



DELIVERABLE

Project Acronym: Europeana Cloud

Grant Agreement number: 325091

Project Title: Europeana Cloud: Unlocking Europe’s Research via The Cloud

Deliverable 1.5 (2 of 4): Expert Forum Tools & Content for Humanities Research Report

Revision: 2

Authors:

Kees Waterman (KNAW - DANS)

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
1	December 7, 2013	Kees Waterman	KNAW - DANS	First draft of report sent to relevant colleagues in WP1.
2	December 11, 2013	Kees Waterman	KNAW - DANS	Circulated to participants for comments.
3	December 19, 2013	Kees Waterman	KNAW - DANS	Final, comments and remarks of readers incorporated.

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This deliverable contains original unpublished work except where clearly indicated otherwise.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 (2 of 4):

Expert Forum Tools & Content for Humanities Research

Executive summary

This report deals with the Expert Forum that took place in Amsterdam November 11-12 to assess humanists 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 humanists experts.

The forum had two themes; (a) what kinds of tools would be useful, and could potentially be developed, for humanists 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 humanities research.

The discussions provided the following key recommendations for Europeana:

- 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.

Table of Contents

1. Introduction to the task	5
1.1 Expert Forum 3 in the context of Work Package 1.....	5
1.2 The purpose of the forum.....	6
1.3 How the results will be used.....	6
2. Selecting the research areas and the participants	6
2.1 Research areas	6
2.2 Selecting the experts	7
2.3 Participants	7
3. Designing the forum	8
3.1 Europeana Treasure Hunt.....	8
3.2 Session 1 – Tools	9
3.3 Session 2 – Content	10
4. Results	11
4.1 Tools	11
4.1.1 Summing up the Tools session discussions.....	16
4.2 Content	17
4.2.1 Summing up the Contents session discussions	21
5. Conclusion	23
Appendix I – List of participants	25
Appendix II – Agenda of the Expert Forum	26
Appendix III – The Europeana Treasure Hunt	27

1. 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.

1.1 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.¹

Expert Forums 2 and 3 (held in Amsterdam, project month 10, and Gothenburg, project month 9²) 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,

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

² 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.

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.

1.2 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.

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

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.³ 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

³ These were, in alphabetical order, Archaeology, History, Law, Linguistics, Musicology and Philosophy.

not employ digital tools in their work would probably contribute less to the discussion than researchers well versed in deploying digital research.

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 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.

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.⁴ 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, 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.

⁴ See Appendix I for a complete list of participants.

2. 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.⁵ 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.

3.1 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.⁶

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

⁵ See Appendix II for the entire agenda.

⁶ The complete instructions given to the participants can be found in Appendix III.

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.

3.2 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.

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 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:

text, image, video and sound.

The fifth content type, 3D visualizations or constructs, is still relatively scarce in Europeana.}

4. 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.

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.

4.1 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 16th-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, 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 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.

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.⁷ 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?"⁸ RISM plays a role here, building inventories to know where the sources are. Hence, developing a partnership with RISM seems a viable option.⁹

⁷ <http://wp1187670.server-he.de/e4d/> (accessed December 1, 2013).

⁸ Compare the IMLSP, International Music Score Library Project, <http://imslp.org> (accessed, December 1, 2013).

⁹ RISM, Répertoire International des Sources Musicales - International Inventory of Musical Sources, <http://www.rism.info> (accessed, December 1, 2013).

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,¹⁰ 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.¹¹ 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).

4.1.1 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

¹⁰ <http://www.perseus.tufts.edu>, <http://pleiades.stoa.org> (both accessed December 3, 2013).

¹¹ 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).

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.¹² Some participants expressed the opinion that Europeana’s 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.

4.2 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.”

¹² 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.’

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 musicology¹³ or even other disciplines, i.e. to datasets of performance history, or connect it to church history. Europeana might be able to enhance its

¹³ Links to and cooperation with for instance RISM and IMLSP were among the suggestions here, see notes 8 and 9.

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).¹⁴

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

¹⁴ 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).

digital repositories have already joined such projects, aiming to develop as one of these put it “an innovative *cross-archival semantic content discovery platform*.”¹⁵

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 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.

¹⁵ 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).

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.¹⁶

4.2.1 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

¹⁶ 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>.

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 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.”¹⁷

¹⁷ 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.

5. Conclusion

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.¹⁸

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;

¹⁸ For a ‘shining example,’ see this US government website, <http://www.data.gov> (accessed December 5, 2013).

- 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 *“tap in”* to existing [research] communities.”¹⁹ 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.²⁰ 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.²¹ 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.

¹⁹ D1.2 State of the Art Report on Digital Research Practices, Tools and Scholarly Content Use, 51, emphasis added.

²⁰ D1.5 (3 of 4) Expert Forum Tools & Content for Social Sciences Research Report, 11.

²¹ *Ibid.*, 15.

Appendix I – List of participants

Name	eCloud status	Institution	Research field/Expertise
Agiati Benardou	eCloud WP1	DCU Athens	Ancient History
Alastair Dunning	eCloud, programme manager	The European Library	Programme management
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 University Rotterdam	Oral History; audio-visual
Peter van Kranenburg	non-eCloud	KNAW - Meertens Insituut	Data retrieval; musicology
Vicky Garnett	eCloud WP1	Trinity College Dublin	Linguistics

Appendix II – Agenda of the Expert Forum



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	<i>Hosted by</i> Alastair Dunning / Agiati Benardou
14.00	The Europeana Treasure Trail	<i>Hosted by</i> Vicky Garnett
14.20	Introducing the assignment (brainstorm) on tools	<i>Introduced by</i> Kees Waterman
14.35	Coffee break	
14.50	Session 1: assignment in breakout sessions	<i>Discussion groups accompanied by eCloud representatives</i>
16.15	Session 1: reporting back on breakout sessions	<i>Discussion facilitated by</i> 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 content /new material	<i>Introduced by</i> Kees Waterman
9.30	Session 2: assignment in breakout session	<i>Discussion groups accompanied by eCloud representatives</i>
11.00	Coffee break	
11.15	Session 2: reporting back on breakout sessions	<i>Discussion facilitated by</i> Eliza Papaki / Kees Waterman
11.45	Summary, general feedback	Agiati Benardou, Karina van Dalen-Oskam, Kees Waterman
12.00	Lunch (on your own)	

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.