User Required? On the Value of User Research in the Digital Humanities

Max Kemman
University of Luxembourg
www.maxkemman.nl
max.kemman@uni.lu

Martijn Kleppe
Erasmus University Rotterdam
www.martijnkleppe.nl
kleppe@eshcc.eur.nl

Abstract

Although computational tools play an increasingly important role in the humanities, adoption of tools by scholars does not always reach its potential. One approach to this problem is user research to uncover the needs of the users. However, it is uncertain whether such user requirements can be generalized to a wider group of humanities scholars, and whether users are able to explicate their requirements for methodological innovation. We ask what the role of user research is in the Digital Humanities by discussing gathered user requirements for two projects. We categorized the requirements as within- or out-of-scope of the projects’ goals, and found a tension between the specificity of humanities’ research methods, and generalizability for a broader applicable tool. With the out-of-scope requirements we are able to map the wider research workflow, showing DH tools will most likely take a spot in the wider workflow, and that it is infeasible to create a tool for the entire workflow that is generic enough for a larger user group. However, the within-scope requirements led to features that were sufficiently generic for the tool to be adopted, also for unintended purposes. These insights show user research has a clear benefit for DH projects.

1 Introduction

The development of tools plays an important role in the Digital Humanities. With the increasing quantities of digitised as well as born-digital source material, computational tools have become necessary for exploring, analysing and enriching this material. While many tools have been and are being developed, adoption by the target audience, i.e., humanities scholars, does not always reach its potential (Edwards, 2012; Gibbs & Owens, 2012; Warwick et al., 2007). In projects where the research data is published within a tool, this can result in neither the tool nor the research data being fully used by other scholars. One partial solution to this problem is to publish research data separately from the tool, as advocated by Borgman (2012), and Kansa et al. (2010).

Furthermore, in order to create tools that will be adopted by scholars, development should take into account the practices and conventions adhered to in subdisciplines of the humanities (Bradley, 2005; Kemman et al., 2014b). One approach is to focus on the users, actively involving them during development and evaluation of designs, known as user-centred (systems) design (Gulliksen et al., 2003). To achieve this, user research is performed (Warwick, 2012), for which one of the tasks is to uncover the needs and wishes of the user group, commonly referred to as user requirements (e.g., Sweetnam et al., 2012).

There is however an ongoing debate whether such user requirements can be sufficiently generalized to a wider group of humanities scholars. On one end of this debate, we see the suggestion that research contains generic tasks called scholarly primitives, defined as “basic functions common to scholarly ac-
Activities across disciplines, over time, and independent of theoretical orientation” (Unsworth, 2000). Unsworth presented a non-exhaustive list of primitives, summarized by Martin Weller as follows (Weller, 2011):

1. discovering – knowledge either through archives or research;
2. annotating – adding layers of interpretation;
3. comparing – for example, texts across languages, data sets;
4. referring – referencing and acknowledging;
5. sampling – selecting appropriate samples;
6. illustrating – clarifying, elucidating, explaining; and
7. representing – publishing or communicating.

Tools can be developed to support these primitives, and are thus applicable to a broad community of scholars. For the development of an infrastructure for the arts and humanities, the scholarly primitives have been combined into discovering, collecting, comparing, and delivering (Anderson et al., 2010; Blanke & Hedges, 2013). The idea is to create a user-centric infrastructure to support the entire research process with primary source material.

At the other end of the debate, we see the suggestion that scholarly practices are very specific and that a “one size fits all’ approach would be a disastrous underestimation of the specific needs of humanities research” (van Zundert, 2012). Van Zundert suggests that insofar methodological innovation is desired, generalization and standardization might be detrimental.

Whether user research enables targeted users to explicate their requirements for methodological innovation is furthermore met with scepticism in literature. Although interviews are regularly used as a method for gathering user requirements (Benyon et al., 2005), users supposedly do not know what they want, and cannot predict their own future behaviour (Nielsen, 2001). Moreover, innovation is said to be driven by focusing on new technology, even though people do not yet need such technology, nor have a clear use case for it (Norman, 2010). Nevertheless, in the wider Human-Computer Interaction literature, user research is regarded crucial during development (Hofmann & Lehner, 2001). Following from the above discussion, we ask what the role of user research is in the Digital Humanities. Our research question is: what is the added value of user research for developing tools aimed at digital research methods?

To address this question, we will discuss results from user research for gathering user requirements for two Digital Humanities projects we coordinated; PoliMedia and Oral History Today. In these projects, we held interviews with scholars to inform development. We will show user requirements that were within- or out-of-scope, where the scope is determined by feasibility and the project goal, and examine how many user requirements were common to multiple participants. By doing so we aim to provide insight into the added value of user research for these two case studies.

This paper is structured as follows: first, we will introduce the research projects and their goals. Second, we will explain how scholars were involved in these projects to voice their needs and wishes. Third, we will review the user requirements that were collected and whether these were determined to be within- or out-of-scope. Fourth, we will discuss how our findings relate to the literature. Finally, we will discuss what we learned from the user requirements, and what the added value was of user research.
2 User requirements for PoliMedia and Oral History Today

The PoliMedia project\(^1\) aimed to facilitate a digital research method for large-scale cross-media analysis of the coverage of political debates (Kleppe et al., 2014). Investigating how political debates are covered in the media required scholars to explore three distinct collections: 1) the minutes of the Dutch parliament, 2) Dutch newspapers and 3) Dutch radio bulletins. Additionally, a fourth dataset of interest is the Dutch television broadcasts, but due to a lack of links found between the proceedings and television broadcasts, this dataset was dropped from implementation, although it was included in the interviews. In order to present a dataset with as much overlap of these three collections as possible, we set the timeframe from 1945-1995.

Although access to the collections has already improved with digitization, each collection still required scholars to learn and use three different user interfaces, as well as redo searches for the same subject in each system.\(^2\) To better facilitate such research, PoliMedia provides a search user interface where scholars can explore the minutes of the Dutch parliament with integrated links to media coverage, see figure 1.

For each speech in the parliament, information was extracted to represent the speech; the speaker, the date, important terms (i.e., named entities) from its content and important terms from the description of the debate wherein this speech was held. This information was then used to query the archives of the newspapers and radio bulletins, and links were created to items that correspond to the query (Juric et al., 2013). The debates and links were then represented as RDF, a Semantic Web standard (Juric et al., 2012). By employing Semantic Web technology, information about entities (such as people, places, subjects) can be aggregated from multiple collections to gain a broader perspective. The scope of the project could thus be described as follows: automatically creating links between debates of the Dutch parliament to media items, made available in a search user interface in which debates of the Dutch parliament can be explored.

The Oral History Today project\(^3\) aimed at facilitating a digital research method for exploring and searching of aggregated, heterogeneous oral history content (Kemman et al., 2014b). Discovering interesting oral history interviews is a difficult task, as many small collections are available at many different locations: sometimes digitized, sometimes annotated by archivists, and sometimes available through an online portal. To better facilitate this process, Oral History Today provides a search user interface where scholars can search through over fifty oral history collections containing over four thousand interviews, enabling scholars to discover interviews across several collections, see figure 2. The collections were aggregated in a previous project (Ordelman & De Jong, 2011), and are hosted by DANS (DANS, 2012), where the collections were annotated to fit this archive’s schema. The metadata was then indexed and made searchable through a search user interface with a focus on usability. Since Google is immensely

---

\(^1\) [http://www.polimedia.nl](http://www.polimedia.nl)

\(^2\) Shortly after the PoliMedia project, the Dutch National Library launched a new search system that integrates the newspapers and radio bulletins, Delpher ([http://www.delpher.nl](http://www.delpher.nl)).

\(^3\) [http://zoeken.verteldverleden.org](http://zoeken.verteldverleden.org)
popular among scholars (Kemman et al., 2014a), the search system was designed to be like ‘a Google for oral history interviews’, i.e. the system would provide a simple search bar and a high recall of results ranked by relevance. This was extended with several filtering and ranking features. The scope of the project could thus be described as follows: a search user interface similar to Google but including advanced filter options, in which oral history interviews and collections can be searched and explored to discover topics across a multitude of collections.

![Figure 2: Oral History Today user interface. Left: search results page. Right: interview page.](image)

3 Methods

In the PoliMedia project, before development commenced, we held semi-structured face-to-face interviews with five scholars. Interviewees were invited from our own network and represented both qualitative and quantitative methods. The interviewees worked at different universities. There were no further selection criteria regarding demographics. One interview was with two scholars simultaneously, and is treated as a single interviewee, thus leading to four interviewees in our data. Interviewees talked about their research questions, methods and requirements for cross-media analyses. Questions were specifically related to their general research problems and approaches, which databases and search engines scholars used, what they liked or disliked about these, and asking feedback on a verbal description of the PoliMedia plans.

In the Oral History Today project, we held semi-structured interviews with fifteen scholars via Skype. Interviewees were selected from our own network as well as via the oral history working group of the Dutch Research Institute and Graduate School for Cultural History. We selected interviewees in all stages of careers from project assistants to PhD Candidates to Professors. There were no selection criteria regarding other demographics. All interviewees were given a monetary reward for their participation. Interviewees talked about their research questions, methods and requirements for a federated search engine for oral history collections. Questions were specifically related to how they performed Oral History research, which collections they used, and asking feedback on a rudimentary search user interface that was created before the interviews, particularly regarding their first actions in the interface, how they explored collections, how they did more directed searches, and how they evaluated interviews. Interviewees were mainly knowledgeable in employing the oral history method; less than half of the interviewees created or reused oral history collections.

After each interview, the interviewer summarized this information into functional requests, which was then sent back via e-mail to the interviewee for approval, allowing edits where needed. These functional requests were then categorized into user requirements by the interviewer, where similar statements were combined. These user requirements were finally discussed by the project team to classify them as within- or out-of-scope, determined by feasibility and the project goal. The within-scope requirements were then prioritized for development.

For PoliMedia, after developing the user interface, 24 scholars evaluated the usability of the portal (Kemman et al., 2013). Feedback voiced during this evaluation led to an improved final version of the search interface.

4 http://www.huizingainstituut.nl/werkgroep-oral-history/
For Oral History Today, after an update of the search interface, five scholars were interviewed via Skype to explore the collections, try search questions of their own interest and provide feedback. The results of these evaluations were then considered for the next update; we repeated this process a second time leading to the final version of the search interface.

In this paper we report the user requirements that we gathered and classified for the first round of interviews for both projects.

4 Results

4.1 PoliMedia

The interviews for PoliMedia led to 39 user requirements. A total of 21 requirements were deemed within-scope of the project, and were related to functionality such as:

- gaining insight into contextual information (e.g., Function of actors, Party of actors, or Type of programme (news, talk show, late night, etc.)),
- the frequency of terms (e.g., Mathematical queries, Frequency of searched, related, and important terms in documents, Comparing/sorting search results by frequency of terms),
- search operators (Boolean operators and Google search operators (esp. the combination of a string with quotation marks)), and
- analysis of the debates (e.g., Length of document per actor, Ability to export non-formatted text).

18 requirements were deemed out-of-scope. These requirements were related to computational analysis of the sources with advanced techniques:

- image processing of newspaper pages (e.g., Size of headers, Number of columns on a page, Presence and size of photographs),
- audio-visual processing of television programmes (e.g., Length of talk, Presence of music, Use of filming techniques), and
- linguistic analyses of debates (e.g., Speech functions, Type of speech fragments (interruptions, questions, jokes, etc.)) and of newspaper articles (Genre (report, comment, letter to the editor, etc.)).

The project scope, as described above, focused on creating links between collections, and developing a search user interface to explore the proceedings to which media items are linked. The computational analysis of these items then is clearly out-of-scope. Moreover, such tasks are far from trivial considering the size of the collections: eight million pages from newspapers, 1.8 million radio bulletins (Delpher, n.d.), 2.4 million pages of parliamentary proceedings (Staten-Generaal Digitaal, n.d.), and 2500 hours of television material (Academia.nl, n.d.). Finally, computer vision tasks such as the classification of filming techniques are research problems not yet solved.

27 requirements were unique, i.e., voiced by a single interviewee. The most common requirements were the inclusion of Media output about subject before debate, Names of actors (people) involved, and Location in the newspaper (page number, location on page), each mentioned by three interviewees. The first two were deemed within-scope, while the third was deemed out-of-scope due to required image processing as described above.

Some user requirements that we had not considered before the interviews, but that were considered within scope and made a big impact on our thinking about the tool:

• Function of actors (e.g., minister, member of parliament, but also show host, interviewer, etc.) – voiced in one interview.

• Party of actors (e.g., VVD, PvdA, but also Greenpeace or other lobby groups) – voiced in two interviews

• Media output about subject before debate – voiced in three interviews.

The first two requirements could be addressed without too much difficulty, since this information was already part of the dataset. Making this information available at the front-end for interaction introduced the opportunity to explore the proceedings on the level of the speaker’s role (in our implementation, as member of government or of parliament). The third requirement introduced a different perspective on the interaction between politics and media than was envisioned. Not only do newspapers report on what happens in parliament, parliament discusses events in society by referring to newspaper reports. Newspaper articles regularly set the stage for parliamentary debates. Unfortunately, due to technical reasons, it was ultimately not implemented.

4.2 Oral History Today

The interviews for Oral History Today led to 75 user requirements. A total of 33 user requirements were deemed within-scope of the project, and were related to:

• more instructions and clearer details of functionality and collections (Support page describing interviews and search technology, Description of project (within which collection was created) and how collection came to be, Organization behind collection (management/creation)),

• more advanced searching with filters (e.g., Locations, Collection, Topics, Year Event, Access conditions),

• navigation within the search user interface (e.g., Navigate from interview to interview collection, Clicking a topic should result in all interviews with the same topic, Links between related interviews), and

• workspaces (Search trail (i.e., a history of queries), Bookmark functionality for interview).

42 user requirements were deemed out-of-scope. These requirements were mainly related to:

• features of the search technology (e.g., Boolean operators, Search explicitly for broad or narrow terms, Detect synonyms of search terms), and

• additional metadata on the interviewee (e.g., Age/Year of birth, Gender, Religion, Community of experience, Social class), interviewer (e.g., Age, Gender) and the interview (e.g., Research question underlying interview, Location of interview, Description of interview per 10 minutes).

On first sight, such requirements might appear well within the scope of the project. The decision to categorize them as out-of-scope was mainly due to limitations of what we had available. The requests regarding search technology were dependent on the search technology provider we had chosen before the start of the project. The search technology that we used focused on high recall with relevance ranking, i.e., adding more search terms broadened the result set but improved the search results ranking. This conflicted with the search behaviour we observed from the interviewees who aimed at reducing the search result set until it became a manageable set that could be assessed interview by interview. This wish for precision is also reflected in the wish for more metadata to assess relevance and the broader context of the interview. However, since we used a dataset created in a previous project, we could only use the metadata that was made available then. We cannot provide information we do not have ourselves, and enriching the metadata was out-of-scope.

34 user requirements were unique, i.e., voiced by a single interviewee. The most common requirement was a filter for Year event, voiced by ten interviewees, and was deemed within-scope.
Some user requirements that we had not considered before the interviews, but that were considered within scope and made a big impact on the final tool were the following:

- **Description of project** (within which collection was created) and how collection came to be – voiced in five interviews.
- **Organization behind collection** (management/creation) – voiced in one interview.
- **Distinguish facets between relating to content or general** (where content relates to the contents of the oral history interviews, e.g., Year or Location, while general is about the interview files, e.g., Open Access, or Audio/Video) – voiced in one interview.

What is interesting about the first two requirements is how oral history interviews are understood within the context of their collection. While we started with the assumption of a keyword search bar, we learned that a significant portion of the interviewees wanted to browse and view the interviews in the context of their collections. Our observations showed that while half the interviewees (8/15) started by typing interesting terms into the search bar, the other half (7/15) started by browsing the collections. Knowing this, we introduced more fine-grained exploration of the collections, and navigation controls to move from an interview to a collection page. The third requirement described gave us input to further consider the search filters we provided; what type of filter is this, and how should the search filters thus be presented? Considering such questions ultimately led to a better search filter interface than we initially provided.

## 5 Discussion

What can we learn from the gathered user requirements? The user requirements show that our users, the humanities scholars, are very aware of what they want, agreeing with the findings of Warwick (2012). In PoliMedia many user requirements reflect the research methods of the interviewed scholars, who would like their heuristic process simplified, i.e., the discovery of primary and secondary sources for investigation. Automatic analysis was perceived as helpful for this process to easily discover e.g., debate sentiments, framing of topics by media, and topic importance. In Oral History Today the requirements reflect the fine-grained control oral historians desire during their heuristic process: being able to find interviews related to a specific place, time and event. Additionally, insight into the background of both the interviewee and interviewer is desired to properly understand the interview.

Still, to some extent, our results agree with the criticisms of asking users about their requirements (Nielsen, 2001; Norman, 2010). First, in the case of PoliMedia only three uniquely voiced user requirements, out of 39, were related to the project’s technological goal of linking debates and media items and publishing these as RDF:

- External linking to databases about persons (e.g., www.parlement.com).
- Function of actors (e.g., minister of defence, member of parliament, show host, interviewer).
- Search on committee.

Second, in the case of Oral History Today the user requirements are based on current, rather than future practices, and even show a distrust of potential innovations. The idea of a simple Google-like search bar and high recall ranked by relevance did not appear to match the desire for high precision resulting in manageable sets, reminiscent of the “perfect thirty-item” online search identified by Bates (1984). Interviewees explained they could not trust the search ranking in a way to be confident search results further down the list would not have to be looked into, regardless of the performance of the ranking mechanism. This seems to show a tension between a need for completeness of search results, while at the same time keeping the number of search results manageable. Potential innovations in the discovery of oral history interviews are deemed undesirable, despite the proven utility of other search engines with high recall ranked by relevance (cf. Kemman et al., 2014a).
Finally, our results agree to a large extent with the suggestion that the humanities are too specific for generic innovations (van Zundert, 2012). A large number of user requirements we found were unique, underscoring the specificity of humanities research.

6 Conclusions

What is the added value of user research for the development of tools aimed at digital research methods? In our investigation of user requirements for two Digital Humanities projects, we found scholars have a clear idea how they perform their research, and how tools could simplify some steps in the process of discovering and analysing sources. On the other hand, we hardly see scholars immediately embrace the full potential of the projects’ goals in their user requirements: i.e., semantic web technology in the case of PoliMedia, and simple Google-like searching in the case of Oral History Today. Whether this means that scholars are unaware of how such facilities might help them, or whether scholars are aware that such goals do not match with their methods, remains an open question. To answer this question requires a deeper understanding of how (digital) technology is adopted by scholars. A study of how historians adopt digital technology and how it affects their practices is the topic of PhD research by Kemman currently in progress.6

Alternatively, perhaps the scope we chose was already too much tied to specific requirements dependent of the researcher, i.e., the linking between different collections is perhaps already a specific rather than a generic research method.

The findings and such questions seem to confirm the criticism that interviewing users for their requirements might not be the most effective method to advance methodological innovations. Instead, alternative approaches such as observations might give more insights into practices. Another promising approach is to move beyond the list of user requirements, and emphasize participatory design as a negotiation between users and developers (Muller, 2003), e.g., as done in the HistoGraph project (Novak et al., 2014).

Still, despite the specificities of the user requirements, we also find that the tools contain generic features. For example, we were happy to find an article in a Dutch newspaper in which the author stated to often use PoliMedia. This author however mainly used the tool for its search and filter options, without using the linked media coverage (Sanders, 2014). The within-scope user requirements helped to improve the tool to be used even for purposes not specifically intended.

The out-of-scope user requirements on the other hand provide hints of what the wider research workflow consists of for the different participants. That is, after finding the related media items with PoliMedia, scholars want to analyse these media items, or annotate it with their observations. With Oral History Today, we see that after finding an interesting video, scholars want to contextualize it and come to a full understanding of the interview. User researchers should thus keep in mind that the tool will most likely take a spot in a wider research workflow, and that it is infeasible to create a tool for the entire workflow that is generic enough to be applicable to a larger user group. In this sense our conclusions are in opposition with the ambitions of e.g., Blanke and Hedges (2013). Our findings instead suggest to focus on a single task within the workflow, which is reminiscent of the old adage do one thing and do it well (McIlroy et al., 1978). This proposition is compatible with Van Zundert’s suggestion of light-weight tools for specific humanities tasks (van Zundert, 2012). To some extent it seems compatible with Unsworth’s suggestion of tools for specific scholarly primitives (Unsworth, 2000), in that the research workflow is split up in a set of primitives. However, what we have learned from our user research is that the requirements for a tool related to certain tasks are related to what tasks come further down the workflow. As such, these tasks are not true primitives since their implementation is dependent of the rest of the workflow. To what extent certain tasks are generalizable is a question that requires further user research.

These insights furthermore lead us to conclude that in order to enable a workflow with multiple tools, Digital Humanities projects should separate the tool and the data. Even when the tool would not be compatible with a specific scholar’s research methods, the data should still be usable. In PoliMedia, not only was a tool created, but also a dataset, which was made available via a SPARQL-endpoint.7 Altern-

---

6 For more information about this PhD research and for future updates, see http://www.maxkemman.nl/category/phd-thesis/
7 http://data.poliMedia.nl
native approaches are an API or a downloadable dataset. This introduces a new continuum in the innovation of digital research methods, namely that from developing tools for scholars, via developing tools with scholars, to scholars developing tools. An in-depth discussion of this continuum is beyond the scope of this paper, but in receiving feedback on our published datasets we do observe that many humanities scholars have difficulty using data without an accompanying tool. Data reuse is not as simple an undertaking as one might hope (Borgman, 2015; Edmond & Garnett, 2015).

We note that there is a tension between the specificity of humanities’ research methods, and generalizability for a broader applicable tool. Our findings suggest however that user research has a clear benefit for Digital Humanities projects: first, the out-of-scope user requirements give insight into the tool’s compatibility with existing research practices. Second, the user requirements that were within-scope led to usable features that were sufficiently generic for the tool to be adopted, also for purposes for which it was not specifically intended. User research thus proved useful for the development of tools to be compatible with specific research methods of scholars, taking a place in a wider research workflow.

Acknowledgements

The PoliMedia and Oral History Today projects were funded respectively by CLARIN-NL (http://www.clarin.nl) and CLARIAH (http://www.clariah.nl). We would like to thank Joris van Zundert for his helpful feedback on an earlier version of this paper, as well as Andreas Fickers who also inspired the title of this paper.

References


