DELIVERABLE

Project Acronym: **Europeana Cloud**

Grant Agreement number: **325091**

Project Title: **Europeana Cloud: Unlocking Europe's Research via The Cloud**

Deliverable 1.5 (3 of 4): Expert Forum Tools & Content for Social Science Research

Revision: 2

Authors:

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

Revision History

Revision	Date	Author	Organisation	Description
1	21 Nov, 2013	Stefan Ekman and Björn Sjögren	Swedish National Data Service/University of Gothenburg	First draft of report sent to RCAB
2	29 Nov, 2013	Stefan Ekman and Björn Sjögren	Swedish National Data Service/University of Gothenburg	Draft revised following comments from RCAB and the Project Coordinator

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This deliverable contains original unpublished work except where clearly indicated otherwise.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.

D1.5 (3 of 4):
Expert Forum Tools & Content for Social Science Research

Executive summary

This report deals with the Expert Forum that took place in Gothenburg October 24-25 to assess social scientists needs and requirements for using Europeana as a research source. The 19 participants either belonged to the Europeana Cloud projects or were invited as external social scientists experts.

The forum had two themes; (a) what kinds of tools would be useful, and could potentially be developed, for social scientists carrying out research on the Europeana material; and (b) 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.

The discussions 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.

Table of Contents

1. 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".

1.1 Expert Forum 2 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.[1]

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.

---

[1] See Deliverable 1.5 (1 of 4): *Expert Forum Case Studies Report*.

1.2 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.

1.3 How the results will be used

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

2. Selecting the research areas and the participants

2.1 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.[2] 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.

---

[2] These were, in alphabetical order, Economic & Social History, Gender Studies, Human/Economic/Political/Cultural Geography, Political Science, Social Anthropology, and Sociology.

2.2 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.

2.3 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.[3] 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, 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.

---

[3] See Appendix I for a complete list of participants.

3. 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.[4] 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.

3.1 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.[5]

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

---

[4] See Appendix II for the entire agenda.

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.

3.2 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.

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.

3.3 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

---

[5] The complete instructions given to the participants can be found in Appendix III.

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?

4. 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.

4.1 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?

4.1.1 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 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.

4.1.2 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.

4.1.3 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 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 mention 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.

4.1.4 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 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.

4.2 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.)

4.2.1 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.[6] 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).

---

[6] Deliverable 1.5 (1 of 4): 8.

4.2.2 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.

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[7], 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.

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[8], 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

---

[7] Ibid.
[8] Ibid.

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 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.

4.2.3 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 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.

4.2.4 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.

5. Conclusion

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."[9] 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
- 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.

---

[9] Ibid: 13.

Appendix I – List of participants

Name	eCloud status	Institution	Research field[10]
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	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	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

---

[10] 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 II – Agenda of the Expert Forum

Expert Forum – Tools & Content for Social Science Research
October 24-25, 2013
Lundgrensgatan 7 "Språkskrapan" – 8th 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 **tools** 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 **content**/new 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 R/ Agiati/Stefan/Björn/Ilze
12.00	Lunch at the forum venue	

Appendix III – The Europeana Treasure Hunt

The Europeana Treasure Hunt!

Before you start:
- Log in to 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.