

Optimizing Explorative Search for the Needs of Media Professionals: The DIVE+ Use Case

Justin Verhulst
MSc Information Studies
0031643854823
justinverhulst@gmail.com

ABSTRACT

The work practices of media professionals involve sense-making, contextualization, and storytelling. An important part of their job is to tell fact-based stories on a wide variety of topics that are often not very familiar to them. Explorative search is, as such, an important part of their job, as this allows them to understand and deepen their knowledge on a topic. This research investigates whether the explorative search practices of media professionals can be supported by online tools - specifically, DIVE+ is being adopted as a use case. This Linked Data browser guides users in their explorative search through cultural heritage collections and it could therefore be useful to media professionals. Event-based exploration, the construction of narratives, and serendipitous discovery are central aspects of DIVE+, and we want to see if these aspects are suitable to media professionals. The following research questions guide this research: (1) *What are the digital search and exploration practices of media professionals?*, and (2) *To what extent are the exploratory search requirements of media professionals supported by the events, narratives and serendipity in DIVE+?* In order to answer these questions, multiple user studies were conducted which provided us with data about media professionals' exploration practices and experiences with DIVE+. These user studies revealed that the search requirements that media professionals have are not optimally supported by DIVE+. In order to better align the requirements with the search browser, some recommendations are given that will help developers to make informed decisions about how to improve the DIVE+ interface and adapt it to the needs and wants of media professionals.

Keywords

DIVE+; Media professionals; exploratory search; user studies

1. INTRODUCTION

Media professionals and scholars are avid users of cultural heritage collections. The vast amount of material available in archives can provide these users with a better understanding of history and can form the basis for research projects and stories. In order to improve access for these and other user groups, libraries and museums have embraced innovative Web technologies. Cultural heritage has become part of the Web of Data in initiatives like Europeana¹ with the aim to open up collections to the public. A central research challenge in this regard is how to best facilitate

access to the end users, such as the media professionals and scholars. The wealth of available material does require smart ways of linking and presenting the information so that people are not only able to find it, but are also able to understand and interpret it. As opposed to keyword-based search, which requires a clear, well-defined information need on part of the user (White & Roth, 2009), explorative search can help the user place cultural heritage collections and objects into context when the information need is less well-defined. As such, explorative search, and tools that facilitate this, allow for meaningful interactions with linked cultural heritage.

DIVE+² has taken on the challenge to improve access to and understanding of cultural heritage. Digital hermeneutics forms the foundation of DIVE+, which theorizes that by linking collections through events and narratives, users are better able to interpret and make sense of cultural heritage objects (Van den Akker et al., 2011). This is facilitated by DIVE+ through a web interface that allows users to interact with events and narratives. Concretely, DIVE+ aims to:

- facilitate a serendipitous browsing experience by means of events-enriched Linked Data
- guide the user in their exploration by means of personal narratives
- support users in their interactions with cultural heritage through event-based browsing

The current research contributes to the academic field in general, and to the DIVE+ project in particular, in the following ways:

- By investigating the research needs and search practices of a new user group, media professionals, we can determine if exploratory search and DIVE+ could be useful to them.
- By adopting a diverse set of user testing methods, in-depth insights are retrieved about the differences in the usability of DIVE+ for media professionals and the existing user groups.
- By translating search requirements into concrete recommendations for the interface, DIVE+ can be further improved.

¹ <http://www.europeana.eu/>

² <http://diveplus.frontwise.com/>

1.1 Research problem

In the last 3 to 4 years, DIVE+ has undergone a number of transformations of both the interface and back-end data support, in order to provide the most optimal support for the users. Recommendations from the numerous user studies provided the input for this. These user studies mainly involved humanities scholars, as this group works a lot with historical material and archival collections. As such, search tools that support them in this are very useful to them, and they will remain an important focus of the DIVE+ research community. Therefore, while the current research focuses on a user study that was done with media professionals, a user study with humanities scholars was conducted as well, which allows us to compare the two groups. More in-depth results from the user study with the humanities scholars can be found in the research by Cheng (2017).

While previous research has demonstrated that events and narratives within the DIVE+ browser indeed help scholars to better contextualize and interpret the collection, more research on the relationship between narratives and the interpretation process is needed (Kruijt, 2016). Also, as Collijn (2016) argues, more user studies are needed to test the DIVE+ interface so that it can be improved. A potentially relevant user group that could be part of the DIVE+ user studies are media professionals, since they share a lot of the exploration practices of humanities scholars. Just as humanities scholars, sense-making and contextualization are central in their work practices of media professionals. Investigating if, and how, events, narratives and serendipity are useful to them thus perfectly fits as a DIVE+ research project and helps us to determine what aspects and features of a search tool media professionals value when they conduct exploratory searches.

1.2 Research questions

First, in order to find out if DIVE+ is suitable for media professionals, we need to investigate how this group engages in exploratory search practices. The following research question is proposed to address this information need:

1. What are the digital search and exploration practices of media professionals?

More specifically, by investigating the search behavior of media professionals we can come up with a list of requirements that media professionals have when engaging in (exploratory) search practices.

Second, narratives, events and serendipity were identified as central features in DIVE+, which are supportive in the exploratory search process as it places collections and media objects into context. In order to find out if these aspects are helpful for media professionals as well, the following research question is proposed:

2. To what extent are the exploratory search requirements of media professionals supported by the narratives, events and serendipity in DIVE+?

2. RELATED WORK

In section 2.1, we first explain exploratory search. Then, in section 2.2 a look is taken at the existing search practices of the main target group of this study, media professionals. Section 2.3 explains the value of narratives and events based on the theory of digital hermeneutics. Finally, in section 2.4, a brief summary is given on some common user testing methods, as described in the literature, that are applicable to the current study.

2.1 Search stages and exploratory search

When people engage in online search activities, they apply different kinds of search strategies in different stages of the research process (Kuhlthau, 2004). A person that is in an advanced stage of collecting information is able to narrow the focus and is therefore able to specify information needs. In this stage, the searcher engages in ‘lookup’ tasks, which is suitable for basic search activities where the user is looking for specific information (Marchionini, 2006). However, a person who has just begun to decide on a theme or topic has to retrieve information from a broad set of potential topics, before a focus can be taken. This latter group engages in search activities which Marchionini (2006) calls ‘learn’ and ‘investigative’ tasks. In this stage of research, the search for information can be described as explorative.

2.1.1 Exploratory search

White and Roth (2009) explain the difference between traditional, look-up type of search and exploratory search. For look-up type of tasks, the information need is clear in the mind of the user. For example, a journalist could be interested in where or when a news event happened - fact-finding and question answering are central. The authors refer to this as the query-response paradigm: The user has a clear, well-defined information need that can be immediately resolved by a specific query. On the other hand, White and Roth (2009) state that the information need is less well-defined for exploratory search. The user is unfamiliar with the topic of interest, has vague and complex information needs, and learning and understanding are more important than getting answers to specific questions. A person can try to resolve these type of information needs through exploratory browsing sessions that last over a longer period of time and that involves multiple iterations and query reformulations. During these sessions, a user collects and assimilates a lot of new information, which could result in uncertainty and confusion. Tools for exploratory search should help users to deal with this.

2.1.2 Tools for exploratory search

As Marchionini (2006) notes, search tools in which users can conduct exploratory searches to make sense of information are needed. This is also noted by White and Roth (2009), who state that tools should “define the problem, make sense of the encountered information throughout the current session and across multiple sessions, and handle uncertainty and confusion by providing progress updates, explanations for system actions, and summaries of major themes present in encountered information” (p. 13). In line with these ideas, DIVE+ was developed in order to enable scholars and the general public to explore and understand cultural heritage collections. The focus is on learning and investigating, rather than on finding specific information.

Now that the concept of exploratory search is clarified, we will look into related work concerning the (exploratory) information needs and search behaviors of media professionals.

2.2 Search practices of media professionals

2.2.1 Creative retrieval practices

Search practices and information needs of media professionals have been studied before. For example, Sauer (2016) conducted research on the creative retrieval practices of this group. Through expert interviews, a range of factors were identified that influence the search behavior of media professionals. Time allotted to search was identified as the most important factor that affects search behavior. Some media professionals work under time pressure and this shapes their search practices. For example, documentalists who work at a news organisation have to gather material rapidly as the item has to be finished before the broadcast. Other factors that affect search behavior are the budget, the genre of the media text, the intended audience of the media text, and personal interest.

Furthermore, media professionals engage in serendipitous search behavior during the creative retrieval process. The concept of serendipity refers to the discovery of pieces of useful information that were not initially sought for by the user (Toms, 2000). As Sauer and De Rijke (2016) note, when media professionals search through archives, they adopt strategies that allow them to find unknown audiovisual material, which they can use in the creation of stories. This reliance on serendipitous findings is evidenced by the interviews that Sauer and De Rijke (2016) conducted with media professionals: the interviewees noted they not only look for things that they know, but that they also like to discover new things, things they are not aware of at the start of the search.

A range of studies looked more specifically into the search practices of one group who also can be considered media professionals, namely journalists. Lecheler and Kruikemeier (2015) analyzed and synthesized the available research on this topic and they note that the relevancy to study the use of online sources by journalists is evident. The use of online sources have become embedded in the daily work practices of journalists. Lecheler and Kruikemeier (2015) note that journalists develop routines for dealing with these sources, in which journalistic values such as objectivity are being strived for. The routines that journalists employ for accessing and assessing online information can have an impact on the quality of news reporting. The skills to verify online information and assessing the credibility of sources have become of crucial importance (Lecheler and Kruikemeier, 2015; Schifferes et al., 2014; Kemman et al., 2013). In line with this is the discussion of van der Haak, Parks and Castells (2012) about the future of journalism. They argue that “transparency and independence are vital for journalism to be credible in the 21st century” (p. 2931). Moreover, they note that journalism will increasingly focus on storytelling, and in order to do this effectively, “professionals should be liberated to focus on explanation, contextualization, [and] sense-making” (p. 2935). Thus, online tools and sources should help journalists to deal with the richness of data and information out there so that they are able to tell meaningful, fact-based stories. As such, examining

journalists’ use of online sources and tools is a relevant and useful activity.

2.2.2 Search phases

Sauer (2016) discusses the process of creating an audiovisual narrative by media professionals and defines three phases. In the first phase, the overarching story is researched and the primary sources are sought. In the second phase, archives are searched and sources are selected. In the last phase, ideas for stories about a subject are developed based on the collected materials. While there are differences between media professionals, these three phases are the main steps media professionals go through when they develop an audiovisual story.

Kemman et al. (2013) studied the online search strategies of Dutch journalists. They see similarities between the production process of a journalistic article and Kuhlthau’s (2004) search stages. This process is described by Kemman et al. (2013) as a ‘search diamond’, in which a broad information need is followed by a more focused information need. More specifically, in the first two stages, initiation and selection, the journalist identifies a general idea for a news story. This is followed by an exploration and formulation stage in which the journalist explores potential angles and formulates ideas for an article. In the collection stage, information that relates to the chosen angle is collected and finally, in the presentation stage, the information is assembled in a unified form, the news article. The current research is mainly interested in the first part of the diamond, where the researcher broadens the search through initiation, selection, and exploration. Within DIVE+, this type of search is characteristic. It is good to be aware of the stages as they indicate the information needs of media professionals at different steps in the production process.

2.3 Digital hermeneutics

As Van den Akker et al. (2011) describe, the appropriate framework to think about the human interpretation process in a web setting is digital hermeneutics. The scholars see it as a central theory for studying the access to and interpretation of online digital heritage collections. It is a relevant theory in light of the current research, because we investigate if, and how, media professionals can be supported in their online interactions with cultural heritage collections, so that they are better able to interpret it. Digital hermeneutics states that by explicating relationships, between objects and events and between two events, a user is able to better understand cultural heritage collections. Van den Akker et al. (2013) also note that digital hermeneutics deals with “the study and theory of the design and evaluation of Web applications as a means of interpretation” (p. 432). DIVE+ aligns with these ideas as this search browser revolves around the idea that the interpretation of heritage collections can be improved by providing users with a web interface in which they can interact with objects in the collection. An important question for the DIVE+ research community, related to digital hermeneutics, is whether events and narratives are useful methods to enhance the interpretation process.

2.3.1 Usefulness of narratives and events in the interpretation process

The research that has been done so far suggests that narratives and

events are useful mechanisms for improving the understanding and interpretation of cultural heritage collections by humanities scholars. Van den Akker et al. (2013) evaluated the usefulness of narratives on Agora, a predecessor of DIVE+, and came to the conclusion that narratives have a positive influence on interpretation. Kruijt (2016) also theorized that implementing narratives in DIVE+ can support researchers to interpret the collection. In her research, Kruijt (2016) aimed to find out whether a set of features in DIVE+, such as narratives, can help scholars to better contextualize and interpret the collection. Her main finding was that when researchers explore the collections in DIVE+, narratives do help them to interpret it. However, she encourages further research on this since there is a limited amount of applied studies that analyse the relationship between narratives and the interpretation process. For a rationale of how narratives should be implemented in DIVE+, she argues, it is important to test actual prototypes with users. On the same line, Collijn (2016) encourages to do more user studies on DIVE+, so that the tool can be improved. It should be noted that these studies looked into the interpretation process of scholars, and more specifically humanities scholars and students. For these user groups, the narratives help them to interpret the collection. However, this does not necessarily mean that this is also the case for other user groups as well. The claim of Van den Akker et al. (2013) that event-based narratives help the users to make sense of objects in collections should therefore be investigated for other user groups, such as media professionals, as well. This is furthermore recommended by Collijn (2016), but she stresses that the group to be studied should not be too broad. She argues that ‘researchers’ as a user group is too broad and that the focus should therefore be on a specific group of researchers.

2.4 User studies methods

The current research aims to determine how the previously described concepts - events, narratives and serendipity - are useful to end-users of DIVE+. As such, the research positions itself as a user study in which we investigate how these features are used and evaluated by different user groups, specifically media professionals. Literature on information retrieval research distinguishes two methodological paradigms for assessing user interactions with an information system: system-centered quantitative research and user-centered qualitative approaches (Vassilakaki, 2014). The current study mainly utilizes qualitative methods, and specifically tests an interface with users by giving them an exploratory search task. This type of user testing is a common approach to investigate (exploratory) search behavior (Kules et al., 2009).

2.4.1 Exploratory search task

The exploratory search task, also referred to as ‘simulated work task’, is “a short textual description that presents a realistic information requiring situation that motivate the test participant to search the IR system” (Borlund, 2016: 395). The main goal is to provoke a genuine information need on part of the user, so that the

search engine can be used realistically. Borlund stresses that a simulated work task should be realistic: it should be tailored to the test participants. In other words, the described task should be something that the test participants can relate to. Moreover, the topic should be interesting to the test participants. Allowing the test participants to choose a topic is therefore recommended, as this allows for a personalized search experience. Furthermore, pilot testing of the task should be conducted in order to make sure that the task is suitable for the test participants. When these requirements are followed, a task that enables the capturing of the ‘real’ search behavior of users should be the result.

An exploratory search task, which is used in the current research, has its specific set of requirements. As we have seen, exploratory search can be characterized by uncertainty, ambiguity and discovery (Kules et al., 2009). Therefore, a task intended to evaluate exploratory search behavior should induce these information needs. The task should be, as Kules et al. (2009) argue, broad in scope, so that exploratory search is promoted. In developing the tasks, these recommendations were followed in order to provoke realistic and genuine search behavior.

3. The DIVE+ Use Case

The exploratory search browser adopted in this research as a use case is DIVE+. The browser contains four datasets: Dutch news broadcasts from the Netherlands Institute for Sound and Vision, ANP Radio News Bulletins from the Dutch National Library, cultural heritage objects from the Amsterdam Museum, and cultural heritage objects from the Tropenmuseum (De Boer et al., 2017). As we have discussed, DIVE+ aims to, in line with the ideas from digital hermeneutics, support users in their interpretation of cultural heritage collections. Three essential components, or features, of the interface are contributing to this: data is modeled and structured through events, it is possible to construct narratives by means of a visualised search path, and it stimulates serendipity through the (implicit) connections it forms between the data. The workings of these three components within DIVE+ will be demonstrated below.

3.1.1 Events in DIVE+

Historical events and related objects form the basis of DIVE+ and can be seen as the building blocks for the narratives. There are lots of different types of events, and events and their interrelations can be represented in many ways. DIVE+ uses the Simple Event Model to formalize the relationships between objects and events. This model postulates that an event can be related to an object, and an event can be related to another event. More specifically, as Van den Akker et al. (2011) describe, object-event relationships can be formalized in three ways: (1) The object depicts an event, (2) The object is used or functions in an event, and (3) The making, collecting, or exhibiting of the object itself is an event. The event-event relationships are identified through a mutual location, actor, or concept. Figure 1 illustrates these object-event and event-event relationships concerning a painting of the attack on Yogyakarta.

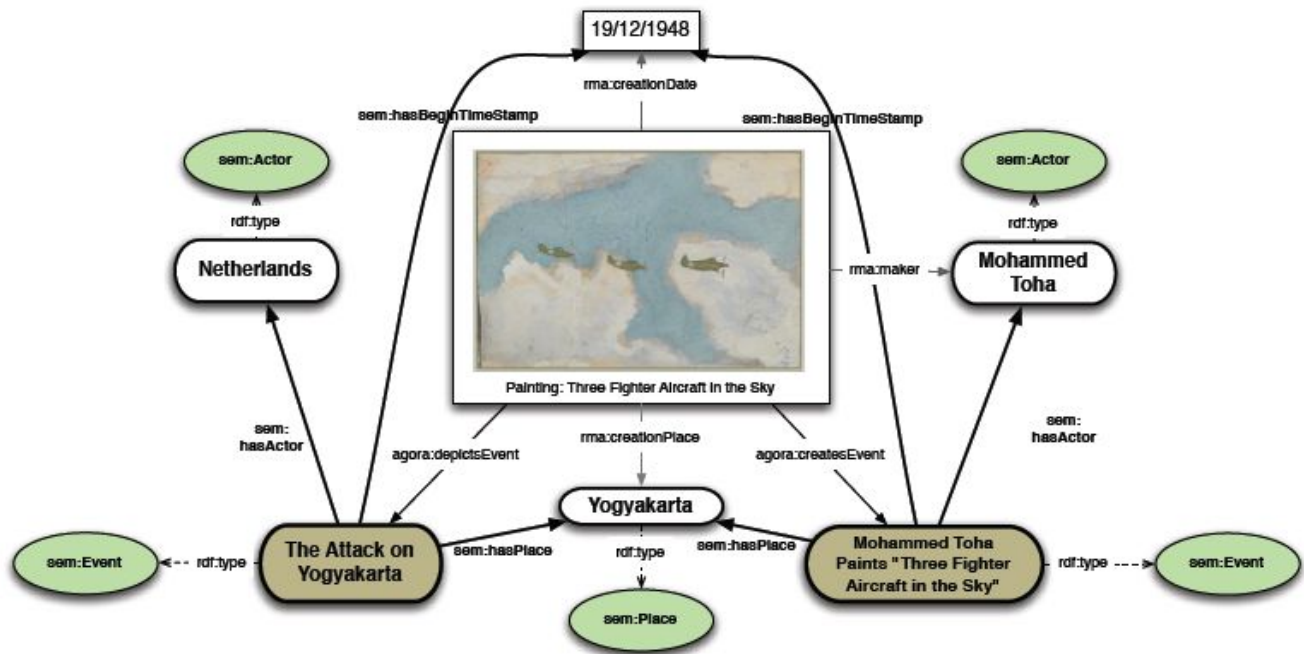


Figure 1: Event dimensions of an object (Van den Akker et al., 2011)

The cultural heritage objects in DIVE+ are enriched with metadata that describes what type of object it is and which type of event it belongs to. A user can apply filters based on these demarcations (see figure 2), and can in this way create historically meaningful narratives from different perspectives.

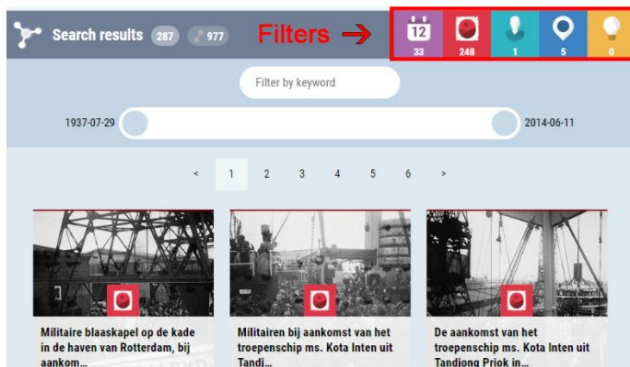


Figure 2: The event filters in DIVE+

3.1.2 Narratives in DIVE+

Van den Akker et al. (2011) describe that a narrative is formed by placing historical objects or events into relations. The idea, then, is that the relations between different events and objects, for example in an interactive search browser like DIVE+, allows the user to place objects and events into context. More specifically, a narrative can be seen as a particular sequence of objects and events which together constitute one angle on a (sub)set of the cultural heritage collection. For example, a user that is interested in the second world war might choose to select and link events and associated objects that relate to military actions during the

war. Another user might instead focus on the role of women during the war. In this way, different users are able to construct their own ‘story’ by following a path of related entities.

Within DIVE+, the narrative construction is facilitated through the ‘exploration path’ (see figure 3). When a user browses the collection, the entities that are selected and followed are automatically added to the exploration path in the DIVE+ interface. Moreover, the user is able to save the current narrative and can load existing narratives.

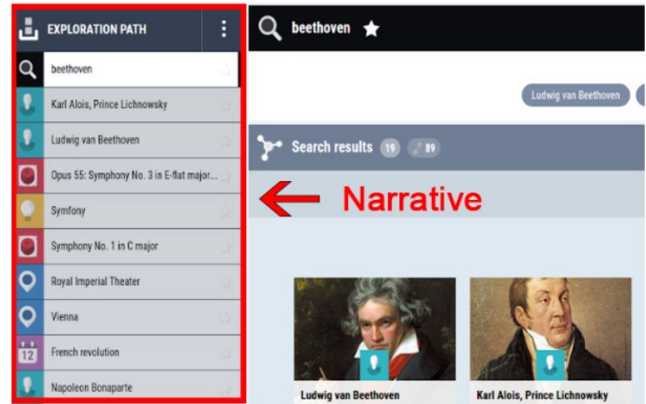


Figure 3: The exploration path in DIVE+

3.1.3 Serendipity in DIVE+

DIVE+ aims to support serendipitous information retrieval. Serendipity is stimulated with the help of the previously discussed features of narratives and events. By linking events and objects and by visually presenting them, the exploring user can stumble

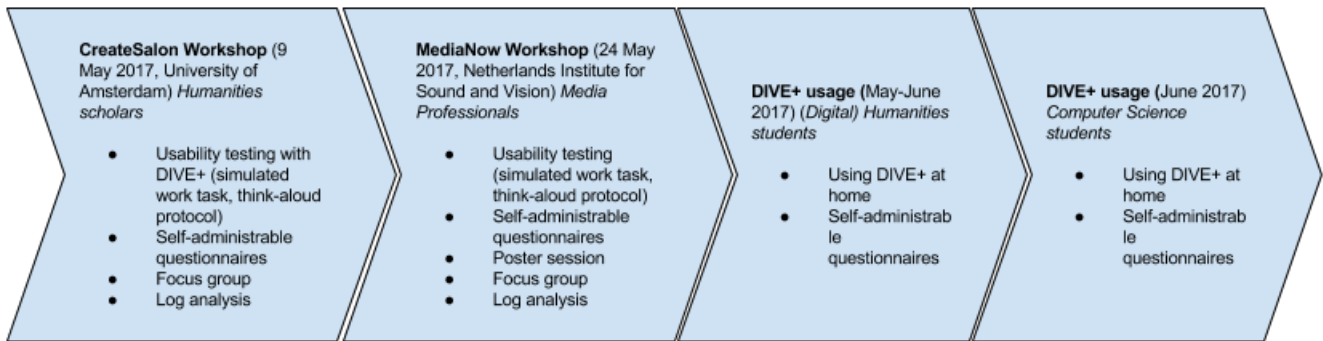


Figure 4: Overview of the different user studies and methods

upon unexpected, but useful items. Moreover, narratives can be saved and compared with other narratives, which also facilitates unexpected discoveries. As Sauer and De Rijke (2016) demonstrated, serendipity is an important part of the search process of media professionals, because they partly rely on ‘hidden gems’ when they search through archives. The serendipitous browsing path of DIVE+ is therefore, we suppose, a valuable asset for this group.

4. METHODS

In section 4.1, the set of methods used in this research is listed, and we explain why, and when, we used the different methods. Then, in section 4.2, we explain in detail how the methods were developed and how data was collected with the various methods.

4.1 Research design

This study’s research goals are twofold, namely (1) to investigate the digital search and exploration practices of media professionals and (2) to assess the usability of an exploratory search browser, DIVE+, with regard to the digital search needs of media professionals. A range of methods is used to achieve the research goals. As Kules and Shneiderman (2007) argue, the situated nature of exploratory search makes it difficult to evaluate this by quantitative measures alone. The relevancy of search results is hard to assess, as searchers’ information needs may be ambiguous and vague and may be different from person to person. Looking at aspects such as dwell time or click-through-rate are therefore not enough to fully capture users’ satisfaction with an exploratory search tool. Using a multitude of methods, both quantitative as qualitative, is therefore warranted as this will provide better means to assess exploratory, creative search.

Specifically, the following data collection methods were adopted:

- Usability testing of DIVE+ by means of simulated work tasks, in which think-aloud protocols were followed
- Self-administrable questionnaires
- Focus group
- Poster session
- Log analysis

The methods were used in different compositions and in different settings in the period from May to July 2017, as illustrated in figure 4. Four user studies were conducted, with four different user groups: humanities scholars, media professionals, (digital) humanities students and computer science students. This unique combination of user tests allowed us to retrieve in-depth insights on the exploratory search practices of media professionals and their use of DIVE+, while it also gave us the possibility to determine how this group compares with students and humanities scholars.

4.2 Tool development

During the workshops with the humanities scholars and the media professionals, the participants were invited to engage in usability testing with DIVE+. The testing was done by means of a simulated work task and a questionnaire. Both the task and the questionnaire were the result of an iterative process of development and refinement, which took place in April and May 2017. The visualisation of this process (see figure 5) shows that a first toolset was created by analysing a dataset that contained students’ perspectives on DIVE+ (Hagedoorn & Sauer, 2017). By pilot testing this toolset with six Information Science students, the tools were further refined.

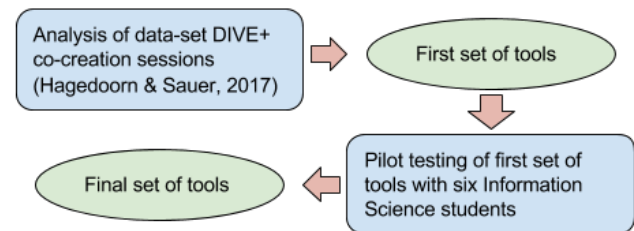


Figure 5: Tool creation process

4.2.1 Simulated work task

As the basis for the simulated work task, we used and adapted the framework developed by Kules et al. (2009), and we followed the recommendations by Borlund (2016). In the task description, first the context of the information need was given, after which more detailed search instructions were outlined. The context and the search instructions were quite broad so that exploratory search behavior was promoted.

The following information need was given to the test participants. As two test sessions were organized, with different types of participants, we slightly adapted the task to fit the specific information need. In appendix C, the full task, with specific descriptions, can be found.

Workshop MediaNow (media professionals): *Imagine that a media company is going to produce programs about Jakarta, Beatrix, Islam and Watersnoodramp. Your goal is to propose an interesting angle for one of the programs.*

Workshop Createsalon (humanities scholars): *Imagine that an academic journal has opened up a call for papers on the topics of Jakarta, Beatrix, Islam and Watersnoodramp. Your goal is to propose an interesting research angle for a paper.*

4.2.2 Think-aloud protocol

During both the workshop with the humanities scholars as the workshop with the media professionals, the researchers involved in this project walked around and sat by different participants while they were testing DIVE+. A think-aloud protocol, a common usability testing technique (Nielsen, Clemmensen & Yssing, 2002), was followed, which allowed us to retrieve concrete information about the experiences of the users. Before the session began, the participants were encouraged to think aloud during the testing. These interactions were recorded and transcribed.

4.2.3 Questionnaire

The questionnaire served as a tool for capturing the experiences of the participants in a more structured way. In the two workshops, a link in the task description guided the participants to the online questionnaire, so that they could fill it in individually, directly after they completed their test with DIVE+. Also, (digital) humanities students and computer science students were asked to test DIVE+ at home and to fill in the questionnaire. The team of researchers collectively developed this questionnaire and questions were included that addressed a range of research needs. Specifically, three main question categories were included, which related to DIVE+ and the interface in general and narratives and events in specific. Moreover, background information about the participants and former knowledge of exploratory search and linked data was inquired. In appendix D the full questionnaire can be found.

4.2.4 Focus group & poster session

During the workshops, additional qualitative data was collected in order to obtain more in-depth insights about the experiences of the participants in using the exploratory search browser. In a poster session, participants worked together in groups to sketch and discuss their search flows. Due to time limitations, this was only done during the workshop with media professionals. A focus group, held at both the CreateSalon and MediaNow workshops, served as a way to capture information about experienced with DIVE+ and search behavior in general. As these focus groups were held at the end of the sessions, it allowed us to address topics that were not addressed in detail earlier during the session.

4.2.5 Log analysis

Finally, transaction log data of users' interactions during the workshops was obtained as this provides insights into the actual use of DIVE+ (see appendix E & F). While this research is interested in the perspective of the users themselves, 'objective' data such as transaction logs can provide relevant information as well. The transaction logs were derived in two ways. First, DIVE+ provides a real-time action log³, in which interactions with DIVE+ are automatically added as JSON data. In order to identify different users, session tokens were automatically added to an individual test session. Second, within the interface of DIVE+, it is possible to save and export the existing exploration path that the user has constructed as a JSON file. In order to retrieve this exploration path separately from each of the test participants, we asked them during the test sessions to export their exploration path.

5. RESULTS

The results of the various user studies can be found in appendix G. In this section, we give the most prominent results. Specifically, section 5.1 describes the demographics of the people that participated in the different user studies. In section 5.2, the results are listed that relate to the general search practices of media professionals and the requirements that they have when searching. Section 5.3 deals with the opinions of the media professionals on DIVE+. As one of the research goals is to understand how media professionals differ with the other user groups of DIVE+, results from other user groups, such as humanities scholars, will be noted as well when relevant. Refer to Cheng (2017) for more in-depth results of the group of humanities scholars. In describing the results, the participants are anonymised and are referred to by a combination of the user study the result is from, the method of gathering the data, and a number. For example, focus group participant 3 in the workshop with media professionals is referred to as WS_MP: focus group_3.

5.1 Demographics Across user Studies

- **Media professionals:** A total of 11 media professionals participated in the workshop. All of them participated in the different sessions (poster session, focus group, and the user test), but the questionnaire yielded a total of 8 respondents. The difference in the total number of participants and the number of participants that filled in the questionnaire is due to the fact that some of the participants worked together on one laptop and some had to leave early. Their background was varied as they occupied different positions in media- and archival organizations. They worked as information specialists, filmmakers, editors, media and access managers, and innovation specialists.
- **Humanities scholars:** Another workshop was held with scholars, specifically from the humanities domain. In this session, 18 people participated and 11 filled in the questionnaire.

³ http://diveplus.frontwise.com/logs/Y2fPEHdjmgBGP_dive.log

- **(Digital) humanities students:** A total of 16 students from the humanities and digital humanities domains tested DIVE+ at home and filled in the questionnaire.
- **Computer science students:** Finally, 22 bachelor students from the computer science domain tested DIVE+ at home and filled in the questionnaire⁴.

5.2 Media Professionals Search Practices

The first research goal was to find out what the search and exploration practices of media professionals are. This was investigated by looking at the **qualitative data** retrieved from the focus group and poster session during the workshop with media professionals. By coding the output of these sessions, common categories were identified that relate to the needs and wants of media professionals when they conduct exploratory search. In Table 1, an overview of the **most important findings**, related to the search and exploration practices of media professionals, is given. The full document can be found as an online appendix⁵. Two overarching categories were found which will be explained in more detail below: work-related influences on search and general search flow and search strategies.

Table 1. Main themes and findings identified regarding the general search practices of media professionals

Theme	Finding
Work-related influences on search	The target audience of a program determines what kind of material or point of view is sought
	There are time and budget restriction for searching
	Transparency within a search tool or archive is considered as important when searching
General search flow and search strategies	When searching for a story, the search process goes from macro-level to micro-level
	A range of search sources and strategies are mentioned that relate to the broad, macro-level start of the search (e.g. own background knowledge, previous coverage, Wikipedia, Youtube) and the more specific, micro-level of search (e.g. newspaper databases, locations, names, eyewitnesses, jargon).

5.2.1 Work-related influences on search

The media professionals were divided into three groups during the poster sessions and all of the groups stated that the search process is to a large extent influenced by a range of work-related

constraints. The listed constraints were **time, budget, and target audience of a program**. Also, they mention the importance of **transparency** in their jobs.

Budget and time

The **budget, and related to this, the time that is available for searching** are often mentioned as factors that influence search. For example, when group 3 (WS_MP: poster_3) explained how they would normally conduct searches, they argued that *“eventually, the assignment is the most important. We identified a number of parameters. So budget, maybe the first one is time, how much time do I have, the medium, with which medium do I work, radio, television”*. The time and budget constraints differ depending on the **type of program**. An information specialist working at a Dutch news broadcaster, part of group 1 in the poster session (WS_MP: poster_1), mentions the difference between searching for material for the evening news and for documentary-style programs: *“what we really have is the time restriction. If I would work for ‘watersnoodramp’ at Andere Tijden, I would do research in the archive, look what they have there”*.

Target audience of a program

Multiple participants stated that the **target audience of a program** influences the search to a large extent. If a program has a younger audience, material that is recognizable is suitable, while programs that have an older target audience rather use material that is new and original. So there is a clear difference between the material sought when the target audience is different: *“We were talking about restrictions that influence your search. I work at the NOS, when you search for the Youth News, you are looking for the cliché explanation story, so you are looking for facts and cliches, while for the Evening News, you are looking for an original angle”* (WS_MP: poster_1).

Transparency

Transparency within a search tool is highly valued by the media professionals. In order to conduct reliable searches, they need to **know why results are presented to them**. If this is not evident, they might **lose trust** in the search tool. That transparency is highly valued in their work is evidenced by the fact that nearly all of the participants mentioned that they need to know ‘what’s behind it’, and that they want to ‘see it with their own eyes’. A participant in the focus group exemplified this point when discussing the results that were shown to him when using DIVE+: *“to see which keyword did you use, or what is the basis when I enter something as a keyword, I don’t see it coming back in the results. Those are the things you are used to. [...] I want to have that double check, I don’t trust that it happens behind the scenes”* (WS_MP: focus group_5).

5.2.2 General search flow and search strategies

During the poster session, media professionals described that when they search for material on a topic that they are not familiar with, **they start with a broadly oriented search**, for example by using general query terms. In this search phase, general sources such as YouTube, Wikipedia, or newspaper databases are consulted in order to generate initial ideas. After this initial stage of search, the media professionals **narrow their search and search more specifically** by using more specialized

⁴ Due to time limitations, only the answers of the computer science students on the closed-ended questions were analysed.

⁵https://docs.google.com/spreadsheets/d/1rDQjH9cdNs0PkifNJDYyD88Fxdty10FB_ev7tRHd-E/edit#gid=0

search tools and by using terms relating to for example locations, names, eyewitnesses and jargon. This helps them to get a ‘micro-story’ to the surface: “*specifically searching on the villages that were involved so that you are able to find a story, reports of eyewitnesses, newspaper databases, so multiple sources combined, so that you can retrieve some kind of a micro-story*” (WS_MP: poster_1). This search flow, which start at a general and broad level and becomes more focused during the process of search, can also be seen when looking at the DIVE+ log data of the media professionals (see appendix F). The queries show that their initial searches are broadly oriented, for example indicating one general concept or place. Later queries are much more specific and contain multiple terms.

5.3 DIVE+

The second research goal was to find out if DIVE+ is suitable for the search and exploration needs of media professionals. In particular, we want to know whether the exploratory search practices of media professionals are supported by the events and narratives in DIVE+, and whether DIVE+ stimulates serendipitous browsing for this group. The questionnaire, in which questions were added about the usefulness of the exploration path and the event representation in DIVE+, helped to investigate this. More specifically, a set of closed-ended questions inquired about participants’ opinion on the exploration path and narratives, the events and event entities, and DIVE+ and the user interface in general. In order to make sure that the participants understood that the questions were about these specific features, screenshots of the interface were added to the questionnaire. Additional information was captured through the open-ended questions, the think-aloud protocol, and the focus group. By coding the output of these sessions, common categories were identified that relate to the opinions of the participants on DIVE+. Transaction logs of the actual search behavior were captured as well for triangulation purposes. In appendix E and F, the full transaction log can be found for both the group of humanities scholars as the media professionals. This shows that on average, the humanities scholars included 11 items in their exploration path, while the media professionals included on average 17.7 items in their exploration path - containing a combination of entities and queries. It is likely that the latter group included more items because their testing time with DIVE+ was a bit longer. This data suggests that both groups thoroughly tested DIVE+ and as such their opinions on the different features of DIVE+ are grounded.

In appendix F, a table provides the results (and demographics) for all the studied groups. Here, we summarize the most notable results for the media professionals (Table 2), and we make comparisons with the other groups when there are notable differences. Furthermore, an online appendix⁶ lists all the results and evidence (mostly quotes from the participants). The most notable results will be briefly explained here with regard to the answers of the respondents on the closed-ended questions about DIVE+ in general and the events and narratives in specific. In the discussion section, more detailed explanations are given.

⁶<https://docs.google.com/spreadsheets/d/1xL3sNIhOaCrAZAOP-s1pIkoOi6B1d0ewDFmhcs2bSIQ/edit#gid=0>

Table 1. Overview Results for Media Professionals

Theme	Finding
Transparency of information	More information is needed about which collections are in DIVE+
	Respondents needed more (meta)data and topic descriptions
	Respondents needed more information about how relationships between entities are generated
	Respondents needed more information about why certain labels (e.g. actor, concept) are given to entities.
Filtering options	More fine-grained sorting of the media types is needed
	More fine-grained sorting of the collection is needed
Exploration path and narratives	Respondents find that the exploration path does not result in narratives

Events and event entities in DIVE+

Figure 6 shows the answers of the respondents regarding their opinions on the events and event entities in DIVE+. The second chart for each statement refers to the media professionals. The majority, or even all of the media professionals answered ‘neutral’ or ‘disagree’ on the first three statements. It should be noted that the amount of neutral responses is relatively large (up to 50%), which could indicate that the respondents were either indifferent or the statement was not clear to the respondents. Still, as just a minor part, or even not one of the respondents agreed on the first three statements, we can conclude that media professionals generally found that **the meaning of the different event characteristics was not clear, it did not help them to learn about historical events, and the event representation in DIVE+ was not considered useful** when researching topics of interest. Furthermore, most of the media professionals (86%) **missed some characteristics of events**. Concerning the other user groups, a similar trend is visible: **the current usefulness of events was considered limited**. However, an interesting difference is visible when comparing the first two groups (media professionals and humanities scholars) with the last two groups ((digital) humanities students and computer science students). The latter groups were less negative towards the (usefulness of the) events and event entities. For example, while a large part of humanities scholars and media professionals found that the event characteristics were not useful to learn about historical events (70% and 50% respectively), this was less so for the student groups (25% and 27%). A striking difference is visible regarding the statement “I missed some characteristics of events”. While 86% of the media professionals agreed with this statement, only 23% of the computer science students agreed with it.

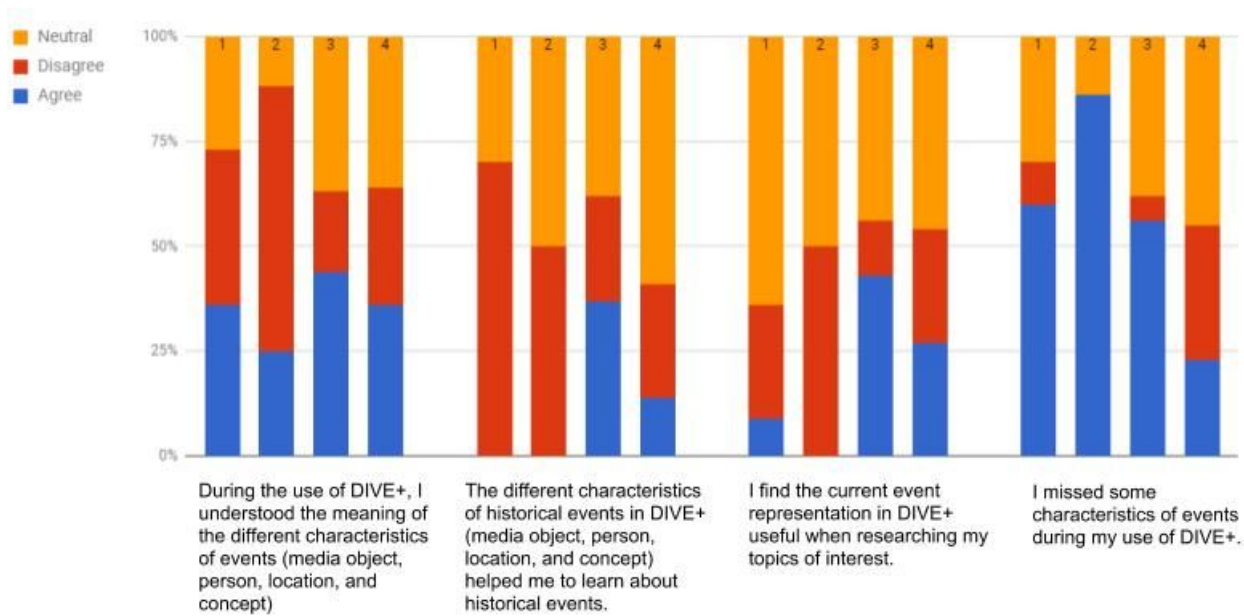


Figure 6: Answers of the four groups of respondents (1 = humanities scholars, 2 = media professionals, 3 = (digital) humanities students, 4 = computer science students) on the closed-ended questionnaire questions about the events and event entities in DIVE+.

Narratives in DIVE+

Figure 7 shows opinions of the respondents on the exploration path and narratives in DIVE+. A majority of the media professionals indicated that they **did not discover different narratives** in DIVE+ (57%) and that **the exploration path did not help them to learn about historical events** (67%). Furthermore, 72% of the media professionals **found suggested narratives not helpful** to explore a topic in depth (72%). Interestingly, only a small part disagreed, and more than half is neutral, concerning the statement “I find the exploration path a

useful feature”. When comparing the results with the other user groups, a **striking difference can be found concerning the opinions on the usefulness of suggested narratives for exploring topics in depth**. While 72% of the media professionals do not find suggested narratives useful, the numbers are reversed for the other groups (respectively 18%, 12%, and 23% for the humanities scholars, (digital) humanities students, and computer science students). Thus, media professionals were most critical of the four groups concerning the suggested narratives by a search tool.

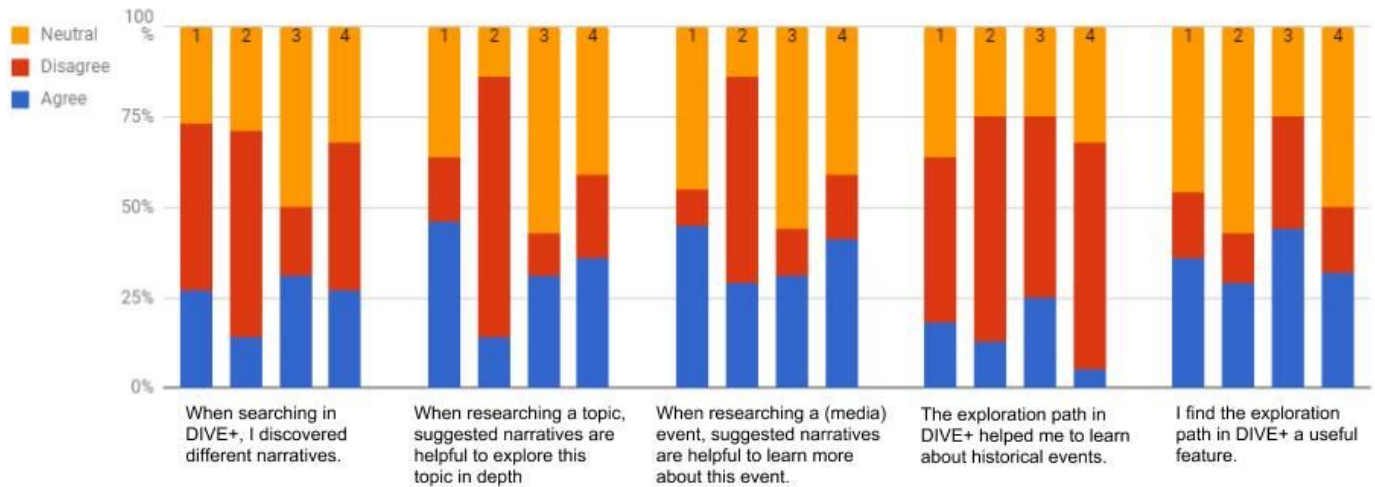


Figure 7: Answers of the four groups of respondents (1 = humanities scholars, 2 = media professionals, 3 = (digital) humanities students, 4 = computer science students) on the closed-ended questionnaire questions about the exploration path and narratives in DIVE+.

Usability of DIVE+

When looking at the general evaluation of DIVE+, we see that the the larger majority of the media professionals were quite **critical and assessed the search tool negatively** when it comes to the **exploration of topics**, the potential for **idea generation**, and the experience with **the interface in general** (see appendix I). A similar trend is visible for the other studied groups: only a minor part assessed DIVE+ positively.

6. DISCUSSION

The aim of DIVE+ is to help users in their interpretation and understanding of archival collections and historical events. It does so, it is argued, by linking events and event entities with each other, so that the user can browse through these events and can form narratives, and as such can stumble upon unexpected, but useful, findings. However, as shown in the results section, the users that participated in our user studies did not uniformly see or experience this, and as such they assessed the tool negatively. This is not only true for the main user group of this study, media professionals, but also for existing user groups of DIVE+, namely humanities scholars and students. In this section, we give a more detailed description of these negative experiences and the issues that are the cause of it. Specifically, in section 6.1 a model is introduced that visualizes the gap between the theorized working of DIVE+ and the working of DIVE+ for the users in practice. Then, in section 6.2 the factors that could explain the negative assessment of DIVE+ are discussed. Section 6.3 describes how the issues relate to events, narratives, and serendipity, and finally, in section 6.4 some recommendations are given for the DIVE+ interface which could resolve some of the issues.

6.1 Difference between theory and practice

The issues in DIVE+ are be explained by a model that illustrates the difference between the theorized working of DIVE+, and the actual working of DIVE+ for the end users. As Figure 9 shows, if the user experience is optimal, the (representation of) events within DIVE+ lead to narratives and a serendipitous browsing path for the users. Together, this will result in enhanced information interpretation: understanding of the collections is improved and the user is better able to place cultural heritage into context. This is considered the added value of the exploratory search browser.

Figure 10 shows that this theorized and optimal search process does work out differently in practice for the group of media professionals. A range of problems surfaced during the user tests which had a negative influence on the narrative construction and serendipitous browsing path of the user, therefore imposing a limit on the interpretation process. The identified problems can be summarized by a lack of transparency and limited user control. Moreover, some external influences, which were evident for the group of media professionals, resulted in a negative assessment of DIVE+. These three factors will be explained in more detail below.

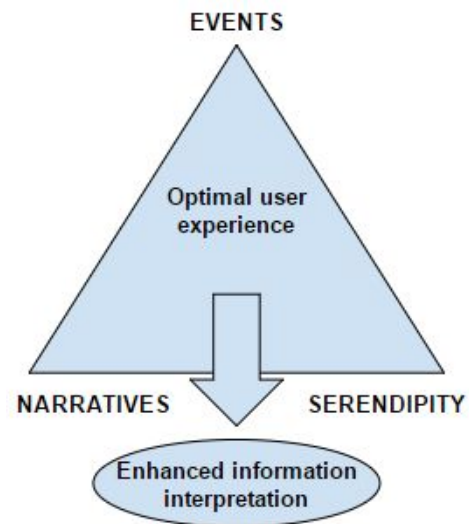


Figure 9: theorized working of DIVE+

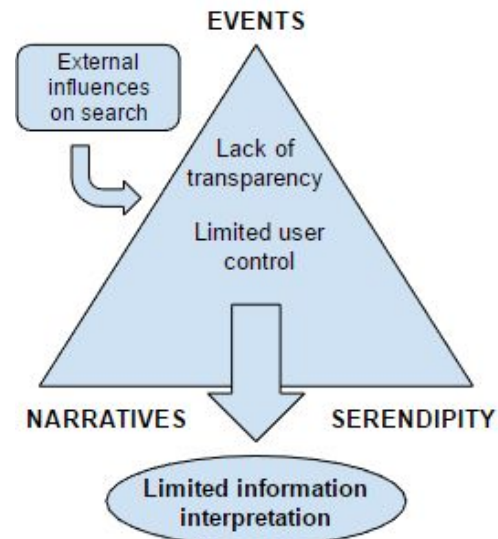


Figure 10: model of the working of DIVE+ for media professionals

6.2 Explaining the negative assessment of DIVE+

6.2.1 External influences on search

First of all, in investigating the search process of media professionals, we found that they start at a broad investigative level and narrow their search along the way. This finding is in line with the search flow of journalists, described by Kemman et al. (2013) with the metaphor of a search diamond. The scholars explain that journalists usually start with a broad information need, which is followed by a more focused information need. In the first phase, the media professionals that were studied in this

research referred to sources like YouTube, Wikipedia, or newspaper databases as ways for them to get inspired, to gather initial ideas for a story. This finding suggests that exploratory search - essentially a way to gather ideas in this first phase of search - could be useful for them.

One of the problems, however, with the group of media professionals is that there are some external influences that can affect the search process to a great extent, which can have consequences for the usefulness of exploratory search for them. The factors that we found were time, budget, and target audience of a program, which are work-related constraints that were also found in the research by Sauer (2016). Also they require transparency when searching. These factors determine to a large extent whether an exploratory search tool will be used and in what ways. Especially the work-related constraints of time, budget and target audience of a program imposes limits on the search process of media professionals: when there is little time and budget, the broad information need at the start of the search needs to be resolved as soon and as efficient as possible. This was nicely illustrated by a group in the poster session (WS_MP: poster_1), who, concerning the restrictions of time and budget, stated that *"this is important, so there is a budget available, and for others this is less important, so then we use the pool of recent news, then it also cannot cost any licences. So in that case you already have much more specific search queries"*.

It should be noted that these constraints are very different depending on the type of media professional: people that work for the evening news usually have a very short time span to finish a story, while documentary makers have a lot more time to gather ideas. Likewise, we have seen that the material sought depends on the target audience of a program. A media professional that works on a program geared at children wants to find recognizable images, while a media professional that works on a program geared at adults is rather interested in original material. A search tool should be able to cater to these differing information needs

Another external factor that influences the search process is trust, and related to this, transparency: they need to be sure that the information presented to them is correct, or that it can be verified in some way. This need for transparency is logical when taking into consideration that for media professionals, such as journalists, telling fact-based stories is their *raison d'être*, and it will become even more important in the future (Van der Haak, Parks & Castells, 2012). As such, being able to check the credibility of sources and verifying information are essential. A search tool should therefore assist them in these occupations. If this is not done correctly, issues of trust emerge, therefore limiting the usefulness of the tool for media professionals. The importance of transparency and trust in the search process can be illustrated by a comment made by a participant during the focus group (WS_MP: focus_3): *"the first thing you hear: don't trust anyone, nothing. Everyone has a hidden agenda. And you have to find it. A database also has a hidden agenda, so to speak. An algorithm, how does that work here. Can I trust it. [...] We want to see it with our own eyes"*. This requirement leads us to the first identified issue in DIVE+, namely the lack of transparency.

6.2.2 Lack of transparency

One of the most prominent issues in DIVE+ identified by the media professionals was a lack of transparency at different levels in the tool, mainly related to how events are represented in DIVE+ and how they are presented to the user. This is where the problem lies: DIVE+ did not offer enough clues to the media professionals that allowed them to 'check' the information. It was not evident where the information was coming from and why it was presented to them. Particularly, participants needed more transparency on the following aspects:

- How are relationships between entities generated
- What is the rationale behind the labels given to entities
- Metadata and information that describes individual entities

Lack of transparency: relationships

The media professionals stressed that they wanted to know how one entity is related to the other. In order to use the tool effectively, information about the relations should be made available and should be transparent, so that it is clear to the user why certain results are presented to them. Currently, DIVE+ does not provide this information. The unclarity of the search results and the relations between them was mentioned in the focus group (WS_MP: focus_6): *"what's behind it. On the basis of which concepts, which tags, is a relationship generated?"*. Another participant (WS_MP: focus_1) mentioned a Linked Data search tool that he once used, which provided more information about the relationships and as such gave him more inspiration: *"more data triplets were shown there. The entities and the relations between them were shown explicitly, it was stated that this is the one, this is the other, and this is the relationship between them. That gave me already more inspiration."*

Lack of transparency: entity labels

Another thing that was not clear to the media professionals was how entities were given their labels (e.g. concept or actor). One respondent (WS_MP: questionnaire_2) stated that, in answering a question about the usefulness of events and event entities, that it is *"unclear why search results get certain tags. That makes it impossible to ascribe value to it"*. The fact that the majority of the media professionals, 63%, did not understand the meaning of the different characteristics of events supports this finding. Again, they wanted more transparency about why results are presented to them: *"you say making it easy, is actually showing exactly how it is composed, the results, then you make it easy. Because then you are able to get insights, then you can trace it back"* (WS_MP: focus_5).

Lack of transparency: object/entity information

When the users were interested in one particular result and clicked it to retrieve more information about it, they were often disappointed by the minimal metadata available that describes this object. A participant (WS_MP: think aloud_3) mentioned: *"this is something I came across often, that nothing is there. Not an image, but also not a description. A bit more descriptions/metadata would be desirable, so that you know why certain things/events are being linked."*

6.3 Limited user control

Another limitation of DIVE+ mentioned by the media professionals was the limited amount of control they could exercise to tailor the search tool to their specific needs. Filtering options were missed and more control was requested concerning the exploration path.

Filtering options

Most prominently, the fact that they could not filter on the type of media and the collection was considered as a drawback. Their background varied and one can imagine that some media professionals are more interested in images, while others are more interested in videos. In the think-aloud protocol and in the questionnaire, comments were repeatedly made that more filtering options should be available. For example, a participant (WS_MP: think aloud_4) said: *“can you select on image? I would find that very useful, all those sources in there, it only expands. And there I want to exclude things as well”*. These filtering requirements also closely reflect the work-related constraints of media professionals discussed previously. Particularly, the type of program (e.g. Youth news or Evening News) affects the type of material that is sought for, and media professionals wanted to have more options in DIVE+ to tailor to these needs, even in this initial stage of search.

Exploration path control

More user control was also requested for the exploration path by the media professionals. For example, a respondent (WS_MP: questionnaire_2) stated *“it would be more useful if DIVE+ would not save everything by itself, but only on the request of the user”*. Furthermore, the media professionals mentioned that they wanted more control over the exploration path so that only ‘relevant’ results would be saved. A similar result is visible for the humanities scholars, and they more specifically go into details of how to achieve this adaptability of the exploration path, such as renaming and dragging and dropping functionalities.

Currently, in DIVE+ there are too little ways of how the user can take control over, or adapt the search experience. The searchers expect such functionalities, and if they are not there, this limits the user experience.

6.3 Do events, narratives and serendipity support media professionals?

6.3.1 Utility of event modeling

The current representation of events in DIVE+ was not considered as very useful to the media professionals. The aforementioned issues, mainly regarding transparency, are an important cause of this attitude. Interestingly, the student groups were less critical than the media professionals about the current usefulness of events in DIVE+. Perhaps this is because media professionals have more experience with collecting and verifying information, as these are important parts of their job. Transparency in how events are generated and linked are therefore highly valued, while this is more taken for granted by the student groups. This is also evident from the result that 86% of the media professionals missed characteristics of events, while this was only the case for 23% of the computer science

students and 56% of the (digital) humanities students. Especially for the media professionals it became clear that the current implementation of the event model is limited because essential information is missing which allows them to make sense of it. This finding is especially relevant because sense-making and contextualization are considered the added value of DIVE+ by the DIVE+ research community, but the way that events are currently presented to the users does not contribute to this. Exploring in DIVE+ therefore does not lead to enhanced information interpretation.

6.3.2 Added value of narratives

Media professionals were most critical concerning the suggested narratives by a search tool. Moreover, they did not discover narratives in DIVE+ and it did not help them to explore topics in depth. The aforementioned transparency issues could be an important reason for this: they simply could not reasonably connect different events together and as such, they also did not find narratives. Furthermore, explanations could lie in the search requirements that we identified for media professionals: they work with time constraints, but at the same time, they need to deliver (material for) a story that can be trusted and verified. They thus need to be sure that the information, and in this case the narratives, are correct and make sense, because it needs to be told to a wide audience and there is little time to fact-check. DIVE+ was not effective for them in this regard. It is true that the other studied groups of scholars and students also have to deal with time constraints, but these are less restrictive and immediate, because they are usually engaged in research projects that span multiple months, or even years. The work of media professionals, on the other hand, is more output-oriented and immediate. They, for example, have to deliver a story for a news item in a short amount of time. Their search process is therefore often much shorter than the search process of scholars and students. So, it could be argued that for media professionals, it is of uppermost importance that a search tool is effective in resolving immediate information needs efficiently, and that the information can be trusted. It seems that the narratives in DIVE+ are currently not supporting this need.

Furthermore, the limited user control concerning the exploration path is also limiting the potential of this feature in the formation of narratives. It is currently seen as a way of saving the thought process, where too much irrelevant information is saved, even though entities can be deleted manually. The value of the exploration path in the formation of narratives is therefore questioned. However, several respondents noted that the exploration path has potential and could be interesting, but currently it is not. For example, one of the respondents (WS_MP: questionnaire_8) stated *“the exploration path could be interesting if the basic functionalities would be ok. It could be interesting to share with colleagues which searches you have conducted, which paths you have taken. But this is a need which only surfaces when you were able to search well in the first place”*

6.3.3 Serendipity

The identified problems of a lack of transparency and a lack of user control also impose limits on the serendipitous browsing path of the users. When looking at opinions of the respondents

on the interface, a recurring finding is that there are too many search results, which cannot be filtered, and therefore it is difficult to see the relevance. For example, one of the media professionals argued: *“The search results are a big mess. It is difficult to assess the value of the results”* (WS_MP: questionnaire_2). Right now, the user has little means to ascribe value to the search results and the links between them. The usefulness of the results during browsing is therefore difficult to determine. So, while DIVE+ does provide unexpected results, it currently is not able to ensure that this is of use in the search process of the users.

6.4 Recommendations for DIVE+

The previous discussion shows that DIVE+ currently is limited in its ability to facilitate understanding and interpretation of archival collections to its users. Regarding the media professionals, the requirements that they have for searching are not supported by the different features in DIVE+. Their requirements for an exploratory search tool can be summarized as follows:

- It should be time-efficient to use
- There should be enough information available that describes collections and objects
- The information should be transparent and verifiable
- There should be options to tailor the search to the specific (needs of the) media professional.

Some of these demands can easily be solved by implementing small changes in the user interface of DIVE+. Other demands, such as that it should be time-efficient to use, are more difficult, or even undesirable, to meet. After all, the aim of DIVE+ is not to provide quick solutions to information problems, but rather to support users in their exploration of cultural heritage collections. In adapting the search browser to new user groups, developers should keep this in mind. Still, we think that a lot of the issues can, and should, be solved so that DIVE+ is better able to fulfill its potential. Therefore, some recommendations for DIVE+ (mainly regarding its user interface) are proposed, which can be found in appendix H.

7. CONCLUSION

This research presented an extensive investigation of the search needs of media professionals by using DIVE+ as a use case. Various user studies were conducted, both with media professionals as with other groups, that allowed us to see how their search needs differ and to what extent an event-based exploratory search tool fits their needs when they search through cultural heritage collections.

There were some limitations to this research. First, the time available for actual testing with DIVE+ during both workshops was rather short. The participants had around 15-20 minutes to actually use the interface. This is quite limited, especially when considering that the participants were not familiar with the tool. Their opinions are thus based on this short and rather shallow experience. While we were able to retrieve some useful insights out of this, future researchers are encouraged to devote more time to testing with users, for example through longitudinal studies where the same group of users participate multiple times.

More profound insights could be retrieved from these tests. Another limitation is the slightly different set-up of the user studies, which diminishes their comparability. For example, while the media professionals quite extensively discussed their existing search practices in the poster session, this was not done in the session with humanities scholars, because of time limitations. Comparing the two groups on this aspect was therefore not possible.

In investigating the search practices of media professionals, we found a set of requirements, relating to their work practices, that they have for searching: a tool should be time-efficient, there should be enough information available in the tool and this should be transparent, and there should be enough options to tailor the search. These requirements could not be fulfilled by the exploratory search browser DIVE+. Particularly, the way in which events and event entities were presented provide users with too little means to understand and grasp why it was presented to them in this way. This results in a lack of trust and a skeptic attitude, which is problematic as it has negative consequences for the usefulness of the constructed narratives and the serendipitous findings. All in all, this let us conclude that currently, narratives, events and serendipity in DIVE+ are not supportive to the exploratory search needs of media professionals.

The humanities scholars and students also found DIVE+ not optimal for exploring, partly for the same reasons given above. However, media professionals were very critical about the usefulness of suggested narratives for exploring a topic in depth, while the other groups were not very critical. This is a remarkable difference that is difficult to explain. Perhaps, because their jobs revolve around the production of fact-based stories, media professionals do not trust the narratives presented by a search tool, especially if they cannot trace it back. Still, the difference with particularly the humanities scholars is notable because the work ethics of this group also heavily involves truthful and reliable reporting, and one would assume that they thus have similar requirements regarding suggested narratives as the media professionals. Future researchers are encouraged to shed more light on this difference.

Right now, the main challenge that lies ahead for the DIVE+ research community is to increase the transparency of the tool, so that users are better able to understand what is presented to them, and why it is presented to them. The current research has proposed some recommendations for the interface that should give a first step in the right direction. Future user studies should test whether these features have their intended outcome. We are confident that, once the recommendations are implemented, DIVE+ is better able to support users in their explorations through cultural heritage.

8. ACKNOWLEDGMENTS

This work has been made possible with the help of the DIVE+ project team. In particular, I would like to thank Lora Aroyo, Sabrina Sauer and Berber Hagedoorn for their guidance and feedback. Also, during this research project, I closely worked together with Minyi Cheng. It was a joy to work with her, and I would like to thank her for the good cooperation.

9. REFERENCES

- Borlund, P. (2016). A study of the use of simulated work task situations in interactive information retrieval evaluation: A meta-evaluation. *Journal of Documentation*, Vol. 72 Iss 3 pp. 394-413.
- Collijn, A. (2016). *DIVE+: An exploratory search system for Humanities scholars*. (Master's thesis, Vrije Universiteit, Amsterdam).
- Cheng, M. (2017). *DIVE+: User Interaction Design with Digital Humanities*. *Forthcoming*
- De Boer et al. (2017). Enriching media collections for Event-based exploration. *Forthcoming*
- De Boer, V., Oomen, J., Inel, O., Aroyo, L., Van Staveren, E., Helmich, W., & De Beurs, D. (2015). DIVE+ into the event-based browsing of linked historical media. *Web Semantics: Science, Services and Agents on the World Wide Web*, 35, 152-158.
- Hagedoorn, B. and Sauer, S. "Getting the Bigger Picture: Evaluating Narrative Creation and Media Exploratory Search for AV Research" (in preparation, 2017).
- Kemman, M., Kleppe, M., Nieman, B., & Beunders, H. (2013). Dutch Journalism in the Digital Age. *Icono 14*, 11(2), 163-181.
- Kruijt, M. (2016). *Supporting exploratory search with features, visualizations, and interface design: a theoretical framework*. (Master's thesis, Universiteit van Amsterdam, Amsterdam).
- Kules, B., & Shneiderman, B. (2008). Users can change their web search tactics: Design guidelines for categorized overviews. *Information Processing & Management*, 44(2), 463-484.
- Kules, B., Capra, R., Banta, M., & Sierra, T. (2009, June). What do exploratory searchers look at in a faceted search interface?. In *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries* (pp. 313-322). ACM.
- Lecheler, S., & Kruikemeier, S. (2015). Re-evaluating journalistic routines in a digital age: A review of research on the use of online sources. *new media & society*, 1461444815600412.
- Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, 49(4), 41-46.
- Nielsen, J., Clemmensen, T., & Yssing, C. (2002, October). Getting access to what goes on in people's heads?: reflections on the think-aloud technique. In *Proceedings of the second Nordic conference on Human-computer interaction* (pp. 101-110). ACM.
- Sauer, S. (2016). Searching for a story: Creative retrieval practices of media professionals. Paper presented at the IAMCR 2016 conference, Leicester, UK.
- Sauer, S., & de Rijke, M. (2016, July). Seeking Serendipity: A Living Lab Approach to Understanding Creative Retrieval in Broadcast Media Production. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 989-992). ACM.
- Schifferes, S., Newman, N., Thurman, N., Corney, D., Göker, A., & Martin, C. (2014). Identifying and verifying news through social media: Developing a user-centred tool for professional journalists. *Digital Journalism*, 2(3), 406-418.
- Toms, E. G. (2000, December). Serendipitous Information Retrieval. In *DELOS Workshop: Information Seeking, Searching and Querying in Digital Libraries* (pp. 17-20).
- Vassilakaki, E., Garoufallou, E., Johnson, F., & Hartley, R. J. (2014). Users' Information Search Behavior in a Professional Search Environment. In *Professional search in the modern world* (pp. 23-44). Springer International Publishing
- van den Akker, C., van Erp, M., Aroyo, L., Segers, R., Van der Meij, L., Legêne, S., & Schreiber, G. (2012). Understanding objects in online museum collections by means of narratives. In *Proceedings of the Third Workshop on Computational Models of Narrative (CMN'12)*.
- Van Den Akker, C., Legêne, S., Van Erp, M., Aroyo, L., Segers, R., van Der Meij, L., ... & Jacobs, G. (2011, June). Digital hermeneutics: Agora and the online understanding of cultural heritage. In *Proceedings of the 3rd International Web Science Conference* (p. 10). ACM.
- Van Den Akker, C., van Nuland, A., van der Meij, L., van Erp, M., Legêne, S., Aroyo, L., & Schreiber, G. (2013, May). From information delivery to interpretation support: evaluating cultural heritage access on the web. In *Proceedings of the 5th Annual ACM Web Science Conference* (pp. 431-440). ACM.

Van der Haak, B., Parks, M., & Castells, M. (2012). The future of journalism: Networked journalism. *International Journal of Communication*, 6, 2923-2938.

Beyond the query-response paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1), 1-98.

White, R. W., & Roth, R. A. (2009). Exploratory search:

Appendices

Appendix A: Protocol workshop CreateSalon (Humanities scholars)

- Online available at: <https://docs.google.com/document/d/1u78OhHeHCh2QGA3KKwkAn-DzdtVenAtmSoJCDm7VQjk/edit>

Appendix B: Protocol workshop MediaNow (Media Professionals)

- Online available at: <https://docs.google.com/document/d/1dwq-laOHyn-nfSaCneR0Gqar17pe62K0uKK63crCZYZ/edit>

Appendix C: Task description MediaNow workshop (media professionals)

- Online available at: <https://docs.google.com/document/d/1c2usGrKygaLWZr0TVF7F8lHaRyCFEiuhjDTv1fm7j3o/edit>

Appendix D: Questionnaire MediaNow (Media Professionals)

- Online available at: <https://docs.google.com/forms/d/1IGqA0bTJtFtgjGs4vqQOwLMOFVJU342h-VVEXb6ZYqM/edit>

Appendix E: DIVE+ transaction logs CreatSalon (Humanities scholars)

- Online available at: https://docs.google.com/document/d/1ZmUAwoh1FtKCb-O_DTP53cz2MkU54ZPq-IDYyTQ3S7g/edit

Appendix F: DIVE+ transaction logs MediaNow workshop (Media Professionals)

- Online available at: <https://docs.google.com/document/d/1fch2ipNdSPnkkuaSzFLRSq7n9KkvqQLE4FwCMwEiJKM/edit>

Appendix G: results

- Online available at: <https://docs.google.com/document/d/1Uq4eGAmAD7oWa3zX6fQPWLEI5yVkhj-1I-Ku3SRo8sk/edit>

Appendix H: Recommendations for DIVE+

- Online available at: <https://docs.google.com/document/d/1-yCKdMI27UMeH1KLVCmG488kiJAqNGneIRVYJmbP7OY/edit>

Appendix I: Answers of the four groups of respondents on the closed-ended questionnaire questions about DIVE+ in general

- Online available at: https://docs.google.com/drawings/d/1zARBxW3M5y6fYFk33OAJe_uoGLsviCsen7Xrx6zyAeE/edit

Online Appendix:

- **Results: spreadsheet of the user studies results (search practices of media professionals):**
https://docs.google.com/spreadsheets/d/1rDOjH9cdNs0PkiFNDYyD88Fxdty10FB_ev7tRHd-E/edit#gid=0
- **Results: spreadsheet of the user studies results (DIVE+ usage of all groups):**
<https://docs.google.com/spreadsheets/d/1xL3sNlhOaCrAZAOP-s1pIkoOi6B1d0ewDFmhcs2bSIQ/edit#gid=0>

Appendix A: Protocol workshop CreateSalon (humanities scholars)

Date: 9 May 2017

Location: UvA Turfdraagsterpad 15, Amsterdam

Time: 15.00 - 16.00

Duration workshop: 1 hour

15.00 – 15.05 Introduction/presentation

(5 minutes)

Sabrina Sauer

15.05 - 15.25 Performing a task with DIVE+ ([Digital Humanities search task](#))

(20 minutes)

Sabrina Sauer, Berber Hagedoorn, Justin Verhulst & Minyi Cheng

15.25 - 15.35 Questionnaire

(10 minutes)

All

15.35 – 16.00 Focus Group of search case/task

(25 minutes)

Sabrina Sauer and Justin Verhulst

Appendix B: Protocol workshop MediaNow (Media Professionals)

Date: 24 May 2017

Location: Netherlands Institute for Sound and Vision, Hilversum

Time: 09.00 - 12.00

Duration workshop: 3 hours

9.00 – 9.15 Walk-in

9.15 - 10.00 Introduction - Presentation update MediaNow

Maarten de Rijke, Sabrina Sauer, Ilya Markov and Nikos Voskarides

10:00-10:10 Questions about the project updates

10.10 - 10:50 Scenario drawing (poster session) of search flow

Sabrina Sauer and Justin Verhulst

11:00 - 11.30 Use DIVE+ Task Performance and questionnaire (all fill the questionnaire)

Sabrina Sauer and Justin Verhulst, Minyi Cheng

11.30 - 12:00 Overall discussion with all participants about their search behavior

Sabrina Sauer and Justin Verhulst, Minyi Cheng

Appendix C: Task description MediaNow workshop (media professionals)

Exploring linked open data with DIVE+:

searching for audio-visual sources to discover ideas for programs and create narratives

Thank you for testing the exploratory search engine DIVE+.

While performing the tasks with DIVE+ please consider the following:

- talk aloud when you're interacting with the DIVE+ browser (only when you are approached by one of the interviewers)
- provide us with any feedback, both positive and negative, you have.

To start your DIVE+ session go to <http://diveplus.frontwise.com/#live> in the browser.

Task: Imagine that a media company is going to produce programs about Jakarta, Beatrix, Islam and Watersnoodramp. Your goal is to propose an interesting angle for one of the programs.

- Use DIVE+ to explore your topic of interest (please choose one between: Jakarta, Beatrix, Islam or Watersnoodramp).
- DIVE+ keeps track of your search history with a so-called 'exploration path', which you'll see on the left side of the screen. Make sure to include at least five related entities in this exploration path, as this is a way for you to capture your potential research angles.
- Note that DIVE+ contains audiovisual material from diverse cultural heritage institutions. There are no recent materials available - most materials come from the time period of 1980s and earlier. Keep this in mind when conducting your research.
- When you are done, please export your exploration path (see image) and upload it to the following folder, with this format: [exploration_path_<yourname>]



Questionnaire about DIVE+

Finished? Please fill in the questionnaire in the other browser tab to let us know your experiences - comments or suggestions for improvements are highly appreciated.

Appendix D: Questionnaire MediaNow (Media Professionals)

SECTION: Share Your DIVE+ usage experience

Thank you for taking part in the DIVE+ testing! We hope it was a pleasurable experience. Your feedback and comments are very important to us, in order to keep improving the DIVE+ exploratory search engine. Below, we would like to ask you some questions about your experiences with DIVE+.

SECTION: DIVE+ task performance

During the workshop, you have received a task to perform with DIVE+. In the following questions, we'd like to know a bit about your familiarity with the topic you performed your research on.

Which topic did you do your DIVE+ research on?

- Jakarta
- Beatrix
- Islam
- Watersnoodramp
- Other

Are you interested in this topic?

- Yes, I am interested in the topic
- I am not particularly interested, but the topic is OK
- No, I am not really interested in this topic

How much did you know about this topic before this session?

- I knew a lot about this topic before I started with DIVE+
- I knew something about it, but not much
- I didn't know anything about this topic before I started with DIVE+
- Other

Were you able to come up with an angle for your chosen topic with DIVE+?

- Yes
- No

SECTION: Finding an angle with DIVE+ didn't work

Please tell us more about your experience using DIVE+ for finding your program angle, and tell us why you were not able to come up with an angle by using DIVE+

SECTION: Finding an angle with DIVE+ worked.

Please tell us more about your experience using DIVE+ for finding your program angle, and tell us what angle you found.

SECTION: Explorative Search Tools

Besides DIVE+, there are also other exploratory search tools. We are interested in your experiences with these other tools. In the following questions, please tell us about your experience with other explorative search tools.

Are you familiar with other exploratory search tools (except DIVE+)?

- Yes, I am familiar with exploratory search tools
- Yes, I am familiar with exploratory search tools, but I have never used them
- I am not familiar with exploratory search tools

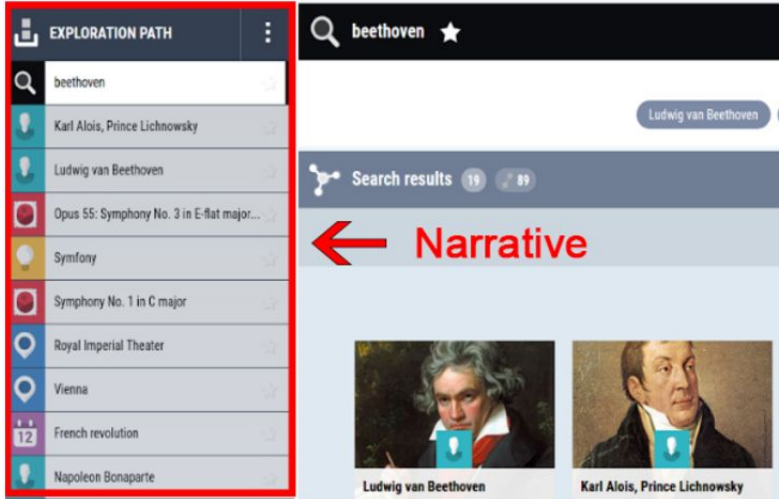
Which exploratory search tools have you used?

How frequently do you use exploratory search tools in your daily job?

- Often
- Sometimes
- Never

SECTION: DIVE+ Narratives

The following questions relate to the narratives within DIVE+. Narratives can be seen as stories or perspectives around objects or topics, constructed by the events that are related to those objects or topics. For example, a battle in a war could be looked at from the angle of the military or the citizens. This would give two different perspectives (narratives) on the same event, which means it would include different other related events, people or locations. In DIVE+, the exploration path on the left visualises the narrative you follow while searching (see example below, the questions continue after the image).



Please indicate your opinion on the following statements (disagree / neutral / agree)

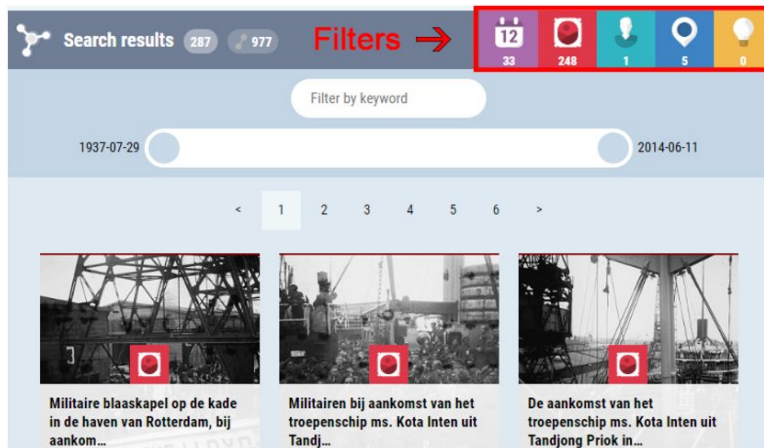
- When searching in DIVE+, I discovered different narratives.
- When researching a topic, suggested narratives are helpful to explore this topic in depth
- When researching a (media) event, suggested narratives are helpful to learn more about this event.
- The exploration path in DIVE+ helped me to learn about historical events
- I find the exploration path in DIVE+ a useful feature.

Please elaborate on your choices above.

Do you have any recommendations for improvements with respect to the DIVE+ exploration path?

SECTION: DIVE+ events

In DIVE+, historical events are linked with media objects and are furthermore identified by their person, location, and concept. For example, a painting that depicts a soldier in a battle could be characterized by the name of this soldier, the place the battle was fought, and the 'concept' of this event (e.g. a battle). In DIVE+, you can filter on these characteristics (see example image below). In this section, we'd like to ask you about the usefulness of these event characteristics.



Please indicate your opinion on the following statements (disagree / neutral / agree).

- During the use of DIVE+, I understood the meaning of the different characteristics of events (media object, person, location, and concept)
- The different characteristics of historical events in DIVE+ (media object, person, location, and concept) helped me to learn about historical events.
- I find the current event representation in DIVE+ useful when researching my topics of interest.
- I missed some characteristics of events during my use of DIVE+.

Please elaborate on your choices above.

Do you have any recommendations for improvements with respect to events in DIVE+?

SECTION: Evaluation of DIVE+ and DIVE+ User Interface

Please share your opinion and usage experience with DIVE+ in general.

Please indicate your opinion on the following statements (disagree / neutral / agree)

- DIVE+ is useful for the exploration of my topics.
- DIVE+ enabled me to generate ideas for a story
- I would use DIVE+ in my daily work.
- I had a pleasant experience with the user interface of DIVE+.

Please elaborate on your choices above.

Do you have any recommendations for improvements of DIVE+ in general?

SECTION: Tell us a bit about yourself

Thank you for filling in this questionnaire! Before you submit, please provide us with some of your personal details. Your personal data will be anonymized and strictly used only in an aggregated way to explain the background and demographics of the study participants. If you consent to it, we would like to contact you for follow up communication about this study. We will not share your personal data for any other purpose outside of the scope of this project.

What is your name?

What is your age?

- 20 to 34
- 35 to 49
- 50 or older

What is your occupation?

What is your specific domain expertise?

How much experience do you have with linked data?

- A lot
- some

- Little

If you are interested to hear about the results of this study, as well as to participate in future research about DIVE+, please let us know your email

Appendix E: DIVE+ transaction logs

About this data: The documents below list the log data captured during the CreateSalon workshop. The data was retrieved by means of the DIVE+ action log⁷ and by looking at the final exploration path, which the participants were asked to upload. The amount of unique sessions (identified by the session ID) is not equal to the amount of people attending the workshops. This is due to the fact that some participants did not upload their exploration path, and through the DIVE+ action log, not all sessions could be retrieved or identified correctly.

CreateSalon workshop (humanities scholars)				
Session ID	No. of queries	No. of entities	Length of exploration path (queries + entities)	Queries
2f5ragytrl	3	0	3	1 Islam 2 art 3 ceramic
1ykxti385y	11	5	16	1 Beatrix 2 schaatsen 3 Beatrix 4 Beatrix paleis 5 islam integratie criminaliteit 6 islam 7 beatrix 8 button 9 Beatrix kroning 10 Beatrix 11 geen woning geen kroning
16n48h81lw	3	19	22	1 islam 2 chinese moslims 3 zijeroute
c1oxrzrpmu	8	9	17	1 islam 2 china 3 portret 4 portret moslim 5 portret islam china 6 portret islam 7 marabout 8 cheikh amadou bamba
8khoucmx5p	2	5	7	1 watersnoodramp 2 deltaplan
28cbqp2g2y	3	2	5	1 watersnoodramp 2 zeeland 3 1953
ljkt5tvko2	3	3	6	1 Jakarta 2 Beatrix 3 Wim Kok
cl9iva5etw	1	10	11	1 Islam
blytninuq5	3	6	9	1 Jakarta 2 Jakarta Cultuur 3 Jakarta

⁷ http://diveplus.frontwise.com/logs/Y2fPEHdjmgBGP_dive.log

b2n2rs1onc	12	3	15	1 kroningsrellen 2 kroning beatrix 3 beatrix 4 jakarta 5 Islam 6 nalatigheid 7 dijkdoorbraak 8 ceramic 9 wim kok 10 spinoza 11 Rotterdam 12 rembrandt
AVERAGE	4.9	6.2	11.1	

Appendix F: DIVE+ transaction logs MediaNow (Media professionals)

About this data: The documents below list the log data captured during the CreateSalon workshop. The data was retrieved by means of the DIVE+ action log⁸ and by looking at the final exploration path, which the participants were asked to upload. The amount of unique sessions (identified by the session ID) is not equal to the amount of people attending the workshops. This is due to the fact that some participants did not upload their exploration path, and through the DIVE+ action log, not all sessions could be retrieved or identified correctly.

MediaNow workshop (Media Professionals)				
Session ID	No. of queries	No. of entities	Length of exploration path (queries + entities)	Queries
3u8tzun1fu	14	4	18	1 islam 2 terrorisme islam 3 daroel islam 4 westerling 5 Raymond Westerling 6 Wim Kok 7 Ruud Lubbers 8 batavia jakarta 9 jakarta and not foto 10 video jakarta 11 Rotterdam 12 krakersbeweging 13 watersnoodramp 14 video
24ek97p7pg + 1b4dep83pm	12	1	13	1 watersnoodramp 2 koeien 3 watersnood AND ooggetuigen 4 watersnood inwoners 5 watersnoodramp inwoner 6 watersnoodramp 7 Ruud Lubbers 8 watersnood ooggetuige 9 watersnood 10 watersnood 1953 11 zeeland 1953 12 zeeland and 1953
bd tj8pnahw	13	18	31	1 Beatrix 2 kroning beatrix 3 beatrix 4 woningnood beatrix 5 verstorning inhuldiging beatrix 6 inhuldiging beatrix 7 prinses beatrix 8 beatrix 9 woningnood 10 krakersbewegin 11 geen woning geen kroning 12 krakers 13 beatrix
tu8uy2eu2a + 78fdvt77qd + 8q0v8c6avh	6	6	12	1 jakarta 2 jakarta 3 jakarta onafhankelijkheid 4 batavia

⁸ http://diveplus.frontwise.com/logs/Y2fPEHdjmgBGP_dive.log

				5 jakarta 6 batavia jakarta
98fyhzmagi	20	7	27	1 Jakarta 2 Indonesie 3 watersnoodramp 4 opstand 5 Siloengkang 6 knil 7 jakarta onafhankelijkheid 8 onafhankelijkheid 9 opstand 10 ramp 11 verhuizing 12 geen woning geen kroning 13 watersnoodramo 14 watersnoodramp 15 rotterdam 16 voorburg 17 hilversum 18 jakarta 19 watersnoodramp 20 beatrix
2l2lft2d9i	10	6	16	1 watersnoodramp 2 islam 3 bekering 4 bekering islam 5 bekeren 6 islam 7 bekering 5 islam video 6 islam 7 watersnood 8 watersnoodramp 9 1953 10 1959-01-21
2716k4owcx	7	0	7	1 België 2 Beatrix 3 Koningin Beatrix in België 4 Beatrix België 5 Kasteel 6 Brussel Beatrix 7 Beatrix radio
AVERAGE	11.7	6	17.7	

Appendix G: Results


	CreateSalon workshop: humanities scholars	MediaNow workshop: media professionals	(Digital) humanities students	BSc computer science students ⁹
Demographics	No. of participants: <ul style="list-style-type: none"> - workshop: 18 - questionnaire: 11 - transaction logs: 10 Background of participants (questionnaire): <ul style="list-style-type: none"> - Scholars: 6 - Students: 5 Linked data expertise: <ul style="list-style-type: none"> - Little: 18.2% - Some: 72.7% - A lot: 9.1% 	No. of participants: <ul style="list-style-type: none"> - workshop: 11 - questionnaire: 8 - transaction logs: 7 Background of participants (workshop): <ul style="list-style-type: none"> - Broadcasting organisations: 5 - Libraries/archives: 6 Linked data expertise: <ul style="list-style-type: none"> - Little: 25% - Some: 50% - A lot: 25% 	No. of participants: <ul style="list-style-type: none"> - questionnaire: 16 and less Linked data expertise: <ul style="list-style-type: none"> - Little: 18.8% - Some: 81.3% - A lot: 0% 	No. of participants: <ul style="list-style-type: none"> - questionnaire: 22 and less Linked data expertise: <ul style="list-style-type: none"> - Little: 35% - Some: 55% - A lot: 10%
Actual activities (numbers are averages)	No. of queries: 4.9 No. of entities: 6.2 Length of exploration path (queries + entities): 11.1	No. of queries: 11.7 No. of entities: 6 Length of exploration path (queries + entities): 17.7	N/A	N/A

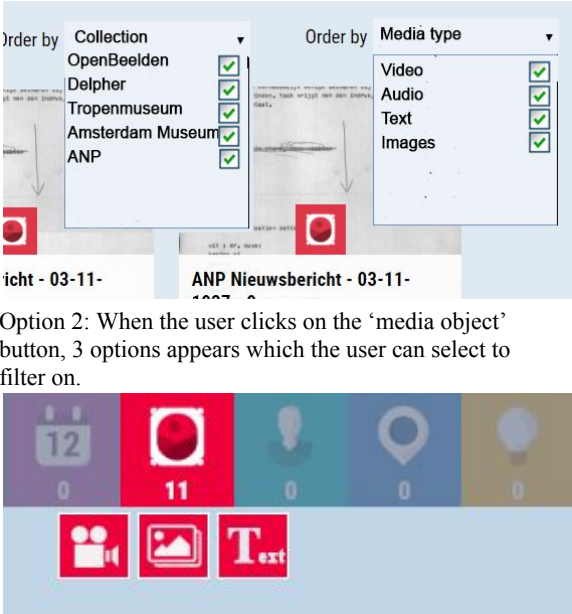
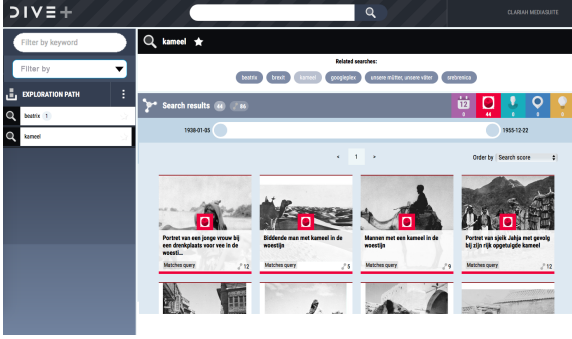
	Result	HSC	MP	DHS
Effectiveness and efficiency of DIVE+ for performing tasks	More information is needed about which collections are in DIVE+	X	X	
	Respondents needed more (meta)data and topic descriptions.	X	X	
	More information is needed about how relationships between entities are generated		X	
	Respondents needed more information about why certain labels (e.g. actor, concept) are given to entities.		X	
	More fine-grained sorting of the media types is needed	X	X	
	More fine-grained sorting of the collection is needed		X	X
	Respondents find that the exploration path does not result in narratives	X	X	X
	The amount of data was experienced limited	X	X	

⁹ Because of time limitations, the answers to the open-ended questionnaire questions could not be analysed for the computer science students, so we do not include this group in the results table.

	Respondents had trouble with the event characteristics	X	X	X
	For the respondents it was not clear how the exploration path is constructed	X		
	Respondents would like to see more available content/information for people who don't speak Dutch			X
General reflections on DIVE+ UI	More interaction with the exploration path needed	X	X	
	For respondents it was unclear what the meaning was of some buttons/icons	X		
	Respondents needed time to figure out the interface	X		
	The 'related searches' was experienced not related to the given query	X	X	
	Respondents had trouble using DIVE+, they find it complicated to use as they also did not find the needed information.			X
	Respondents would like to see the exploration path tool made easier to use with more clarity and instructions			X

Appendix H: Recommendations for DIVE+

Finding	Recommendation	Solution domain: DATA / API / UI
<p>More information is needed about which collections are in DIVE+</p>	<p>When the user clicks on a results, a screen with more information appears. Under ‘sources’, provide details about the collection the result originates from</p> <p>On the startpage of DIVE+, include a section (similar to ‘statistics’) about the collections that are in DIVE+. Hereby an example how it can be included on the startpage.</p> <p style="text-align: center;">Collections</p> <p style="text-align: center;">The following collections can be found in DIVE+</p> 	<p>DATA: Include the data API: retrieve the data UI: [done] already shows the returned ‘source’ fields. These are often empty.</p> <p>DATA: Include collection details API: return collection details UI: display collection details per entity</p>
<p>Respondents needed more (meta)data and topic descriptions, or the given information was not clear enough about the object.</p>	<p>When the user clicks on a results, a screen with more information appears. For some entities, there is very little information available. Add additional metadata by linking the database to other open databases or crowdsourcing.</p>	<p>DATA: include the information API: retrieve the information UI: [done] already shows all information fields. These are often empty.</p>
<p>Respondents needed more information about how relationships between entities are generated</p>	<p>When a user hovers over a result with the mouse, show how the entity is related to the query by showing the triple.</p> <p>On the result pages (after a user queries something), indicate in each result box how it is related to the query, by showing the triple.</p>	<p>DATA: extend data with correct relations API : return these relations UI: [done] the relation is already shown, multiple relations can be displayed</p>
<p>Respondents needed more information about why certain labels (e.g. actor, concept) are given to entities.</p>	<p>Provide more information, see above.</p>	<p>DATA: include entity type reason API: retrieve reason UI: display reason, maybe in a separate field? Or in the sources field?</p>
<p>More fine-grained sorting of the media types is needed</p>	<p>Option 1: Besides ‘order by’, add a ‘filter by’ collection’ / ‘filter by mediatype’ button. When the user clicks on it, a drop-down menu pops up (similar to the ‘order by’ button), in which the collections / mediatype are indicated. The user can turn individual collections or mediatypes on and off. In default mode, all options are on.</p>	<p>Option 4: include the type/collection filter in the filter bar of the related entities (near the keyword / date range filter) - in a way comparable to option 1</p> <p>DATA: include collection info in data API: retrieve collection info UI: make additional filters on collection level UI: make additional filter on media type: (only images or video, or none)</p>

	 <p>Option 2: When the user clicks on the 'media object' button, 3 options appears which the user can select to filter on.</p> <p>Option 3: Reduce the space for the exploration path on the left side and put the filters on the leftside above the exploration path. This will create less distraction on the results page.</p> 	
More fine-grained sorting of the collection is needed	See previous recommendation	UI: [done] you can sort now on title, date, type, #relations asc&desc
'Related searches' is not clear.	Adapt the back-end algorithm such that the related searches are only related to the current query, and make sure that the links generate at least some results when the user clicks on it.	Recommendation not clear, Not sure if this can be done. However it could be smart by the API to base the search ranking on the query; that would be feasible. Elements without data: this can be caused by: 1) search finding orphans 2) related query not returning any data because of a wrong formatted entity id.
Respondents struggled with the interface.	In general, the interface, and the results page in particular, should be less cluttered. The other recommendations solve this to some extent. The meaning of some buttons/icons were unclear to users. This can be solved by putting the title of the button in text	Have the users checked the popup titles? Icon + text combi is useful indeed! Could you indicated more precisely which

	together with the button.	buttons/icons should get this attention?
Respondents find that there is too little content, which results in UX problems	Obviously, add more content / data if this is possible. Try to minimize the amount of dead ends. For example, don't provide a 'related search' if this generates no results.	API/UI: Exclude orphans? Give orphans lower search ranking? This happens often when a library is loaded in the data. Many elements stay unlinked and indeed cause noise.
Need more interaction with the exploration path	Add the possibility for the user to change the name of all the blocks in the exploration path. Add drag-and-drop functionality, so that the user can change the order of the blocks Add possibility for the user to categorize queries together (colorizing, indented tree-structure)	Names in the path are based on the data. However with the annotation options this should be a relevant addition. UI: as the users already seem to struggle with the UI/data; wouldn't it be more useful to make the exploration path more simple (no stacking)? UI: are these features not already covered within the mediasuite (multiple result sets for special queries etc)
Respondents find that the exploration path does not results in narratives	This can be (partly) solved with the above recommendations	UI: good reason to add functionalities as described above to the path
Respondents are mixed in their attitudes concerning the current usefulness of the exploration path	This can be (partly) solved with the above recommendations	UI: this means the exploration path more or less loses its function as browsing history.

Appendix I: Answers of the four groups of respondents (1 = humanities scholars, 2 = media professionals, 3 = (digital) humanities students, 4 = Bsc computer science students) on the closed-ended questionnaire questions about DIVE+ in general

