UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

D3.3 Evaluation report: Evaluation report on integration of tools with Europeana

Benardou, A.; Duval, E.; Parra, G.; Charleer, S.; Jentzsch, A.; van den Berg, H.; Stoitsis, G.; van Berchum, M.; Stavrakaki, M.

Publication date
2016
Document Version
Final published version

Link to publication

Citation for published version (APA):
Benardou, A., Duval, E., Parra, G., Charleer, S., Jentzsch, A., van den Berg, H., Stoitsis, G., van Berchum, M., & Stavrakaki, M. (2016). *D3.3 Evaluation report: Evaluation report on integration of tools with Europeana*. (v21 ed.) Europeana Cloud.

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:18 Nov 2026

DELIVERABLE

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

D3.3 Evaluation report
Evaluation report on integration of tools with Europeana

Revision: v21

Authors:

Agiatis Benardou (Athena RC)
Erik Duval (KU Leuven)
Gonzalo Parra (KU Leuven)
Sven Charleer (KU Leuven)
Anja Jentzsch (Open Knowledge Foundation Deutschland)
Hein van den Berg (Stichting VU-VUMC)
Giannis Stoitsis (ARIADNE Foundation)
Marnix van Berchum (KNAW-DANS / Utrecht University)
Maritina Stavrakaki (ARIADNE Foundation)

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

Revision History

Revision	Date	Author	Organisation	Description
V1	16 Jan 2015	Gonzalo Parra	KU Leuven	Adding the structure to include the work from Year 2
V2	27 Jan 2015	Gonzalo Parra	KU Leuven	Input for the evaluation of Year 2
V3	28 Jan 2015	Erik Duval	KU Leuven	Small updates and corrections
V4	28 Jan 2015	Gonzalo Parra	KU Leuven	Small corrections
V5	30 Jan 2015	Hein van den Berg	VU	Updated introduction and summary
V6	30 Jan 2015	Marnix van Berchum	DANS-KNAW / Utrecht university	Updated musicology related content
V7	1 Feb 2015	Erik Duval	KU Leuven	Consolidation, minor edits
V8	2 Feb 2015	Anja Jentzsch	OKF DE	Minor edits
V9	14 Jan 2016	Sven Charleer	KU Leuven	Input from Year 3
V10	14 Jan 2016	Anja Jentzsch	OKF DE	Update on Year 3
V12	25 Jan 2016	Maritina Stavrakaki	ARIADNE Foundation	Update and revision on Year 3
V13	25 Jan 2016	Sven Charleer	KU Leuven	Revision of Year 3
V14	25 Jan 2016	Hein van den Berg	VU	Rewriting, editing
V15	29 Jan 2016	Anja Jentzsch	OKF DE	Template and typos
V16	29 Jan 2016	Maritina Stavrakaki	ARIADNE Foundation	Rewriting, editing
V17	29 Jan 2016	Sven Charleer	KU Leuven	Rewriting, editing
V18	29 Jan 2016	Hein van den Berg	VU	Rewriting, editing
V19	29 Jan 2016	Sven Charleer	KU Leuven	Rewriting, editing
V20	30 Jan 2016	Maritina Stavrakaki	ARIADNE	Rewriting, editing
V21	31 Jan 2016	Anja Jentzsch	OKF DE	Template and typos

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.

D3.3 Evaluation report
Evaluation report on integration of tools with Europeana

Executive summary

This deliverable reports on the evaluation of the tools developed and interlinked in Work Package 3 (WP3). Three sets of prototype demonstrator tools have been created, integrated and evaluated. In this document, we present an evaluation of these tools.

This document is a companion document to Deliverable 3.2, the actual set of tools itself, with an accompanying document that briefly describes the tools and refers to where the software can be accessed on the Web.

This is the final version of the deliverable.

Table of Contents

1. Introduction 3
2. Methodology 6
3. Evaluation with WP1 7
4. Evaluation with Axiom Group 9
General discussion 9
Other comments 12
Participants in this session 12
Publication: 12
5. Evaluation with Musicologists 13
General discussion 13
Other comments 14
Participants in this session 15
Publication 15
6. Evaluation with the Europeana Newspaper Archive 16
Workshop results 16
General discussion 17
Evaluation results regarding usability of interactive tabletop vs multiple tablets 17
Evaluation results with expert users of the online version of NEE 18
7. Evaluation with Agricultural Researchers 21
General discussion 21
Workshop results 21
Participants in this session 22
8. Conclusion 23
References 24

1. Introduction

WP3 aims at developing services and tools that leverage Europeana content in the Europeana Cloud for researchers. During the first months of Year 1 of the project, WP3 focused on the development of personas, scenarios and use cases, in order to understand and analyse the user needs. This initial work on personas, scenarios and use cases was reported in D3.1.

In the first year of development, our target community of humanities researchers who work with Europeana content was the Axiom Group of Philosophy[1]. For this community, three core problems were identified through the personas and scenarios:

(i) problems with navigating and identifying relevant (digital) content and problems with building corpora;
(ii) a lack of user-friendly tools for conducting fine-grained textual research;
(iii)a lack of appropriate tools and infrastructure that allow members of research groups to work collaboratively.

Having the scenarios and personas at hand, we evaluated tools to support the workflow of our specific target audience. We then compiled a catalogue of tools that could be adjusted to and integrated in the Europeana Research Platform.

Based on this work, we created, integrated and evaluated a first set of prototype demonstrator tools:

- an ARIADNE finder personalised for the Axiom philosophy group to help researchers search and find content coming from Europeana and other sources;
- the TimeMapper, an integrated visualisation tool to visualise the search results on a timeline and an interactive map to further filter the content and get a better overview of the different resources found on Europeana;
- an Activity Stream, integrated in the above tools to capture and present the different actions carried out in this process (search, visualise, explore, annotate, download).

In the second year of work, our target community of humanities researchers who work with Europeana content was a community of (digital) musicologists that focus on Early Music. For this community, four core problems were identified through discussions in which we used the Year 1 toolset to trigger feedback and comments:

(i) difficulty of (meta-)data creation,
(ii) lack of digital corpora with music scores,
(iii) information exchange and linking of data when working with different tools,
(iv) retrieval and analysis of contextual information about the music scores.

In order to help the musicologists with tackling these problems, and building on the work done in the first year, we created, integrated and evaluated a second set of prototype demonstrator tools that extended the Year 1 toolset:

- Ariadne Finder personalized for musicologists;
- TimeMapper;
- the Activity Stream;
- Aruspix, an optimal music recognition (OMR) tool which transforms prints of early music scores into XML encoded music scores[2];

[1] Reference?
[2] http://www.aruspix.net/

- Music 21, a python based set of tools for analysing music, developed at MIT (http://web.mit.edu/music21/).

In the third year of work, our target community of humanities researchers were those working with the Europeana Newspapers Archive. By organising two workshops with researchers, we identified a large list of requests to better support their workflow. Out of this long list, we prioritised which issues should be addressed first:

(i) difficulties navigating the online Newspaper Archive
(ii) poor usability
(iii) limited facets for searching
(iv) automatic suggestions of relevant articles to the one the user is accessing
(v) timeline per newspaper

In order to help these researchers, we built upon the results of previous years and developed the Newspaper Exploration Environment. It offers a more visual approach to exploration of the archive and a a recommender system.

Furthermore we developed a data mining tool, eCloudDM, that

- extracts named entities from the newspaper articles
- defines topic tags for the newspaper articles

Also in the third year of work, we targeted a community of agricultural researchers working at the Agriculture Economics and Policy Research Institute (AGRERI). Through interviews, an online survey and a workshop with these researchers, we were able to identify their needs and requirements as well as some core problems that they face when it comes to using search tools:

(i) problems with navigating and identifying relevant (digital) content from diverse sources
(ii) the need to use specific facets and filters, in order to make it easier to discover resources related to specific themes

In order to help these researchers, we developed the AGRERI Discovery Microsite. It:

- searches predefined collections of datasets based on user input and presents the results in a uniform way
- provides a faceted search interface that allows users to search and quickly filter the results
- allows the discovery of the content from Europeana and other sources, and connects this content to the existing content for agricultural economics

In this document, we present an evaluation of the tools. This document is a companion document to D3.2 which describes the tools and refers to where the software can be accessed on the Web.

Both D3.2 and D3.3 relate to Task 3.2. From the description of work:

> *Task 3.2 Iterative design, development and evaluation of tools [M1-36]*
>
> *Four of the WP partners will build upon 3.1 in order to adapt existing or develop new service/tool demonstrators that can illustrate how Europeana content will be put in use. It will assist WP1, offering the necessary materials, guidelines and facilitation support in order to help them organize iterative design workshops/sessions on envisaged services and tools. It will also feed into WP4 giving, guidance on the related metadata requirements.*
>
> *3.2.1 Initial brainstorming and scenario building [M1-3, M12-15, M24-27]*

3.2.2 Paper prototypes to gather initial feedback on early ideas [M3-6, M15-18, M27-30]
3.2.3 Gradual development of more functional digital prototypes in rapid iteration cycles [M6-9,M18-M21,M30-M33]
3.2.4 Deployment of final implementations in realistic testbeds [M9-12, M21-24, M33-36]. Deliverable 3.2 available at M12, 24 and 36
3.2.5 Evaluation of technical integration of tools with Europeana Content [M9-12, M21-24, M33-36] Deliverable 3.3 available at M12, 24 and 36

2. Methodology

The basic methodology is that of User Centered Design [Abras, Maloney-Krichmar & Preece, 2004]. As mentioned on the Wikipedia article about User Centered Design[3]:

> *The chief difference from other product design philosophies is that user-centered design tries to optimize the product around how users can, want, or need to use the product, rather than forcing the users to change their behavior to accommodate the product.*

In this context, we have had regular formative evaluation sessions over Skype or Google Hangout sessions throughout 2013 and 2014. We also had face-to-face meetings toward the end of the Year 1 and Year 2 cycle to evaluate more in-depth the resulting toolsets. In year 3, we focused on usability evaluations face-to-face with a general public, and online evaluations with expert users.

It is important to note that the evaluation sessions took a broad view of usefulness and usability, i.e. we focused on whether or not the WP3 toolset would actually be of any substantial added value to the researchers involved. We wanted, more specifically, to find out whether our approach could help them to actually *change* the way they work, and whether such an approach would address problems that they may or may not be aware of in their current way of working. We were *not* interested in whether the users could carry out their current way of working more efficiently.

[3] https://en.wikipedia.org/wiki/User-centered_design

3. Evaluation with WP1

Note This section reports on joint work with WP1. Some of this section may also become part of deliverable 1.7 (Research Community Evaluation Report), which reports on that work from the WP1 perspective, due in Month 36.

This section reports on an evaluation workshop, via Skype on November 26th, 2013. The workshop focused on the usefulness of the tools listed in D3.1, which were considered for inclusion in the mash-up then under development in WP3. (See D3.1 for more details on the tools.)

Participants in the workshop were: Lorna M. Hughes (NLW), Erik Duval (KU LEUVEN/WP3 Leader), Vicky Garnett (TCD), Owain Roberts (NLW), Stefan Ekman (SND), Thomas Baldwin (CERL), Eliza Papaki (ATHENA R.C.), Björn Sjögren (SND), Pavel Kats (EF), Gonzalo Parra (KU LEUVEN), Hein van den Berg (Vrije Universiteit Amsterdam), Dimitris Gavrilis (ATHENA R.C.), Andreas Drakos (ARIADNE Foundation), Alastair Dunning (EF/Project Co-Ordinator), Agiatis Benardou (ATHENA R.C./WP1 Leader)

The discussion considered the different tools listed in turn:

- **ARIADNE finder:** Unlike for instance Google, the ARIADNE finder restricts searches to specific collections relevant to the community of researchers involved. One specific example was discussed during the workshop: a search for the philosopher and mathematician Bolzano within the collections aggregated by Euopeana. The participants agreed that this tool serves a major stage in the research cycle 'search and discovery'. They suggested that it would be useful if the ARIADNE finder included thematic descriptions of content and metadata held both within Europeana and elsewhere, regardless of the possibility to access the material itself. It was further suggested to have a demo of not just text, but also of other types of Europeana content, such as image, audiovisual and 3D. The ability to preview audiovisual and 3D records would further nicely complement the view possibility offered by the ARIADNE Finder. Finally, when searching with the ARIADNE finder, it would be useful for a user to have the possibility to search by chronological area as well, and for the results to be sorted by specific thematic tags.

- **Visualisation tools:**
  - **TimeMapper:** There was basic agreement that this kind of tool would be of particular use to historians, art historians, and archaeologists. The main concern was whether this approach would scale up to larger collections.
  - **RelFinder:** The main concern related to this tool is that it draws the data from dbpedia, rather than from Europeana. Moreover, RelFinder does not necessarily guide the user to relevant content, as it is focused more on the relations between the terms considered.
  - **Muse:** There was quite a bit of discussion around whether this visualization tool would be interesting for researchers in Humanities and Social Sciences. In any case, this tool requires a multitouch table for interaction and is thus currently less relevant for experimentation in eCloud.
  - **GlamMap:** Again, discussion focused on whether this approach would transfer beyond the original scope of supporting philosophers.
- **Awareness tools:**
  - **TiNYARM:** There was consensus that this tool represents a good first step toward the promotion of collaboration and awareness, which could, at a later stage, be extended

to material other than reading documents and publications. Similar to reading trends promoted and encouraged through TiNYARM, creating trends inside Europeana based on the most popular content reached or retrieved would be useful to the communities of both the Humanities and the Social Sciences.

- **More!:** As this tool focuses more on real-time collaboration and awareness in a co-located setting, the link with Europeana content and the eCloud project is less clear.

- **Annotation tools:**
  - **AnnotateIt, Textus, Pundit, OpenAnnotation, DocumentCloud, Researchr:** There was agreement that annotation is a key activity in Humanities and Social Sciences and that tools of this nature would be most relevant. However, specific tools often have usability issues for Humanities and Social Sciences researchers. Moreover, they do not transfer well to non-textual material.

4. Evaluation with Axiom Group

After the first round of development (which included regular formative evaluation sessions, following the User-Centered Design approach), new members of the Axiom philosophy group (who did not participate in the formative evaluation studies) were asked to participate in an online session to discuss, evaluate and provide their feedback for the above tools. The meeting took place on the 31st of January 2014.Their feedback is presented below.[4]

General discussion

ARIADNE Finder

The members of the Axiom philosophy group had the chance to look and reflect on the Ariadne Finder. The general idea behind the Finder seems very attractive to them. They also believe that the Finder will help them in their work. The problem they face when searching for (new) resources is that they often end up browsing a lot of unrelated results or have to use numerous different sites in order to be able to search different collections. As a result, they often face a duplication of search results. Having a personalized tool such as the ARIADNE Finder, embedded in their site to search different collections from one point of entrance, would provide them with the ability to reduce the time spent on searching and browsing. They also find it useful to have a uniform way of viewing the metadata of the results, regardless of the initial provider, and value having the metadata in a cleaned format.

The users provided feedback on a number of things that could be adjusted in the Finder to better cover their needs. The need to filter search results per year is very important for them, as well as a facet to filter results based on the author of the resource. These needs are related to the way that philosophers search for resources. They usually start from the work of an initial researcher (philosopher, mathematician, etc.) and then move to work or (secondary) sources related to this person. For this reason, they also asked if a way of prioritizing results could be implemented. As was discussed during the meeting, a facet for authors could provide a suitable way to prioritize results, but other ways to cover this need can also be explored.

Another comment received from the group is related to the type of resources they usually work with. As philosophers, they usually work with books, as opposed to images or audiovisual material, and would like to access books more easily. Hence, they asked whether the respective facet could be pre-enabled when making a search.

Regarding the content itself, the philosophers noted that they would like to able to search more collections and more providers, such as Google Scholar. They remarked that sometimes searches yielded a limited number of results. As was explained during the meeting, these limited results are due to the limited queries the ARIADNE Finder has used to harvest results from different providers. In the future, members of the Axiom Group will provide new queries to harvest and populate the repository.

Finally, they liked linking the ARAIDNE Finder to Wikipedia for immediate access to a philosopher's biography and asked for a small graphical change to get more search results per page.

[4] The results of this evaluation have also been included in the following paper: H. Van den Berg, G. Parra, A. Jentzch, A. Drakos, and E. Duval. "Studying the History of Philosophical Ideas: Supporting Research Discovery, Navigation, and Awareness". In: Proc. of i- Know '14. Graz, Austria: ACM, 2014.

TimeMapper

The philosophers took TimeMapper to be a valuable addition to their current set of tools. They study large collections of textual resources published in different historical periods, such as books on logic published from the 17th to the 20th century. They also study multiple editions of books, published at different times, in different languages, and in different places.

Identifying relevant content and learning about the existence of different types and editions of books is time-consuming and requires a high level of expertise. The TimeMapper provides the philosophers with a quick overview of which textual resources were published when and where, and thus allows them to quickly order and interpret these resources. The TimeMapper was taken to:

(i) support the identification of (novel) relevant content;
(ii) provide quick and easy access to important metadata (e.g., holding of a work, description of content);
(iii) possibly assist historical research by allowing he philosophers to identify and compare works published in the same period.

It was also noted that the tool could significantly benefit students in philosophy, who often have little to no knowledge of different (historical) textual resources.

The philosophers identified a number of features that would help to better cover their needs. They suggested a change of layout of the map, as they were uncertain whether visualizations of large amounts of data would be easily interpretable. They further would like to have immediate access to the ARIADNE Finder when working with TimeMapper, and would like to be able to select and visualize what they take to be important metadata (e.g., only metadata of books of one specific author). The latter might be achieved by linking the TimeMapper to faceted search results provided by the Finder. Finally, the philosophers want to compare timelines of works published by two or more authors over relatively long periods of time.

Activity Stream

After discussing the other tools, the Activity Stream functionality was demonstrated to the Axiom Group. Whereas search or visualization tools may be considered familiar to the members of the group, an awareness tool is a new and possibly interesting addition to their current toolset. In its current state, the Activity Stream captures and presents traces of searches conducted with the ARIADNE Finder and of visualized searches using the TimeMapper. The discussion and feedback therefore mainly concerned these activities. Based on the feedback received, the relevant features of this tool are taken to be:

(i) enhancing group awareness,
(ii) supporting direct collaboration among colleagues, and
(iii) supporting individual research.

Group Awareness: Members of the group remarked that the Activity Stream allows one to obtain an overview of each other's work. They thought this was useful, although they did note that having such an overview might be more relevant to the leader of the group than for junior researchers. Currently, the stream shows different daily activities of the researchers. It shows topics and the time taken to explore different ideas. The philosophers remarked that this overview provides information that might be worth to explore further or to discuss in the group. It was also remarked that the tool might help students to find unknown resources and to gain relevant contextual information regarding a topic.

Collaboration: The Axiom philosophers thought it was interesting to observe what other colleagues were searching and what results they obtained. Furthermore, they liked the possibility of building on research done by their colleagues, and to use search results that they did not think of themselves or that would have taken quite some time to compile.

Individual research: Regarding their individual work, the researchers saw the Activity Stream as an opportunity to save searches or visualizations without the need to always actively conduct these activities themselves.

In general, the philosophers liked the fact that the Activity Stream is presented in a separate screen. This ensures that the use of other tools (such as the Ariadne Finder) does not become more complex. They also look forward to connect other tools to the Activity Stream, so that multiple different kinds of information can be presented in the stream. Other functionalities they would like to see are:

- Ability to prioritize (or rank) and save successful search sessions. This will allow them to immediately continue their work at a later stage, without losing time by repeating previous actions.
- Possibility to search for activities based on a date.
- It would be interesting to include more information with the activities, such as the size of the result set or the different information sources used.

Other comments

They would like to see a way to save a search result, bookmark, and send through an email (annotation for search results).

Participants in this session

- Axiom Philosophers
  - Pauline van Wierst
  - Jeroen Smid
  - Dirk Gerrits
  - Hein van den Berg (also member of WP3)
- eCloud WP3 staff
  - Erik Duval
  - Gonzalo Para
  - Andreas Drakos
  - Anja Jentzsch

Publication:

Work done with the Axiom Group resulted in the following publication: H. Van den Berg, G. A. Parra, A. Jentzsch, A. Drakos, and E. Duval. Studying the history of philosophical ideas: supporting research discovery, navigation, and awareness. Proceedings of the 14th International Conference on Knowledge Management and Knowledge Technologies, pages 12:1–12:8, 2014.

5. Evaluation with Musicologists

As mentioned in previous sections, the mash-up of tools and services (Finder, TimeMapper and Activity Stream) developed for the community of philosophers from the Axiom Group was adapted and deployed to provide resources to musicologists working on early music. Besides these existing tools, two new tools (Aruspix and Music21) were added to the workflow, in order to support specific musicology research tasks: optical music recognition and the analysis of features of music scores.

In initial meetings, members of the research community of musicologists discussed with us the workflow, computational tools, and content.

Towards the end of Year 2, we organised more summative evaluations, some face-to-face and some virtual. They all took place between the 16th and 26th of January 2015. Below we present the feedback obtained.

General discussion

To start the discussion, the complete workflow of tools was presented to the musicologists. Afterwards, questions were asked regarding the usefulness of the current tool setup. In general, the participants agreed that the way in which the tools support the research process is helpful. The connection of existing tools (optical music recognition and processing of encoded scores) and automating the process of data sharing between these tools is of great value for them, as it saves them time with their research tasks, when compared with using the tools individually. Actually, some of the musicologists had not been able to manually feed the output of one tool as input to the next tool in the workflow.

While the participants find the overall workflow useful, they were also interested in details about specific parts of it. Some of them suggested that, in some cases, just one or two tools are more relevant for their research (e.g. converting a score into a computer readable format or importing their own encoded scores to process with Music21). This is mainly related to their technical background and research goals. Some of the participants are computational musicologists that use tools like Music21, while others are more traditional musicologists that work with the original prints.

The participants agree with the added value of the loosely integrated workflow while doing research on a single item (score), but also observed that the workflow could be automated for use at a larger scale (e.g. a large dataset of scores of an specific period or region). This process and the results could be of great value in order to answer research questions about a complete collection or in order to generate new questions for such a collection.

ARIADNE Finder, TimeMapper and Activity Stream

After the musicologists discussed the overall workflow and setup of tools, they were prompted to assess the tools on an individual level.

From the set of tools adapted from last year, the TimeMapper was considered the most interesting and relevant for their research. In its current form, the tool provides a visualization of scores based on location and year of print. The participants suggested extending the functionality of the tool, for example with the use of more information than just the data of publication of the prints (e.g. include the information gathered in the Music21 tools, like parallel fifths, valid melodies, or other species counterpoints of a score or measure) or the possibility to compare different timelines from different

search terms. This feedback basically affirms the relevance and usefulness of information visualisation techniques in general for their research work.

The Finder was mostly seen as a tool that provides existing functionality, similar to what other search engines provide, though the musicologists acknowledged the value of having facets to filter the result set. They suggested to personalize facets to terms that are closer to musicologist research practice, for example, to use 'printed books', 'manuscripts', 'single pieces' instead of 'image' or 'text' classification.

The musicologists were more critical about the usefulness of the Activity Stream (AS) in their research activities. They were not sure that the current actions are relevant for them or even which alternative kinds of activities might be useful to be displayed in the tool. They mostly perceived the AS as an interesting communication device or as a source of high-level information, comparable to a Social Network. The participants suggested functionality to enhance the perceived usefulness of the stream, such as a search for specific activities, the possibility to aggregate activities in order to obtain statistics from them, and the possibility to store results for later use.

Participants also suggested other interesting ways to connect the tools, instead of only having a linear approach, as we do now. For example, they mentioned that it would be interesting to be able to take the output of Music21 (e.g. parallel fifths of a score) and map the results, based on their location, with the TimeMapper. This can provide an overview of specific scores characteristics and relate them to a particular location.

Aruspix and Music21

While the Aruspix version included for Europeana Cloud does not have a visual frontend for the users, the musicologists acknowledge its importance in the workflow. As mentioned, optical music recognition (OMR) is a crucial step for them, in order to decide which research direction to take, such as bibliographic or music intervals research. Regarding the current output of this tool, the musicologists would appreciate to see the encoding result and the percentage of errors after the OMR process. While in other sciences, researchers are used to work with and accept a certain percentage of errors; these may not be well accepted in the musicology domain. Nevertheless, they appreciate what is happening behind the scenes and how good the obtained encoding is, and believe that the results could build trust from the user to the system. Moreover, information about errors can be used as a feedback mechanism for Aruspix (as simple as possible but at the same time complete enough to get the desired information).

The Music21 web interface was one of the most interesting tools for the musicologists. Besides the textual rendition of the analytical results, the participants would also like access to plots or statistics (e.g. note distribution), as these could be more helpful in order to identify characteristics of a score. Currently, the Music21 interface only supports a specific set of generic calculations and processes. The participants would like to have the freedom to build their own queries, via text or through a graphical user interface.

Other comments

During the session, the participants provided suggestions about the tools and the workflow, but also about the work done more in general by WP3. For example, some users suggested being able to push the generated encoded scores by Aruspix (MEI or MusicXML) back into the Europeana repository. This would allow sharing the results with peers. Also, while it was not the direct scope of our work, the participants suggested enhancing the usability of the tools and providing a nicer user interface.

Finally, the participants suggested additional tools or functionality to be considered. These are:

- Possibility to run batch processes to get a broader overview of music characteristics of a set of scores.
- Provide playback mechanisms in Music21 (or Aruspix) to be able to validate and confirm the automatic encoding.
- Possibility to annotate directly into the digital version of a score.
- Possibility to create their own visualizations based on the data obtained from different tools, especially from the Music21 output.
- Inclusion of additional musicology resources, for example from http://www.diamm.ac.uk/.

Participants in this session

- Musicologists
  - Frans Wiering (Utrecht University)
  - Reinier de Valk (City University London)
  - Eliane Fankhauser (Utrecht University)
  - Laurent Pugin (RISM)
  - Peter van Kranenburg (Meertens Institute - KNAW)
- Europeana Cloud staff
  - Gonzalo Parra
  - Marnix van Berchum (KNAW-DANS / Utrecht University)

Publication

The set of tools we built for the musicologists and the findings from our evaluation resulted in a publication at ISMIR2015 [Duval, van Berchum, Jentzsch, Parra, Drakos. 2015].

6. Evaluation with the Europeana Newspaper Archive

Similar to the Finder and TimeMapper services of previous sections, an interactive visualisation tool was developed that facilitates searching across different (meta-)data. The Newspaper Exploration Environment supports text search, time-range filtering, language filtering (as opposed to location, as this meta-data is not available in the Newspaper Archive except for contributor location) and newspaper filtering. More detailed information can be found in D3.2.

Two workshops were organised to get a better understanding of the researcher workflow and their concerns with the Europeana Newspaper archive. The first full-day workshop was organised in Amsterdam, 25th May 2015. The second one-hour workshop was organised during the Europeana Cloud Plenary Meeting in Edinburgh. The Newspaper Exploration Environment was developed and evaluated regarding usability during the summer in face-to-face sessions with 20 people.

Towards the end of Year 3, the tool was evaluated online with five expert users. Below we present the feedback obtained.

Workshop results

Twenty-four people participated in the full-day workshop in Amsterdam, 25th May 2015. All participants were familiar with the Europeana Newspaper archive. After a briefing on the development of personas and scenarios, participants were split up in groups to come up with realistic scenarios for using the Europeana Newspaper archive. Afterwards, a discussion was held to find the common themes among their scenarios and way of work.

The personas and scenarios can be found in D3.1. The common themes resulting from the discussions are the following:

- **Geo-visualisation: results displayed on maps**
- **Temporal visualisation: representing items on a timeline**
- **Serendipity through exploration**Social functions: chat functionalities, annotations for peers
- Feedback features: providing new metadata, corrections to the Newspaper archiveLanguage specific search
- Export functionality
- Sentiment analysis

The second workshop was a one-hour session during the Europeana Cloud Plenary meeting in Edinburgh, UK. Twelve participants were presented with the current state of the Europeana Newspaper Archive and API to help gather ideas on features necessary to build the tools to support Digital Humanities researchers in their work. A summary of ideas is listed below:

- **Geo-visualisation**
- **Temporal visualisation**
- **Serendipity**
- Metadata: keywords, text-type (advertisement, official proclamations), image recognition
- Export functionality
- Support for changes in language over time
- Translation support / multi-lingual search
- Text view vs scanned view
- Weighted search regarding headlines, page number,…
- OCR accuracy
- Semantic analysis

- Crowdsourcing for improvement of metadata
- N-gram
- Relevant articles

This list resulted in feature requests for both the system (Newspaper Archive API, e.g. multi-lingual search), the dataset (the collection of digitized newspapers, e.g. the addition of extra metadata) and the user experience (accessing the data, e.g. visualising the data in a meaningful way. The remainder of this section focuses on user experience.

Based on the feedback from the workshops, the themes serendipity, and geo- and temporal visualisations were noted to be of great interest to DH researchers. These were chosen as the main features of the Newspaper Exploration Environment (NEE). To prevent the limitation of the tool to these features, a modular approach was used: NEE is a widget-based system that facilitates the addition and removal of features depending on the researcher's need. Future development of more widgets can add support for e.g. OCR accuracy visualisation, sentiment analysis.

These widgets form a Coordinated Multiple View system: it provides different views on the same data at once [Roberts. 2007], allows for complex queries across multiple data dimensions, and helps retain a sense of context [Alsallakh, Miksch, & Rauber. 2014]. The scalability of the number of visualisation widgets is limited to screen-size. Large displays such as large wall displays and interactive tabletops [Nagel, Maitan, Duval, Vande Moere, Klerkx, Kloeckl, & Ratti, 2014] can solve this issue but are not always available. With tablets becoming ubiquitous, combining multiple devices to replicate a single, large faceted search interface becomes a more feasible scenario [Rädle, Jetter, Marquardt, Reiterer, & Rogers, 2014]. We therefore looked at the feasibility of deploying a modular, extendable Newspaper Exploration visualisation on multiple tablets.

General discussion

The Newspaper Exploration Environment was evaluated for two different purposes. The first evaluation looked at the difference in efficiency and user experience between using a large interactive tabletop and four tablet devices to search and visualise the Newspaper Archive data. This helped us understand whether or not tablets could be a good substitute when devices such as large interactive tabletops are not available. The second evaluation focused on the usefulness of the application for searching and exploring the Newspaper Archive data. The following sections detail the results of these evaluations.

Evaluation results regarding usability of interactive tabletop vs multiple tablets

The main focus of the evaluation was to observe the difference in search performance and user perception between a single and multi-device faceted search system. We followed a within-subject experiment design, with the order of systems (whether the participant used the single or multi-display system first) fully counterbalanced. Participants each performed two sets of five tasks on both the single screen and multi-screen setup. Before each set of tasks, the current system was explained to the participant, after which she or he received hands-on time until the participant felt familiar enough to proceed to the tasks. Facet widgets for the multi-device setup were placed at the same positions as the single device setup. This was then followed by a questionnaire on user perception, a SUS questionnaire per system, and a semi-structured interview.

The device for the single screen task consisted of a 42" custom-built interactive tabletop with multi-touch capabilities, running at a 1920x1080 resolution, displaying the 4 views simultaneously. The multiple screen setup consisted of four tablets (3 iPads Air 2 and 1 iPad Mini 2). Each tablet visualised one of the four widgets.

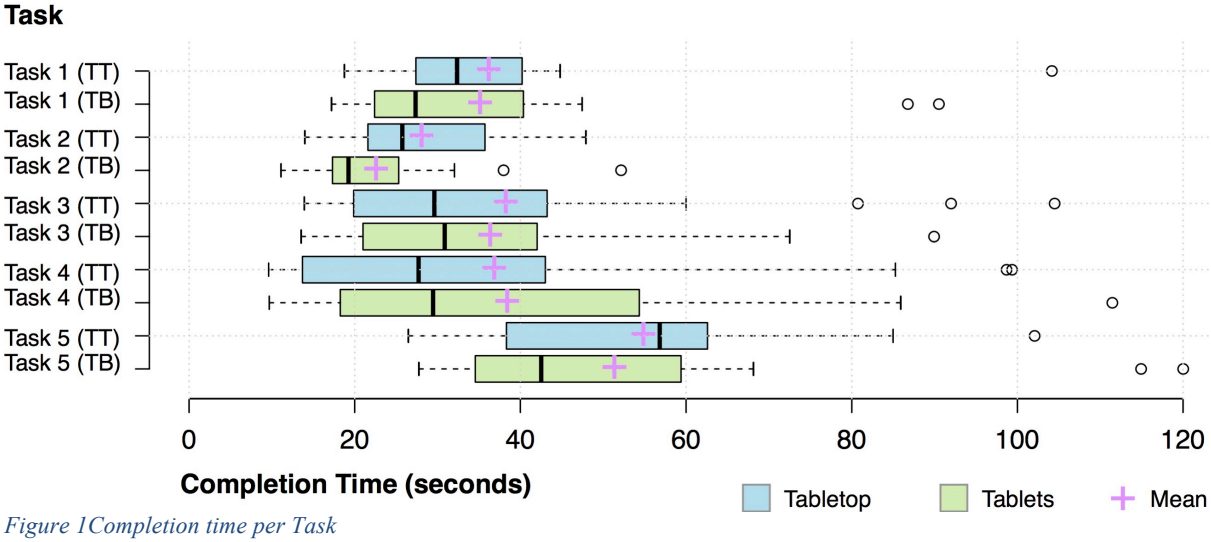

Figure 1Completion time per Task

The evaluation comprised 22 users (6 F, mean age: 34 years, SD:12).

Figure 1 shows the completion time per question for each setup. The difference in time for each task reflects the difficulty of each question. The box-plots show a slight trend for better time-to-task results for tasks 1, 2 and 5, with a lower average time-to-task and lower 25% and 75% quartiles. Task 3 shows a smaller mid-spread for the tablets with a slightly higher median. The 25% and 75% quartile of task 4 are lower for tabletop. Overall there is no significant difference in time-to-task across tasks.

SUS (System Usability Score)[5] questionnaires resulted in similar perceived usability as the tabletop scores (M=82.36,SD=11.83) and tablets (M=80.25,SD=12.87). We can therefore conclude that perceived usability is not an issue with either system. Four participants mentioned advantages for multiple devices as opposed to the tabletop for faceted search ("they facilitate working on individual facets better", "I am more tempted to focus on 1 facet", "I can pick one up to dig deeper in one facet"). Eight participants mentioned portability as another advantage ("I can set this up at home", "I can move more freely around the workspace", "I can take this with me").

Evaluation results with expert users of the online version of NEE

For the purpose of this evaluation, the geo-visualisation was replaced by a language visualisation for the final evaluations, due to the current lack of geo-data relating to historical newspapers. The location metadata is currently limited to the location of the contributors of the different European libraries.

The online evaluation consisted of two options: a short evaluation where the participant watched a video explaining the tool and filled out a questionnaire, and a more elaborate evaluation where

[5] https://en.wikipedia.org/wiki/System_usability_scale

participants used the tool for 15-30 minutes and filled out the same questionnaire with extra questions regarding usability.

Eight people participated, 4 for each evaluation (age 26-53, 5 female). 4 participants were DH researchers, 2 lecturers, 1 manager and 1 developer.

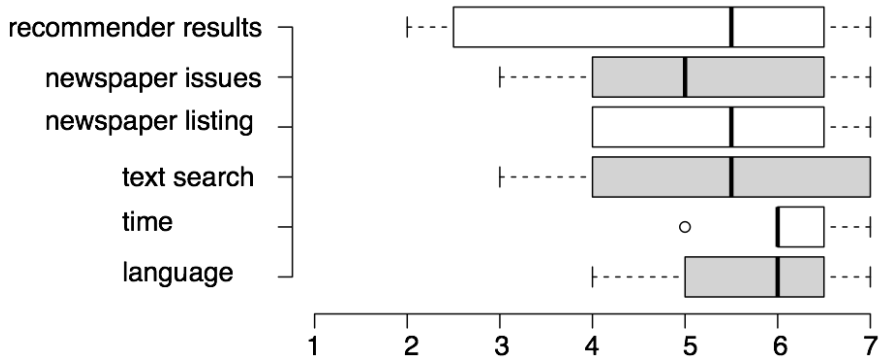

Figure 2 Usefulness of features (1 - Not useful at all, 7 - Extremely useful)

The system was in general regarded very useful (see Figure 2). Language and time visualisation rated highest, followed by text search, newspaper listing and newspaper issues. Two of the eight did not find the recommended results useful.

Table 1 Questionnaire results, 1- Completely disagree to 7- Completely agree

Question	Mean	Standard Deviation
I have the knowledge necessary to use the system	5.5	1.9
I enjoyed working with the system	5.5	1.2
Using the system makes work more interesting	5.5	1.3
Using the system is fun	5.1	1.5
I prefer the filtering approach	5.75	0.9
I prefer the visual approach	5.3	0.7
I prefer the recommendation approach	4.3	1.9

The results of the questionnaires (7-Likert scale, 1 – Completely disagree, 7 Completely agree) are shown in Table 1: participants were confident they had the knowledge necessary to use the system and enjoyed working with it. They also considered it makes work more interesting and is fun . Participants stated they preferred the filtering approach , followed by a visual approach. There is no clear result on whether people prefer a recommendation approach.

Usability received a SUS score of 79 (SD=14): the system was considered consistent and easy to learn.

From the qualitative feedback we earned that the system would be useful for trans-national research to figure out how different countries capture and present different events. One participant mentioned it would without doubt help with explorative aspects and presents a good overview what is in the collection. There was however a lack of trust from 3 participants, regarding the quality and correctness of the results, which can only be as correct as the accuracy of the OCR permits. One user suggested to make it more evident that the results presented were an estimate. Other methods could include presenting estimates of accuracy based on the OCR quality of the presented results.

7. Evaluation with Agricultural Researchers

As mentioned in previous sections, in order to provide resources to agricultural researchers working on various thematics of agricultural economics, the AGRERI Discovery Microsite was added to the workflow, supporting specific research tasks: how to bring content from relevant Europeana resources as well as relevant resources from other external sources, like the AgEcon repository[6] and the FAO AGRIS[7].

In initial meetings, members of the research community of agricultural researchers of AGRERI discussed with us their workflow, problems and data requirements. .

Several face-to-face meetings and an online survey took place between July 2015 and November 2015, in order to collect user requirements, feedback on the AGRERI Discovery Microsite and also evaluation on the final version of the microsite, in the form of a hands-on workshop that took place on January 8th, 2016.

General discussion

The members of the AGRERI team had the chance to discuss and reflect on the AGRERI Discovery Microsite. They were very interested in the potential behind the microsite which could potentially facilitate them with their work. The problem they face when searching for resources, as in the case of researchers in almost all disciplines, is that they often end up browsing a lot of unrelated results or have to use numerous different sites in order to be able to search different collections. As a result, they often face a duplication of search results. Having a personalized tool such as the AGRERI Discovery Microsite, embedded in their site to search different collections from one point of entrance, would provide them with the ability to reduce the time spent on searching and browsing. They also find it useful to have a uniform way of viewing the metadata of the results, regardless of the initial provider, and value having the metadata in a cleaned format.

Workshop results

The workshop that was organised had the objective to provide a hands-on evaluation of the AGRERI Discovery Microsite not only by the agricultural researchers of the AGRERI team but also with real users interested in the tool. These users belong to different user categories than the five agricultural researchers of the AGRERI team mentioned below, such as agricultural advisors, students of agricultural sciences and farmers (who wished to remain anonymous). Fifteen people in total participated in the workshop that took place on the 8th of January 2016.

All participants were familiar with the AGRERI Discovery Microsite, but from different perspectives. After a briefing on the different personas, scenarios and use cases, participants were asked to use the discovery microsite with specific needs in mind, and later a discussion took place in order to evaluate the discovery microsite in detail.

The users provided feedback on a number of things that could be adjusted to better cover their needs. Feedback was received on whether the Europeana collections were useful to users. Not all

[6] http://ageconsearch.umn.edu/
[7] http://agris.fao.org/

participants were familiar with the Europeana collections, but all users liked the idea of having a microsite where external sources can be included with only one query.

The need to filter search results per thematic is very important to them, as well as to filter search results per year. Also, a facet which would enable them to filter results based on the author of the resource could also be of use. Also, something that they think is quite essential in a discovery microsite is to be able to get relevant resources like the one they were searching for (in the same thematic but also based on keywords), and these relevant resources to be presented beneath the resource selected.

Another comment that was made by the users is related to the type of resources that they are searching for. While most of them are typically searching for publications, they would also like to be able to get results of different content type (such as video, image etc.) which would provide them with a better picture of the available resources that exist in the various external sources.

Of course, the idea to be able to browse through even more collections and more providers in the future is quite appealing to all users.

Participants in this session

- Dr Constantine Iliopoulos (AGRERI)
- Dr Irene Tzouramani (AGRERI)
- Dr Irini Theodorakopoulou (AGRERI)
- Magdalene Bardounioti (AGRERI)
- Mimika Anggelopoulou (AGRERI)
- 10 users

8. Conclusion

The work of WP3 was structured around yearly cycles. In the first year, we focused on collaboration with the Axiom group of philosophers and, to a much lesser extent, the DM2E project. In the second year, we worked with a group of musicologists who focus on early music. In the third year we worked with researchers from the Digitial Humantities community and researchers from the Agriculture community.

D3.1 reports on the development of personas, scenarios and use cases. D3.2 summarises the software that resulted from our development and integration work. This deliverable summarises the evaluations that took place throughout the project and in specific evaluation sessions.

To conclude, the user-centered development process worked as intended. The final set of tools was positively evaluated by the intended users. An important issue for the last cycle was to connect the frontend tools for researchers with the backend infrastructure of Europeana Cloud, so that we can work with more comprehensive content collections.

Both, the results from Year 1 and Year 2, have been published to well-known conferences in the corresponding area and we hope to achieve the same for our work on the Year 3 scenario.

9. References

[Abras, Maloney-Krichmar & Preece, 2004] C. Abras, D. Maloney-Krichmar, and J. Preece. User-centered design. Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications, 37(4):445–56, 2004.
[Van den Berg, Parra & al., 2014] H. Van den Berg, G. A. Parra, A. Jentzsch, A. Drakos, and E. Duval. Studying the history of philosophical ideas: supporting research discovery, navigation, and awareness. Proceedings of the 14th International Conference on Knowledge Management and Knowledge Technologies, pages 12:1–12:8, 2014.
[Parra, Klerkx & Duval, 2013] G. Parra, J. Klerkx, and E. Duval. Tinyarm: Awareness of research papers in a community of prac- tice. In Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies, page 21. ACM, 2013.Shneiderman, B. 2008. Science 2.0. Science. Vol. 319, No. 5868, 1349-1350.
[Cuthbert & Ariza, 2010] Cuthbert, M. S., Ariza, C. Music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data. In 11th International Society for Music Information Retrieval Conference (ISMIR 2010), August 9-13, 2010, Utrecht, Netherlands. pp. 637-642.
[Haus & Longari, 2005] G. Haus and M. Longari. A multi-layered, Time-based music description approach based on XML. Computer Music Journal, 29(1):70–85, 2005.
[Pugin, Burgoyne & Fujinaga, 2007] L. Pugin, J. A. Burgoyne, and I. Fujinaga. Goal-directed evaluation for the improvement of optical music recognition on early music prints. In Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries, pages 303–304. ACM, 2007.
[Alsallakh, Miksch, & Rauber. 2014] Bilal Alsallakh, Silvia Miksch, and Andreas Rauber. 2014. Towards a Visualization of Multi-faceted Search Results. In Proceedings of the DL2014 Workshop on Knowledge Maps and Information Retrieval (KMIR), the ACM/IEEE Joint Conference on Digital Libraries. 4.
[Roberts. 2007] Jonathan C Roberts. 2007. State of the art: Coordinated & multiple views in exploratory visualization. In Coordinated and Multiple Views in Exploratory Visualization, 2007. CMV'07. Fifth International Conference on. IEEE, 61–71.
[Nagel, Maitan, Duval, Vande Moere, Klerkx, Kloeckl, & Ratti, 2014]. Till Nagel, Martina Maitan, Erik Duval, Andrew Vande Moere, Joris Klerkx, Kristian Kloeckl, and Carlo Ratti. 2014. Touching transport-a case study on visualizing metropolitan public transit on interactive tabletops. In Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces. ACM, 281–288.
[Rädle, Jetter, Marquardt, Reiterer, & Rogers, 2014] Roman Rädle, Hans-Christian Jetter, Nicolai Marquardt, Harald Reiterer, and Yvonne Rogers. 2014. HuddleLamp: Spatially-aware mobile displays for ad-hoc around-the-table collaboration. In Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces. ACM, 45–54.
[Duval, van Berchum, Jentzsch, Parra, Drakos. 2015] Erik Duval, Marnix van Berchum, Anja Jentzsch, Gonzalo Alberto Parra Chico, Andreas Drakos: Musicology of Early Music with Europeana Tools and Services. ISMIR 2015: 632-638.