliverable of the Europeana Cloud project concluded: “The project needs to think about how it can

---

<sup>18</sup> D1.1 Research Communities Identification and Definition Report, 14 and 13, emphasis added. In the first citation, a reference was made to the project’s DoW; section B.2.1b. “Underlying content”, 68-92.

<sup>19</sup> For a ‘shining example,’ see this US government website, <http://www.data.gov> (accessed December 5, 2013).

## Europeana Cloud

*“tap in”* to existing [research] communities.”<sup>20</sup> Implementing the recommendations from this forum for all domains within the humanities (or, for that matter, the social sciences) seems a daunting task. In this report some communities have been identified where experiments can fruitfully be developed (oral history, biblical studies and musicology). This conclusion differs somewhat from that of the social scientists forum that called on Europeana itself to decide on what fields of research should receive priority.<sup>21</sup> Still, it is considered essential for Europeana to connect more thoroughly with existing digital ecosystems in the humanities. Participation in digital humanities projects and communities is a prerequisite for Europeana to develop its conceptual identity as ‘metadata brain’ for European culture and research.

The concluding remarks in this report on Europeana and researchers in the humanities show a great deal of overlap with the findings of the preceding forum of social scientists in Gothenburg.<sup>22</sup> The unanimity between the two consulted communities adds considerable urgency to their incorporation into the evidence base that will be reported back to the project that requested these consultations. For Europeana to become a trusted repository of (meta)data for these scientific ecosystems it seems a requirement that Europeana Cloud increases its engagement with organized communities in the humanities and social sciences.

---

<sup>20</sup> D1.2 State of the Art Report on Digital Research Practices, Tools and Scholarly Content Use, 51, emphasis added.

<sup>21</sup> D1.5 (3 of 4) Expert Forum Tools & Content for Social Sciences Research Report, 11.

<sup>22</sup> *Ibid.*, 15.

## **Expert Forum 3 - Tools and Content for Social Science Research, University of Gothenburg, October 2013**

### **13. Introduction to the task**

This report focuses on the outcomes of Expert Forum 3 held in Gothenburg, October 24-25 (project month 9). It is the third of four Expert Forums within Work Package 1 of the EU project “Europeana Cloud: Unlocking Europe’s Research via The Cloud”.

#### **13.1 Expert Forum 3 in the context of Work Package 1**

The general purpose of Work Package 1 is to assess the researchers’ needs for future work with the Europeana content. To do this, the work package encompasses a number of tasks and subtasks employing different investigation methods such as desk research, web surveys, and case studies. In addition, four Expert Forums will be held in the course of Work Package 1, each providing input from external experts on different topics.

Expert Forum 1 took place in Dublin, June 18 (project month 5), earlier this year. The Dublin forum focused on the typical needs for researchers in the humanities and social sciences to be able to work with the Europeana content, illustrated in the three created cases.<sup>23</sup>

Expert Forums 2 and 3 (held in Amsterdam, project month 10, and Gothenburg, project month 9) are to be seen as complementary in that they have much the same focus and agenda, although each explored more thoroughly the social sciences and the humanities, respectively.

Finally, the fourth Expert Forum is scheduled for July in 2015 (project month 30). This forum will provide a broad review of the tools and content access and use services provided by Europeana Cloud, and will develop recommendations for future work, including how the engagement of researchers will continue beyond the lifetime of the project to ensure their future use and uptake of the Europeana Research platform.

#### **13.2 The purpose of the forum**

The Expert Forum aimed to generate ideas for how Europeana can be developed into a useful resource for social science research. The discussions focused on what digital tools would benefit research on the Europeana content (metadata, text, images, sound, video, 3D) and how Europeana can be improved as a source of research material.

#### **13.3 How the results will be used**

This report forms part of Deliverable 1.5 together with the three accompanying Expert Forum reports.

---

<sup>23</sup> See Deliverable 1.5 (1 of 4): *Expert Forum Case Studies Report*.



## **14. *Selecting the research areas and the participants***

### **14.1 *Research areas***

Given the wide scope of the social sciences, it was necessary to narrow the concept by selecting a few disciplines of the social sciences. During the work with Deliverable 1.1, six subject domains were identified as most likely to find Europeana material useful in their research.<sup>24</sup> These subject domains were taken as a point of departure for the understanding of the social sciences field in the context of the Expert Forum as well. The six subject domains thus came to inform the selection process, but other variables were also considered when selecting expert participants. One such variable was the range of research areas a particular participant could address; it was felt that a researcher with experience from several social science disciplines (especially through interdisciplinary research) would be in a position better to contribute to the discussions. Another variable was familiarity with digital methods; researchers who did not employ digital tools in their work would probably contribute less to the discussion than researchers well versed in digital research. Finally, it was also felt that researchers who concerned themselves with development of methods as well as with research on digital research methods (for instance within Library and Information Science [LIS]) would be able to provide a broader perspective on the topics of the forum.

Taking these variables into account, the final factors that came to decide part of the selection of experts was the difficulty in identifying “ideal” participants and, once they were identified, their possibility to attend. While not covering all potentially relevant disciplines, the mix of participants that did take part was nevertheless very successful in meeting most of our criteria. The main social science research areas represented by the researchers were: cultural heritage studies, economic history, education and learning research, e-social science, ethnology, gender studies, infometrics, LIS, political science, science and technology studies, sociology, and text mining.

### **14.2 *Selecting the experts***

Potential external experts were identified by desk research and through recommendations from colleagues, both from within and outside the project. Once recognised as potentially relevant for the forum, an invitation e-mail was sent out, briefly describing the Europeana Cloud project and the general aims of the forum. Those who responded that they would like to take part were then provided with detailed information. It should be emphasised that many more than the actual participants reacted positively to our invitations by showing great interest in the forum and the project in general, but were unable to attend due to other duties. In a great number of cases, these people suggested other suitable experts.

### **14.3 *Participants***

The participants of the Expert Forum were either people from Europeana Cloud-related institutions or external experts from institutions not directly involved with the project.<sup>25</sup> In addition, the Research Community Advisory Board of Europeana Cloud was represented by Leif Isaksen. It is worth noting that the participants belonged to institutions from several different EU countries,

---

<sup>24</sup> These were, in alphabetical order, Economic & Social History, Gender Studies, Human/Economic/Political/Cultural Geography, Political Science, Social Anthropology, and Sociology.

<sup>25</sup> See Appendix VII – List of participants Expert Forum 3 – Gothenburg for a complete list of participants.

providing a broad European perspective. More specifically, institutions in the following countries were represented in the forum: Greece, Finland, Ireland, Latvia, the Netherlands, Norway, Sweden, and the UK.

## **15. *Designing the forum***

In line with recommendations from the report of the Dublin Expert Forum, the forum took place during two half-day sessions beginning in the afternoon of October 24 and ending at midday the day after.<sup>26</sup> The forum started with a short introduction of Europeana and eCloud by the leader of Work Package 1, Agiati Benardou. After that there was an ice-breaker activity that was also meant to ensure that all participants had recent experience with the Europeana portal, followed by two main sessions focusing on tools (day 1) and content (day 2), respectively.

### **15.1 *Europeana Treasure Hunt***

Following the successful introduction feature at the first Expert Forum, the forum started with a “Europeana Treasure Hunt”, to break the ice as well as draw attention to some of the aspects of the Europeana portal and content. It used the same structure as in Dublin but the tasks were substituted better to fit the agenda of the Gothenburg Expert Forum.

The participants were divided into groups named after the Europeana content types; image, text, sound, and video. Each group consisted of four participants composed so that they all had a mixture of people from within the eCloud-project and external experts. The first assignment for each team was to create a Europeana profile. Those who had never used Europeana before were thus allowed to familiarise themselves with the user profile function which enables the user to save all previous searches. Having created a Europeana profile, they were asked to work for 15 minutes with three assigned tasks.<sup>27</sup>

The purpose of the first task was to familiarise the participants with the different content types of Europeana by asking them to formulate searches that gave results which contained all five content types. The second task drew their attention to the metadata structure currently used in Europeana. Participants were asked to find as many metadata fields as possible, allowing them to get a rough overview of existing metadata fields. Finally, in the third task the goal was to find the lowest possible number of search results. The purpose of this task was to give the participants the opportunity to work closely with the search tool and gain insight into its current functionality.

### **15.2 *Session 1 – Tools***

Session 1 considered the kinds of tools that would be useful, and could potentially be developed, for social scientists carrying out research on the Europeana material. The central part of both sessions was group discussions followed by a summing-up session with all participants. For the breakout session, the participants were divided into new constellations, decided in advance by the forum organisers. It was deemed important that each group consisted of people both from within and outside the project, and that there was a good mixture of people from different academic disciplines.

---

<sup>26</sup> See Appendix VIII – Agenda of the Expert Forum - Gothenburg for the entire agenda.

<sup>27</sup> The complete instructions given to the participants can be found in Appendix IX – The Europeana Treasure Hunt - Gothenburg.

## Europeana Cloud

A short introduction was given before Session 1, in which the participants also had the opportunity to ask any questions they felt needed any clarifications. To provide a structure to the breakout session, participants were encouraged to discuss the following:

- What tools do you use for your *own* scholarly activity?
- What tools do you use when you *collaborate* with fellow researchers?
- How would you use Europeana content in your research?
- What tools do you think you could use *now* with Europeana content?
- What tools need to be *developed* to enable you to use Europeana content in your research?

The comments arising from the third question came to anticipate some of the contents discussion of session 2. The points made in this context will therefore be accounted for together along with the discussion of Europeana content under 4.2, below.

### **15.3 Session 2 – Content**

Session 2 dealt with the content of Europeana and how it should be developed, in part building on the results from Session 1. The groups from Session 1 the previous day were kept for Session 2. Again, there was a short introduction giving an overview of what content there is in Europeana at the moment and what is planned for Europeana Cloud, and then the following main questions were given to each group:

- Is there content in Europeana that is useful currently? What content?
- Does this content need to be improved or added to?
- What new content (genres, formats) would you like to see added to Europeana?
- What are the biggest gaps in content in Europeana from a social science perspective?

## 16. Results

The following section presents the results from the discussions on Tools (Session 1) and Content (Session 2). Findings of the four groups as well as comments made during the summing-up sessions afterwards are combined in these results.

### 16.1 Tools

The first session focused on tools; what tools do the expert use for their own scholarly activity, and what tools could they use, now and in the future, with the Europeana content?

#### 16.1.1 Current use of tools

Under this heading are summarized the points that came up in discussion of the first two questions (*What tools do you use for your own scholarly activity?* and *What tools do you use when you collaborate with fellow researchers?*).

The digital tools currently in use by social scientists seems to run the gamut from common, proprietary, off-the-shelf programs to specialized solutions, open-source software, and the use of (unintended) affordances of online tools and services. The tools brought up clustered around some central analysis methods.

One such cluster of methods concerns the analysis of quantitative data. Apart from programs that offered ways to structure data and investigate them statistically, such as SPSS, Microsoft Excel, and Filemaker, a number of programs for (social) network analysis and visualization (NodeXL, NetVis, UCINET, Gephi) were mentioned. There was also a mention of Wordle, an online service for the visualization of word frequencies (so-called “word clouds”). In this cluster could also be included the various unnamed tools that researchers used to manipulate data in various ways: combine data, parse metadata into relational databases, and clean data records by identifying when the same person appeared in more than one record.

Another cluster includes software that allowed qualitative analysis, largely of video and audio material but also of text and images. HyperResearch, Nvivo, Atlas TI, Transana, ELAN, and Final Cut were all mentioned as employed for transcription, annotation, and editing of film and sound recordings. Some of these also offered analytical support, or provided the research team with collaborative capability. Even the video editing tools from online services such as Instagram and Flickr were brought up as useful research tools.

A third cluster is made up of methods of automated data capture (“data scraping”), mainly from the Internet. There seemed to be a trend of combining tools to facilitate data management in this regard. It was explained how Yahoo Pipes could be set up with Google Docs to allow for capture of an XML or JSON feed without downtime; how NetVis could capture Facebook data; and how there were tools that allowed of the capture of mobile-phone screens.

The fourth cluster is comprised of programs and services that are not directly related to data capture, structuring, or analysis. Instead, they facilitate or enable other aspects of the research process, such as information seeking, storage and data access, reference management (mainly EndNote and Mendeley), even e-learning. Programs mentioned range from Google’s search engine and the Google Drive tools to JSTOR, ResearchGate, and Seeknet. These tools and services, while

## Europeana Cloud

clearly relevant to the researchers, demonstrate how all aspects of the research process today have been digitized.

The collaborative functionality was stressed for several of the tools in clusters two and four in particular, and several others were mentioned as well. Cloud-based services featured prominently in the discussion, including tools and services like the Google Drive tools, Dropbox, and Mendeley. One researcher brought up Basic Support for Cooperative Work (BSCW), an online cooperation platform for project work. The three aspects of collaboration that were stressed covered identifying/getting in contact with colleagues sharing one's research interests; communicating, often internationally, between project partners; and working, often simultaneously, on the same files.

### 16.1.2 Tools for working with Europeana now

This section summarizes the discussions on the fourth point (*What tools do you think you could use now with Europeana content?*).

There was a general feeling that as Europeana stands today, there are not many tools that researchers would use. Partly, this is a content issue: the only content actually in Europeana today is metadata, and it was generally felt that the quality was too uneven to provide good material for much research (similar points were made at Expert Forum 1 in Dublin); general uncertainty about the context and provenance of the records (see more under 4.2.2, below) was also indicated as a point against using Europeana material for a scholarly analysis.

As for tools possible to use with current Europeana content, the Europeana API came up in several groups as a convenient way of downloading records, and the benefits of aggregation was pointed out as a strong point in Europeana's favour. A feed in XML format was seen as potentially very useful. Other possible tools that came up for use with current content included Google's search engine and Pinterest, to allow researchers to create theme-based collections of images.

### 16.1.3 Tools for working with Europeana in the future

This section gives a summary of the group discussions on the fifth point (*What tools need to be developed to enable you to use Europeana content in your research?*)

The comments in the four groups covered a wide range of possible tools and research strategies, and many of the issues raised necessarily concerned both tools and the content required for those tools to be relevant. For the sake of clarity, this report has kept tools and contents discussions separate, regardless of whether issues were in session one or session two.

The tools suggested for future work do not allow themselves to be easily categorized, and tools for information seeking, (crowdsourced) metadata enrichment and contextualization of Europeana records have been discussed in section 4.2.2. That the discussions returned to the wish for particular tools for evaluating and visualizing the coverage of Europeana in relation to the collections of the institutions that provide content, and for examining the completeness and quality of Europeana content reflect the concern about the overall transparency of Europeana as a source of research material. This issue is accounted for in more detail in section 4.2.3.

Visualization was one of the general areas in which several suggestions were made. Tools that were mentioned rendered visible geospatial information and relationships; word frequencies and results

## Europeana Cloud

of text mining; and networks, systems and hierarchies. For the latter, a possible plug-in for Gephi was proposed. It was felt that there was little need to replicate the functions of Gephi in a Europeana Cloud platform; better to provide an easy way to connect Europeana with the third-party software.

The relationship between a future Europeana and external, third-party research solutions was another area of discussion. A plug-in was also suggested for the qualitative text-analysis program NVivo. Another suggestion that came up was to allow for text and data mining and text analysis by including the capacity in Europeana to reformat material, for instance by converting PDFs to CSV (comma-separated values) or just plain text.

It was widely agreed that the Europeana API could be made easier to work with, for instance by offering a graphical user interface (GUI); as it is, it was pointed out, the API requires a level of hacking ability that works as a barrier to entry for many users. It was also observed that there was a lack of activity at the Europeana API Google Group, which was taken to imply that the API was not widely used and that this gave the impression of lack of support and help to work with the API. Although no specific third-party software was mentioned for use with the API, there was an impression that the ideal would be an easy-to-use API that accepted a wide range of request types and could output data in standard formats that could be rendered by most programs (CSV and plain text formats were mentioned) or metadata in XML or JSON. No one was willing to predict future requirements, but researchers who built their own tools stressed the need for broad functionality, to allow for the development of as wide a variety of future tools as possible.

A third area concerned possible tools for annotating Europeana material. Partly, it was felt that there should be the possibility to annotate material privately and to save tags in your My Europeana account, to facilitate the search for and triage of material. The possibility to share such annotations was also brought up, along with the possibility to filter out annotations in searches. Partly, there was also a suggestion that it should be possible to make annotations openly available. These discussions shaded into the crowdsourcing of metadata and of metadata enrichment; for more comments on that, see 4.2.2.

One particular area in which annotations were seen as desirable concerned tools that managed images. Functionality suggested included the ability to organize images in various collections, for instance through annotations, and the possibility to export references to a collection of images. As one researcher put it, it would be good if there was a way to add “human intelligence” to image interpretation. While face recognition is a fairly well-established technology nowadays, for most other image types and analyses, there is a need for various kinds of user-provided tags and annotations, which can then be analysed.

Some general comments, based on the observation that the current user interface of Europeana was opaque in some of its functionality, stressed that any new additions in terms of digital tools to Europeana must be time-saving and easy to use or researchers will simply stay with the tools (and sources of material) that they already know. It was also proposed that tools should come in the form of widgets for web browsers.

### 16.1.4 Summing up the Tools session discussions

Research today has gone digital, and it hardly comes as a surprise that digital tools and services are brought in at all stages of the research process. Tools that facilitate traditional forms of analysis or

## Europeana Cloud

enables new ways of approaching material are ubiquitous. It is therefore interesting to see the discrepancy between what tools researchers use, and what tools they would like to see implemented in a future Europeana Research platform or added to the Europeana portal.

Apart from calling for ways in which to evaluate the context and provenance of material, the main concerns appear to be possibilities efficiently to search for, collect, and sort material; and possibilities to export data once the material has been found and prepared. In the cases where analytical functionality was suggested, it stayed with fairly low levels of analysis: visualisation tools that could show interesting patterns that would then have to be investigated with more powerful third-party programs.

Time is expensive. A visit to Europeana must mean time saved; and since these researchers already have found an arsenal of tools with which they are familiar, trying to come to terms with new tools, in particular if these tools are not compatible with the familiar ones, using sophisticated analytical software within Europeana would not be time-efficient.

### **16.2 Content**

The second session focused on content; what content in Europeana is useful as it stands, what gaps exist and, most importantly, what changes in addition to Europeana content could encourage future social science research? (This section also includes comments regarding Europeana's content made during session 1.)

#### **16.2.1 Useful Europeana content now**

As was observed at Expert Forum 1 (Task 1.4.1), the majority of the Europeana items lack sufficient metadata to make it a truly useful research tool.<sup>28</sup> That point was made again in the discussions at this Expert Forum, and the most common suggestion was to use Europeana as a tool to provide inspiration or a starting point for research rather than a way to find actual research material. One possible use, which came up also at the previous Expert Forum, was the possibility to use Europeana to identify institutions with large collections of specific content. It was pointed out, however, that this required better context and transparency about the content selection. Suggestions for improvements in this regard will be detailed under 4.2.3, below.

Other possible ways to use current Europeana material for research mentioned the possibility to map phenomena and visualize content (e.g. places of production), and to carry out interdisciplinary studies of objects or artefacts and the social links between them. A concrete project that was brought up involved using Europeana to track pictures from schools and examine how classroom technologies have changed over time. Another example regarded the possibility for comparative studies, for example comparing people from different cultures in specific periods of history. Finally, besides using the Europeana content in the research process as such, it was suggested that it could also be utilised as a teaching resource (e.g. use images in presentations).

#### **16.2.2 Europeana as a source of research material**

This subsection will deal with a number of problems with the Europeana content as it stands today, as well as conceivable solutions to these shortcomings that were identified during the discussions.

---

<sup>28</sup> Deliverable 1.5 (1 of 4): 8.

## Europeana Cloud

*Content:* The discussions on future content addition centered around two possible paths. Europeana could either increase its range by adding content on a great many different topics, or apply a more narrow approach whereby the focus is put on increasing quantities of material on a limited number of topics. The question is thus whether Europeana should have a little about everything or a lot on certain selected topics. While the ideal case, of course, would be to have a great deal about everything, the participants seemed to agree that the more fruitful strategy would be to move forward by prioritising a limited number of areas. Rather than wasting too much energy at filling the Europeana with a vast number of new topics, it may be a better to inform the Europeana users about the limits and instead aim at the having the highest possible quality of the existing content.

The experts gave examples of some topics that were considered to be represented by too little material in Europeana, including material on women's history, the Sami people, and the Romani people. What kind of content that would be the most desirable to ingest into Europeana is ultimately dependent on whom you ask. Researchers from different fields within the social sciences will come up with completely different answers on what content they would like to have added. Europeana will therefore have to decide which fields should receive priority, following the experts' recommendation that it would be better to go for quality over quantity. The vast majority of experts was too little acquainted with Europeana and had not considered it as a source of research material before the Expert Forum, however, and they were unwilling to give any clear indications of which particular areas should be strengthened or focused on. Any expressed desire for particular content was clearly connected to personal research interests rather than to a strategic view of what would benefit their field or discipline at large.

One matter that was brought up and received a fair amount of attention was the inclusion of material that, while available in European institutions, did not actually originate in Europe. Books from the U.S. and objects from Africa were two of the examples brought up. It was felt that searches should be possible to filter by European items – perhaps through the introduction of a “Europe” tag. The question that then arose concerned what should count as Europe over time, a question that did not get resolved.

*Metadata:* In line with the feedback on existing metadata from the Dublin Expert Forum<sup>29</sup>, it was emphasised repeatedly that the standard of the metadata quite obviously needs to be raised. Regardless of how many interesting objects there are in the collection, they will remain useless for researchers as long as they lack sufficient metadata quality.

One problem with the existing metadata is that it varies greatly among different records, reflecting an imperfect ingestion procedure, which allows for records to be ingested without a complete set of metadata. A standardisation of this procedure is therefore called for to force the providers to complement the record(s) with all relevant metadata. At the same time there is the need for improvements of the metadata on the already existing content as well. Crowdsourcing was suggested as an opportunity for a correcting process of the metadata whereby a great many people together could make a valuable contribution to the metadata standard of Europeana.

---

<sup>29</sup> Ibid.



## Europeana Cloud

On a general level, it would be useful to be able to easily distinguish between objects with good and poor metadata quality. As was also outlined in the Dublin report<sup>30</sup>, one possible solution to this would be to implement a metadata star quality system rating the Europeana items based on certain variables such as the completeness of an item's metadata. Not only would such a system help the researchers using Europeana to navigate among items with varying levels of metadata, but it would also be of help in the correcting process by highlighting the objects most in need of metadata improvement.

Some researchers also brought up the issue of time-related or space-related (geospatial) metadata: if it was clear already in the hit list from a search which objects had these kinds of metadata, and possibly allowed for sorting by geographical area (all records pertaining to, say, Scandinavia) or chronologically, it would be very useful. It was stressed that the level of detail needed to be high in order for the researcher to make precise selections. It would also make the process of identifying potentially useful material more efficient, as there would be no need to open hundreds or thousands of irrelevant records in order to find the relevant ones.

Another comment made about metadata concerned the fact that not all metadata tags had information translated into every language. In a multilingual environment such as Europeana, that means that a researcher would have to be a polyglot even to be able to identify potentially relevant material – which in actual fact would probably mean that material would be ignored if the metadata were untranslated.

*Searchability:* Other than the content and its metadata as such, some shortcomings discussed fall into the domain of searchability. The current filtering function allows for search refinements based on media type (image, text, sound, video, and 3D), language of description, year, providing country, copyright, provider, and addition of any keyword(s). While all these constitute useful search refinements, there are several possible improvements and additions that could be implemented in order to develop Europeana into a useful resource for social scientists.

To begin with, given the large amount of objects, especially regarding images and texts, the current search function does not allow for smooth navigation among the objects based on the different content types. All of Europeana's five content types constitute broad concepts encompassing a wide range of different types of images, texts, sounds, videos and 3D objects. Rather than being interested in all kind of images, a researcher is more likely looking for a specific type of image. Choosing the content type image, then, is not much of a filtering process since it brings on more than 17 million search results containing all sorts of images. A recommendation to Europeana is therefore to elaborate subcategories to each content type since they all are kind of heterogeneous concepts. Text, for instance, could be divided into various subcategories such as books, letters, monographs, and so on. Another possible amendment to the current filtering function concerns the refinement by year. Although useful to fill out specific years of interest, it should also be doable to specify a range of years as it would make the refinement process a lot more effective.

In addition to modifications of the current search refinements, it would also be fruitful for Europeana to apply even more filtering options so that the researchers could narrow their searches in a more detailed manner. In fact, some metadata seem already to exist that could be used to

---

<sup>30</sup> Ibid.

## Europeana Cloud

elaborate further refinements. One limitation of the current filtering is, for example, that it is not possible to specify a search by geographic coverage of the objects, even though such a metadata variable exists. The only geographical variable that could be used for search refinements is *providing country*, which presumably is of much less relevance for a researcher compared to geographic coverage of the object. Other helpful additions to the search function would be full text search (of text objects, rather than just the metadata) and semantic indexing.

A final issue of searchability referred to the importance of reliable search terms. That is, a certain search term should give the same results if carried out repeatedly. At least one of the experts managed to get different results (or at least different ranking) for two identical searches and was therefore unable to find a specific image of interest spotted the previous day. Future Europeana should naturally rule out the possibility of similar incidents.

*Contextualisation:* Other than the need for improved searchability, it was emphasized that the content ought to be more contextualised. To a great extent, the Europeana objects seemed to be too much like isolated pieces without sufficient contextual description attached. Regardless of the specific social science discipline or the specific purpose of an investigation, some kind of contextualisation of the data is always an important part of the analytical process, as an object without context would seldom be of any interest for the researcher.

Somehow, then, the objects need to be more related to each other and their social and historical context. One conceivable improvement in this respect would be to implement a function that suggests relevant content based on previously viewed objects. Providing links to related objects increases the possibility for the researcher to find as much interesting data as possible while at the same time helping to place these objects into their proper context.

Moreover, it was proposed that the objects must not remain “dead pieces”, but could in fact gain life through researchers’ engagement. For instance, comment functions would be useful in which users can make various types of annotations on the objects. For more comments on this matter, see 4.1.3.

### **16.2.3 Europeana as a transparent research source**

A central point stressed by several participants was that Europeana as a whole would benefit from increased transparency and clarity. First, issues were raised that Europeana’s homepage does not clearly present the project itself or who is behind it. Indeed, there is a link to the bottom left of the homepage leading to Europeana Professional, where such information can be found. In its current form, however, it was considered too vague and without the help of the project participants, the external experts would presumably not have found the information. Clearer information on this matter, including a guarantee that the content will be permanently available, is therefore required for enhancing the trustworthiness of Europeana as a research portal.

A second theme on transparency concerned the lack of readily available information regarding the general features of the content itself. For those interested in exploring the Europeana portal, it may be rather difficult at first glance to comprehend what is there. Possible solutions to this proposed during the discussions included to offer entry-level tutorials or demonstration projects for showcasing the content possibilities. This could be done favourably within content-rich Europeana topics such as World War I. Another recommendation was to highlight themes that are currently

## Europeana Cloud

the best developed within Europeana in order to attract researchers potentially interested in these areas.

Third, it was emphasized that researchers always need to examine their data critically, and that Europeana in its current form does not facilitate this activity sufficiently well. Material from cultural heritage institutions does not represent the “truth” about the past but one of several possible perspectives. General information about the selection process ought therefore to be easily available (e.g. what criteria exist for something to be included in the Europeana collection?). Another issue on this matter regarded the provenance of the objects included. Importantly, contact details of the providing institution should appear in connection to the record of an object. In order to be able to evaluate any possible bias in the aggregation, a list of all providers (e.g. organised nationally) would likewise be of great importance. Detailed information about what providing institutions had decided to include and what had been left out, and how much of a particular collection had been included would also be desirable.

### 16.2.4 Summing up the Contents session discussions

The Europeana material encompasses more than 30 million cultural heritage objects covering various themes from different periods of history. Given its scope, the collection could potentially attract researchers from a wide range of disciplines within the social sciences. In its current form, however, Europeana faces a number of fundamental challenges that have to be dealt with.

Researchers must always be able to evaluate their sources of data. Therefore, it would be important for Europeana to increase its transparency by providing clear information on who is behind the project, the provenance of the objects, and the selection process.

One of the main challenges for Europeana is to achieve a sufficiently high standard of metadata for its content. Without a general improvement in this regard, Europeana cannot be considered a solid research portal. However, it will not be very helpful to have high-quality metadata without also implementing appropriate search and filtering functions that, among other things, would help to contextualise the huge collection of objects within Europeana.

## 17. *Conclusions from Expert Forum 3*

This Expert Forum report echoes the conclusion of the report from Expert Forum 1: “It is clear [...] that participants do not consider Europeana a truly useful or robust resource without better quality metadata and a suite of user tools to aid browsing and analysis. Moreover, it is essential that this metadata can be viewed, filtered, and manipulated in a variety of ways.”<sup>31</sup> In fact, there was even greater emphasis placed on the need for transparency and contextualization. Europeana is not perceived as a trustworthy source of material, or at least not as a source which provides easy access to all the relevant information about provenance and context.

Key recommendations include:

- functionality for adding annotations, comments, and user-enhanced metadata to records
- robust, user-friendly functionality for export to non-proprietary, ubiquitous, and/or third-party software

---

<sup>31</sup> Ibid: 13.

## **Europeana Cloud**

- refined search functionality and additional filters
- an approach to additional content that focuses on quality in few areas/subjects/topics rather than quantity in many
- improved metadata quality
- clear and easy-to-find information on Europeana, its providers, collection strategies, and inclusion (and exclusion) criteria.

## ***Expert Forum 4 – “So we’ve built it, but have they come? Investigating barriers and opportunities for API usage among the AHSS community”***

A joint NeDiMAH / Europeana Cloud Workshop, 17<sup>th</sup> December 2014, The Hague, Netherlands

### **18. Introduction**

The aim of this workshop was to deliver an event to demonstrate the potential for API usage to non-technical members of the eCloud and NeDiMAH key researcher cohorts and to gather further detail on perceived barriers and possible solutions. We invited researchers and developers to talk about their research practices; and non-technical researchers in Humanities and Social Sciences to tell us whether they find the potential interesting and/or the skills required too difficult.

### **19. Background**

There are plenty of researchers using cultural data, and it may well be that in some cases their systems of data capture do make use, wholly or in part, of an API service. But in our initial studies of digital workflow practices, most of the researchers we were able to identify really only cared about the data, and had no specific opinions about how that data was accessed: they cared about the electricity (data) and the lightbulb (results enabled by the data), but not the plug socket (API or other data transfer service).

The term API, referring as it does to the specific aim of connecting the lamp to the electricity, seems therefore to be *a priori* restricted to the use of and by developers. As such, it seems the majority of users of cultural heritage APIs are still developers and computer scientists, although there is a small group of Humanities and Social Science researchers who are re-using the data they can obtain through a web-service or API. Who is doing the extracting of that data, however makes the difference. In the case of some exemplar digital humanists, they are making use of developers to make the calls to the APIs and obtain the data they need. They have the expertise to know what they can do with the data when they get it, but they don't 'dirty their hands' by writing the call to the API themselves. On the other hand, we have developers who not only write the code to call the API themselves, but also maintain the content for the API. We might call these people the 'data evangelists', as they showcase what can be done with a particular Cultural Heritage Institution's API and data.

But the potential future usage of APIs may not be reflected in the current patterns, as many of the current workflows scholars deploy engage similar functions and steps. This workshop therefore brought together and attempted to shed light on a full landscape of practice and possibility. As such, the event included perspectives of creators and developer/users of APIs, but also support services within the data-intensive humanities research lifecycle as well as those humanists reusing data themselves, with a specific focus on how they would acquire data and what they would want to do with it (that is, in most cases, what structure they would apply after download).

This workshop further investigated the workflows surrounding API use. In doing so, we were in a better position to determine the current state of the art of API use, the barriers, practices and

justifications, and develop a workflow that non-technical as well as technically competent humanists can follow in order to obtain big data sets using web services and APIs.

## 20. Storyboarding

Over the course of the day, the groups were asked to devise a series of research questions that they might want to answer. In doing so, the groups were asked to think of the data that would be required to answer their research question (be it a small or large part of the overall research), and how they might best obtain that information.

Once they had determined this, the groups were then asked to 'storyboard' the process by which they might obtain that information. This could be in the form of a website, an interface, or through coding.

### 20.1 Newspaper API Builder

The Linguistics group decided to work on an existing database of newspaper articles, and develop an interface that allowed the user to put a series of filters in place to generate a URL and/or the data output. These filters included the elements of a newspaper article that the search should concentrate on (e.g. headlines only, first sentences, by-lines, entire article), date of publication, date ranges, locations of publications, etc. Once all filters that are required have been applied, the URL and also is required, data output will then be shown in the right-hand side of the interface. This data output could be selected to appear in either: JSON, JSONP, XML or RDF.

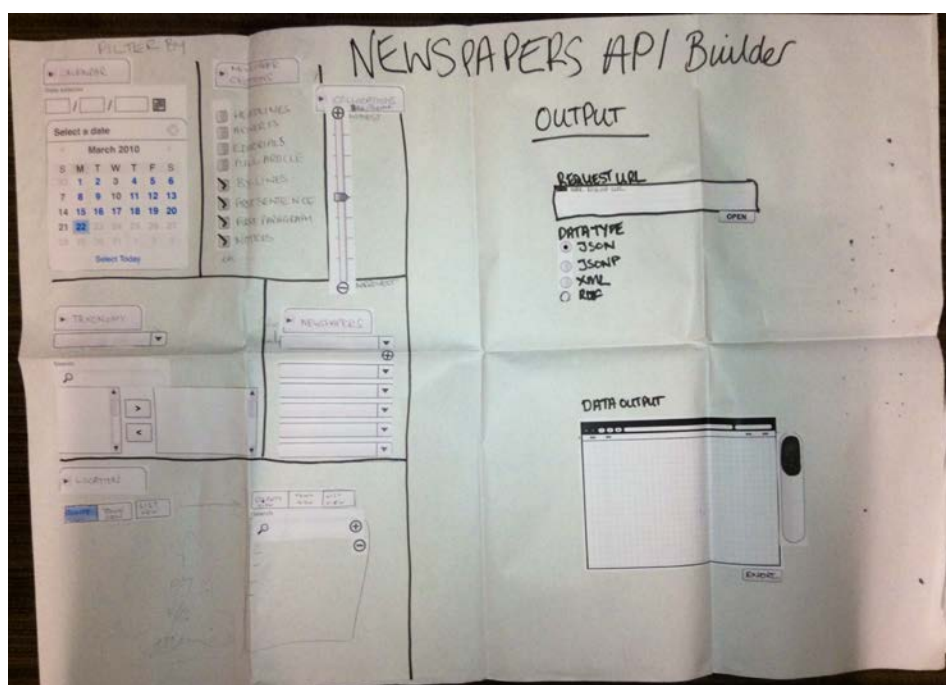


Figure 1 - "Newspaper API Request Builder" storyboard from the Linguistics group.

### 20.2 Image Clusters

The History group decided to look at how to combine all the different datasets that might be associated with images in collections. They devised an API that would visualise the output data in cluster formats. The filters that could be applied included visual information such as 'hair length' of

## Europeana Cloud

the subject and gender. Again, they used a split-screen approach, with the filtration occurring at the top, and the output occurring in a separate frame within the interface.

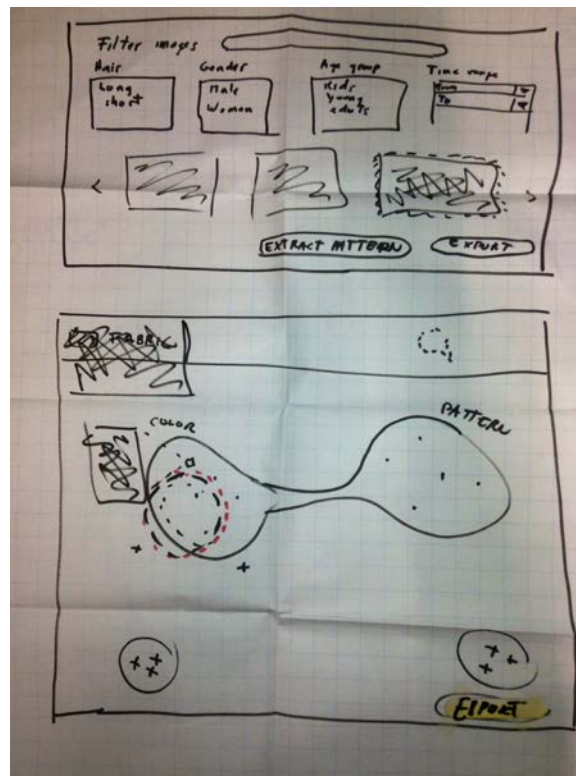


Figure 2 - Image clustering storyboard from the History group

### 20.3 The Allusionator

The Literature group decided to explore whether there were meaningful potential expansion avenues for tools like the Mallet software toolkit. They envisioned a two-layer approach to literary analysis, where by your model (which might be the result of topic modelling one or a small corpus of texts) was then applied as a template to a larger corpus. In this way, the user would be able to adjust the signal to noise ratio both in what they were looking for (granularity of the focus on the model) and where they were looking for it (size and composition of the corpus). By adjusting the controls on the model building kit, users would hopefully be able to gradually determine webs of interconnectedness among texts according to the proximity of the allusions contained in them. To assist in this, they would be supplied with an interface with multi-faceted filtering options, and visualising tools on a 3D plot diagram to show 'closeness' between texts.

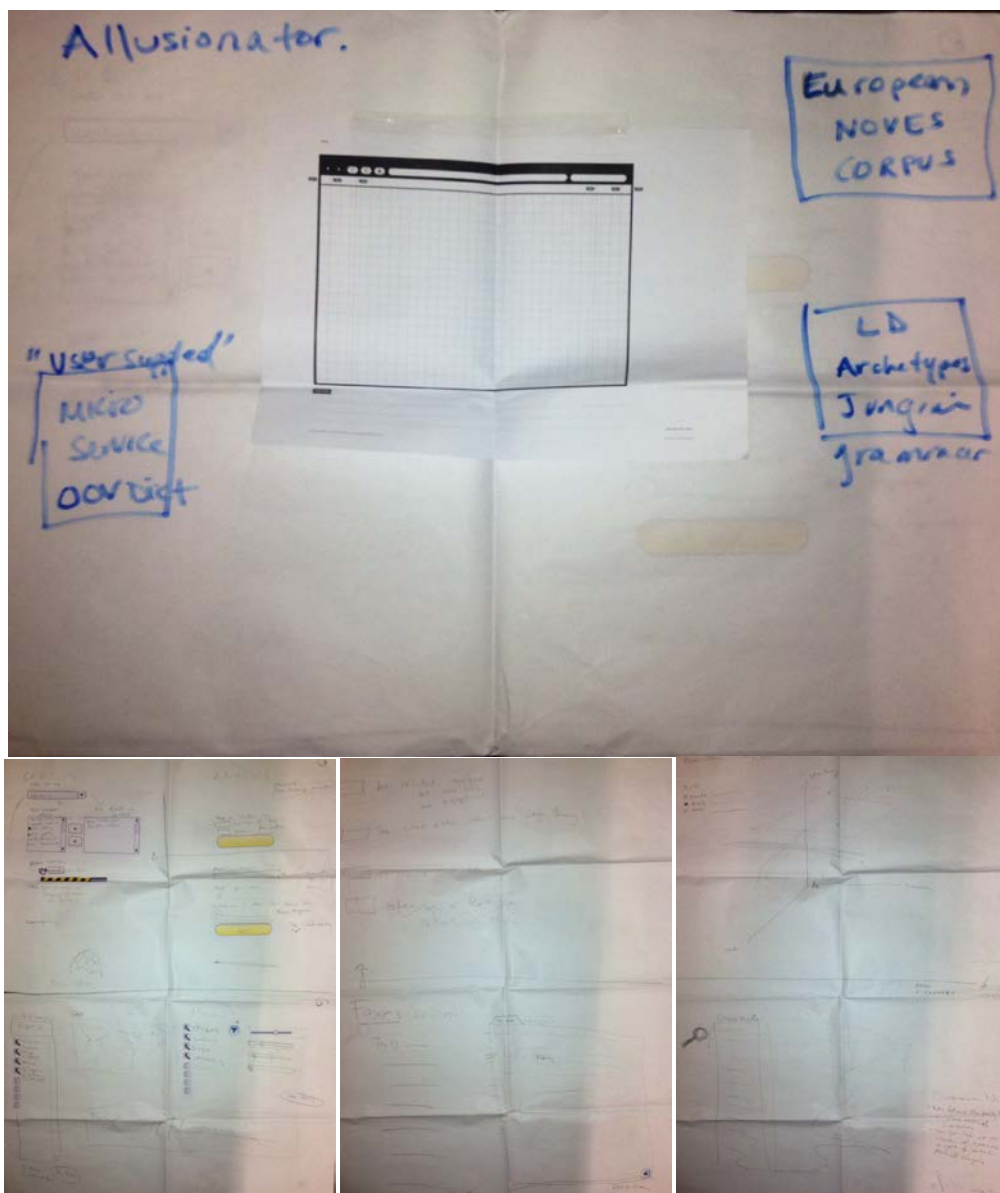


Figure 3 - Images of the Allusionator

## 21. **Recognising the Barriers: Enabling the Participants**

Aside from using the workshop to identify workflows for both technically proficient and non-technical humanists, the workshop also tried to ensure that the participants were able to use what they had learned during their time at the workshop as a launchpad to then go and put APIs into practice. In order to do this, we had to address the fundamentals of API use. In particular, the very basics of making a call to an API, and what that looks like. This small section of the day walked the participants through the process of making a call to an API. This included showing them tools such as code editors, the need for a particular syntax based on the coding language, writing the basic API call request, and how to ensure you get the data in the format you need it. We then showed what could be done with the data from the API call, and how it can be formatted into an excel spreadsheet for further analysis.



### **21.1 *Unknown Unknowns and Very Clever Things***

The response from the participants who had not previously seen this being done was positive, many of whom reported that while they were aware of coding and sites that could help to make an API call, they were completely in the dark about what kind of programme one might use to do that, and what you did with the resulting API call code. This, therefore, tapped into something that might often be overlooked by developers and data providers: the ‘unknown unknowns’ about data retrieval that are experienced by ‘analogue’ humanists. Many humanists who have not previously worked with digital data collections might be aware that Very Clever Things can be achieved using digital techniques, and may have seen those Very Clever Things in action, but they have no idea of the process involved in getting to that stage, and moreover, are not even aware of what they need to find out in order to give them this information.

Europeana therefore needs to look at how it might address these ‘unknown unknowns’ to the Analogue Humanist world in order to ensure that more of its potential users are able to get the most from the data collections on offer.

## **22. *Acknowledgements for the eCloud/NeDiMAH API workshop***

The organisers recognise the generous funding of NeDiMAH and the European Science Foundation for this workshop.

## ***Appendix I – List of Participants - Dublin***

Name	eCloud status	Inst	Research Field
Christine Morris	Non-eCloud	Trinity College Dublin	Archaeology
Erik Champion	Non-eCloud	Digital Humanities Lab Denmark	Archaeology
Ingrida Vosyliute	Non-eCloud	Vilnius University Faculty of Communication (VUFC)	Archaeology
Anthony Corns	Non-eCloud	Discovery Programme	Archaeology
Norman Rodger	eCloud WP1	University of Edinburgh	Archaeology
Mary Rowlatt	eCloud WP1	MDR Partners	Archaeology
Matt Munson	Non-eCloud	Gottingen Centre for Digital Humanities	History
Mark Sweetnam	Non-eCloud	Trinity College Dublin	History
Wilko Hardenberg	Non-eCloud	Rachel Carson Center, LMU	History
Thomas Baldwin	eCloud WP1	The European Library	History
Kees Waterman	eCloud WP1	NIOD	History
Rob Kitchin	eCloud RCAB member	NUI Maynooth	Social Science
Thoa Pam	Non-eCloud	Dublin Institute of Technology	Social Science
Jorge Ricardo da Costa Ferreira	Non-eCloud	Universidade Nova de Lisboa	Social Science
Stefan Ekman	eCloud WP1	University of Gothenberg	Social Science
Louise Edwards	eCloud WP1	The European Library	Social Science
Christian-Emil Ore	Non-eCloud	University of Oslo	Ontologies
Costis Dallas	eCloud WP1	University of Toronto	Ontologies
Øyvind Eide	Non-eCloud	University of Oslo	Ontologies
Alexander O'Connor	Non-eCloud	Trinity College Dublin	Ontologies
Torsten Reimer	Non-eCloud	JISC	Ontologies
Susan Reilly	eCloud WP1	LIBER	Ontologies
Owain Roberts	eCloud WP1	National Library of Wales	Ontologies
Agiati Benardou	eCloud WP1	DCU Athens	eCloud staff
Susan Schreibman	eCloud WP1	Trinity College Dublin	eCloud staff
Vicky Garnett	eCloud WP1	Trinity College Dublin	eCloud staff

## **Appendix II – Expert Forum Running Order - Dublin**

Expert Forum - Case Studies

18th June 2013

IIS Seminar Room (6.002, Arts Block)

Trinity College Dublin

9.25	House-keeping	Vicky Garnett, TCD
9.30	Session 1 About eCloud and the Aims of the Day	Hosted by Prof. Susan Schreibman, TCD Dr. Agiati Benardou, DCU Athens
9.50	Europeana 'Treasure Hunt' game	Vicky Garnett, TCD
10.20	Session 2 About Europeana	Prof. Costis Dallas, University of Toronto
10.40	Coffee Break	
11.00	Session 4 - Case studies - Europeana NOW - 1hr 30mins	Introduced by Vicky Garnett
12.30	Session 4 Reporting back on Scenarios	Facilitated by Vicky Garnett
13.00	LUNCH	
13.45	Session 5 Potential uses of Europeana - Discussion	Introduced by Susan Schreibman/ Agiati Benardou
14.15	Session 6 Revisiting case studies and presenting with how Europeana could/should be used	Introduced by Susan Schreibman
15.45	Reporting back	Facilitated by Susan Schreibman / Agiati Benardou
16.30	Day ends	

## ***Appendix III – Europeana Treasure Trail – Dublin***

### ***The Europeana Treasure Hunt!***



Before you start:

- Log in to [www.europeana.eu](http://www.europeana.eu) and create a 'My Europeana' profile for your team. To do this, you will need to create an account using an email address.
- If you wish, you could create this for one individual, as that person may wish to use Europeana at a later stage.
- If a team member already has a Europeana profile, feel free to use that, but don't use any previously saved searches for the Treasure Hunt!
- Use the 'My Europeana' function to save all your searches.
- You can only use one computer.

You have 15 minutes in total to answer the following questions. The team with the most points at the end of the hunt wins.

#### 1. Multiple Search Terms

How many different search terms does Europeana use for the major European war that began as a result of the assassination of Archduke Franz Ferdinand of the Austro-Hungarian Empire?

1 point for each search term found.

#### 2. Minimum Metadata

## Europeana Cloud

Using at least 3 different search terms, how many examples of items can you find that have a minimum of 10 metadata fields completed.

1 point for each example

### 3. Europeana-Whacks.

Using between 2-5 words, what is the fewest number of results you can find for a search term.

You score **1 point for each item produced** in a single search (for example, if a search term gives you 10 results, you score 10 points). The aim of this task is to score as few points as possible, as points scored in this round will be deducted from your overall score. Select your three lowest scoring search terms (and give the number of items found for each). Scores of '0' will not be accepted, and will instead incur a 5-point deduction from your overall score.

**Rules:** Do not use proper nouns, including place names or people's names. Search terms **MUST** be between 2-5 words long.

## ***Appendix IV – List of participants - Amsterdam***

<b>Name</b>	<b>eCloud status</b>	<b>Institution</b>	<b>Research field/Expertise</b>
Agiati Benardou	eCloud WP1	DCU Athens	Ancient History
Alastair Dunning	eCloud	The European Library Programme management	Programme manager
Claire Clivaz	non-eCloud	University of Lausanne	Early Christianity; digital editing
Eliza Papaki	eCloud WP1	DCU Athens	History
Hein van den Berg	non-eCloud	Technical University Dortmund	Philosophy in the Digital Humanities
Julianne Nyham	non-eCloud	University College London	Metadata for Digital Humanities
Karina van Dalen-Oskam	eCloud RCAB	KNAW - Huygens ING	Analyses, digital text corpora
Kees Waterman	eCloud WP1	KNAW - DANS	Early Modern History
Laurent Pugin	non-eCloud	Independent researcher	Early typographies; Music
Marian Lefferts	eCloud WP1	Consortium of European Research Libraries	Medievalist; rare books specialist
Marijn Koolen	non-eCloud	University of Amsterdam	Data retrieval; cultural heritage
Marnix van Berchum	eCloud WP6	KNAW - DANS	Musicology
Matthew Munson	non-eCloud	Goettingen Center for Digital Humanities	Text mining-mapping/GIS; Biblical/Early Testament
Max Kemman	non-eCloud	Erasmus University Rotterdam	Digital tools; audio-visual
Orla Murphy	non-eCloud	University College Cork	Digital editions; medieval history
Owain Roberts	eCloud WP1	National Library of Wales	Ontologies
Peter van der Maas	non-eCloud	Erasmus Rotterdam University	Oral History; audio-visual
Peter van Kranenburg	non-eCloud	KNAW - Meertens Insituut	Data retrieval; musicology
Vicky Garnett	eCloud WP1	Trinity College Dublin	Linguistics

## Appendix V – Agenda of the Expert Forum - Amsterdam



### Expert Forum – Tools & Content for Humanities Research NIOD, Amsterdam, November 11-12, 2013

#### Day 1

13.30	House-keeping	Dineke de Visser / Kees Waterman
13.35	Introducing Europeana & eCloud The aims of the Expert Forum	Hosted by Alastair Dunning / Agiati Benardou
14.00	The Europeana Treasure Trail	Hosted by Vicky Garnett
14.20	Introducing the assignment (brainstorm) on <b>tools</b>	Introduced by Kees Waterman
14.35	Coffee break	
14.50	Session 1: assignment in breakout sessions	Discussion groups accompanied by eCloud representatives
16.15	Session 1: reporting back on breakout sessions	Discussion facilitated by Agiati Benardou / Kees Waterman
16.45/17.00	Day ends	
(17.00)	WP1 meeting (project participants only)	
18.30	Dinner at “Kantjil & de Tijger”, close to the NIOD	

#### Day 2

9.15	Introducing the assignment on (brainstorm) on <b>content/new material</b>	Introduced by Kees Waterman
9.30	Session 2: assignment in breakout session	Discussion groups accompanied by eCloud representatives
11.00	Coffee break	
11.15	Session 2: reporting back on breakout sessions	Discussion facilitated by Eliza Papaki / Kees Waterman
11.45	Summary, general feedback	Agiati Benardou, Karina van Dalen- Oskam, Kees Waterman
12.00	Lunch (on your own)	

## **Appendix VI – The Europeana Treasure Hunt - Amsterdam**



### **The Europeana Treasure Hunt!**

Before you start:

- Log in to [www.europeana.eu](http://www.europeana.eu) and create a 'My Europeana' profile for your team. To do this, you will need to create an account using an email address. If a team member already has a Europeana profile, feel free to use that, but don't use any previously saved searches for the Treasure Hunt!
- Use the 'My Europeana' function to save all your searches.
- You can only use one computer per team.

You have 15 minutes in total to answer the following questions. The team with the most points at the end of the hunt wins. ONLY ENGLISH WORDS ARE ACCEPTED!

#### 1. Europeana content types

Using at least 2 search terms, can you come up with searches that give results, which contain items of all 5 of Europeana's content types (image, text, sound, video, 3D)?

Points: 3 points for each search with all 5 content types.

#### 2. Metadata

Each Europeana item is described by a number of metadata fields. How many different metadata fields can you find in total? (You can add fields from different items.)

Points: 0.5 points/metadata field

#### 3. Europeana Whack.

Using 2 search terms and searching "All fields", what is the fewest number of results you can get?

Points: Searches that give 1 result = 10 points; 2 results = 5 points; 3 results = 1 point.

*Whack examples:* nice trophy, imprint bike

**Rules:** Do not use proper nouns, including place names or people's names. Make sure that the "Search"/"Search all fields"-function is used.



## **Appendix VII – List of participants Expert Forum 3 – Gothenburg**

<b>Name</b>	<b>eCloud status</b>	<b>Institution</b>	<b>Research field<sup>32</sup></b>
Agiati Benardou	eCloud WP1	DCU Athens	Ancient History
Björn Sjögren	eCloud WP1	University of Gothenburg	Psychology
Dimitrios Kokkinakis	Non-eCloud	University of Gothenburg	Text mining
Ilze Lace	eCloud WP1	University of Gothenburg	Sociology
Isto Huvila	Non-eCloud	Åbo Akademi University	Information studies
Leif Isaksen	eCloud RCAB	University of Southampton	Archaeology
Linda Lane	Non-eCloud	University of Gothenburg	Sociology
Lorna Hughes	eCloud WP 1	National Library of Wales	Digital Humanities
Marisa Ponti	Non-eCloud	Chalmers University	Learning research
Norman Rodger	eCloud WP 1	University of Edinburgh	Archaeology
Owain Roberts	eCloud WP 1	National Library of Wales	Ontologies
Peter van den Besselaar	Non-eCloud	Vrije University Amsterdam	Infometrics
Ralph Shroeder	Non-eCloud	Oxford Internet Institute	e-Social Science
Stefan Ekman	eCloud WP 1	University of Gothenburg	Economics
Susan Reilly	eCloud WP 1	LIBER	Ontologies
Thomas Baldwin (WP4)	eCloud WP 1	The European Library	LIS
Thomas Hillman	Non-eCloud	University of Gothenburg	Education
Wera Grahn	Non-eCloud	Linköping University	Gender studies
Vicky Garnett	eCloud WP 1	Trinity College Dublin	Linguistics

<sup>32</sup> Many researchers were active in more than one field or involved in inter-disciplinary research. The disciplines given reflect a rough disciplinary belonging only.

## Appendix VIII – Agenda of the Expert Forum - Gothenburg



### Expert Forum – Tools & Content for Social Science Research

**October 24-25, 2013**

Lundgrensgatan 7 “Språkskrapan” – 8<sup>th</sup> floor, Conference Room  
University of Gothenburg

#### Day 1

13.30	House-keeping	Björn/Ilze/Stefan
13.35	Introducing Europeana & eCloud The aims of the Expert Forum	Hosted by Agiati/Lorna/Susan R
14.00	Europeana game	Björn/Ilze/Stefan
14.30	Coffee break	
14.50	Introducing the assignment on <b>tools</b> Session 1: assignment in breakout sessions	Introduced by Lorna/Susan R
16.30	Session 1: reporting back on breakout sessions	Facilitated by Lorna/Susan R
17.00	Day ends	
(17.00)	WP1 meeting (project participants only)	
18.30	Dinner at “Dinner 22”	

#### Day 2

9.00	Introducing the assignment on <b>content/new</b> material Session 2: assignment in breakout session	Introduced by Lorna/Susan R
10.45	Coffee break	
11.00	Session 2: reporting back on breakout sessions	Facilitated by Lorna/Susan R
11.30	Summary, general feedback	Lorna/Susan Agiati/Stefan/Björn/Ilze R/
12.00	Lunch at the forum venue	

## **Appendix IX – The Europeana Treasure Hunt - Gothenburg**



### **The Europeana Treasure Hunt!**

Before you start:

- Log in to [www.europeana.eu](http://www.europeana.eu) and create a 'My Europeana' profile for your team. To do this, you will need to create an account using an email address. If a team member already has a Europeana profile, feel free to use that, but don't use any previously saved searches for the Treasure Hunt!
- Use the 'My Europeana' function to save all your searches.
- You can only use one computer per team.

You have 15 minutes in total to answer the following questions. The team with the most points at the end of the hunt wins. ONLY ENGLISH WORDS ARE ACCEPTED!

#### 1. Europeana content types

Using at least 2 search terms, can you come up with searches that give results which contain items of all 5 of Europeana's content types (image, text, sound, video, 3D)?

Points: 3 points for each search with all 5 content types.

#### 2. Metadata

Each Europeana item is described by a number of metadata fields. How many different metadata fields can you find in total? (You can add fields from different items.)

Points: 0.5 points/metadata field

#### 3. Europeana Whack.

Using 2 search terms and searching "All fields", what is the fewest number of results you can get?

Points: Searches that give 1 result = 10 points; 2 results = 5 points; 3 results = 1 point.

*Whack examples:* nice trophy, imprint bike

**Rules:** Do not use proper nouns, including place names or people's names. Make sure that the "Search"/"Search all fields"-function is used.

## **Appendix X – Running Order of eCloud/ NeDiMAH API Workshop – The Hague**

<b>Morning Session</b>	
<b>9.00am</b>	Arrival
<b>9.15am</b>	Welcome from eCloud WP1 Leader, Agiati Benardou
<b>9.15am</b>	Context and Goals for the Day – What have we learned so far on API and web-service use? – Dr. Jennifer Edmond
<b>10.00am</b>	Adam Crymble - Digital histories - Data Mining
<b>10.15am</b>	Paul Rayson - Historical linguistics/Psychology
<b>10.30am</b>	Mark Sweetnam - 1641 Depositions/Cultura
<b>10.45am</b>	Coffee
<b>11.00am</b>	<i>Morning break out groups</i> - Rapid ideation session on humanities research ‘data to knowledge’ pipelines
<b>12.15pm</b>	Presenting the Outcomes
<b>12.45pm</b>	Lunch
<b>Afternoon session</b>	
<b>1.45pm</b>	<b>Basics of API</b> <ul style="list-style-type: none"> <li>- what a call looks like</li> <li>- and what the data looks like.</li> </ul> - What an API <u>CAN'T</u> do.
<b>2.15pm</b>	Gonzalo Parra, Work Package 3 - Europeana work
<b>2.30pm</b>	Dimitris Gavrilis, LoCloud/DCU Athena
<b>2.45pm</b>	Paul McCann, National Library of Wales
<b>3.00pm</b>	Questions for Developers
<b>3.15pm</b>	<b><i>Afternoon break out groups - Storyboarding use cases</i></b>  3.15pm       Decide on which one of the morning use-cases from each group provides the best use of the API, in terms of how it can pull, process, and present data 3.30pm       Storyboard out use case of answering that research question using those tools, in a research environment.
<b>4.30pm</b>	Final Discussion about the process of developing prototypes
<b>5.00pm</b>	Concluding comments – close.

***Appendix XI – List of Participants at eCloud / NeDiMAH API Workshop – The Hague***

<b>Name</b>	<b>Institution</b>
Adam Crymble	University of Hertfordshire
Gonzalo Parra	KU Leuven
Ruxandra Elena Bularca	Transylvania Digital Humanities Research Centre
Teodora Shek Brnardić	Croatian Institute of History
Paul McCann	National Library of Wales
Trilce Navarrete	University of Southern Denmark
Emma Clarke	Trinity College Dublin
Sara O'Sullivan	University College Dublin
Anouk Lang	University of Edinburgh
Agiati Bernardou	DCU Athens
Alastair Dunning	The European Library
Costis Dallas	University of Toronto
Jennifer Edmond	Trinity College Dublin
Vicky Garnett	Trinity College Dublin