

# Access to Archival Material in Context

Khairun Nisa Fachry<sup>1</sup> Jaap Kamps<sup>1,2</sup> Junte Zhang<sup>1</sup>

<sup>1</sup> Archives and Information Studies, Faculty of Humanities, University of Amsterdam

<sup>2</sup> ISLA, Faculty of Science, University of Amsterdam

## ABSTRACT

Archival finding aids are long and complexly structured documents describing archival material—the paper trails of the lives of corporate bodies, persons, and families. Currently, finding aids are encoded in XML using the standard Encoded Archival Description (EAD) and made available to the public on web-sites of archival institutions. But how to provide access to such long and complexly structured documents? On the one hand, users tend to look for specific archival material that may be deeply nested inside the archive. On the other hand, interpreting the meaning of an item is crucially dependent on its context.

Using insights from the field of XML retrieval—a subfield of information retrieval that has recently attracted a lot of attention, mainly through the annual evaluation effort in INEX—we developed three different systems for searching in collections of digital finding aids corresponding to three fundamental choices about archival access. The first system provides access to the fonds or archive as a whole; the second system provides direct access to individual archival material at any level of description; the third system retrieves archival material while preserving the original context. This paper reports on the results of an extensive user study with the three systems. Our main finding is that test persons have a preference for the third system that retrieves archival material in their original context, with test persons indicating that the system assisted them in assessing relevancy, navigation and direct access to relevant parts of the finding aids.

## Categories and Subject Descriptors

H.3 [Information Storage and Retrieval]: H.3.1 Content Analysis and Indexing; H.3.3 Information Search and Retrieval; H.3.4 Systems and Software; H.3.7 Digital Libraries

## General Terms

Measurement, Performance, Experimentation

## Keywords

Focused Retrieval, XML Retrieval, Archival Finding Aids

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*IiX'08, Information Interaction in Context*, 2008, London, UK.  
Copyright 2008 ACM 978-1-60558-310-5/08/10 ...\$5.00.

## 1. INTRODUCTION

Archival institutions have no arbitrary collections of documents, but the unique records of a particular person, family, or corporate body. Archival documents are the paper trails of individuals living their lives, or corporate bodies carrying out their functions. Documents sharing the same provenance are organized in the same fonds or archive. A fonds is defined as the whole of the records (recorded information in any form or medium) created and/or accumulated and used by a particular person, family, or corporate body in the course of that creator's activities and functions [12].

Archival finding aids are descriptions or means of reference created as the primary access tool, allowing researchers to gain access to and understand the material in which they would like to view and use for research. Archival descriptions adhering to the International Standard for Archival Description [ISAD, 12] are multilevel descriptions, proceeding in a top-down fashion, first describing the whole archive (or fonds), then its major sub-components, and so on. It results in a hierarchical structure ranging from the fonds, to subfonds, series, subseries, files, until the individual items. Currently, finding aids are increasingly encoded in XML [24] using the standard Encoded Archival Descriptions [EAD, 7], and made available at the web-sites of archival institutions.

Digital finding aids can significantly improve the access to archival material. However, these finding aids are long documents ranging in length from a few pages up to the size of a multi-volume book. How to provide access to such long and complexly structured documents? On the one hand, users tend to look for specific archival material that may be deeply nested inside the archive. On the other hand, interpreting the meaning of an item is crucially dependent on its context. The overall research question of this paper is to investigate how to provide access to archival material through digital finding aids.

We address this problem using some of the insights from the field of XML retrieval, a relatively new branch of information retrieval that studies focused retrieval in structured documents. The main thrust are the annual evaluation efforts of the Initiative for the Evaluation of XML retrieval [INEX, 11]. Focused retrieval techniques offer new ways for users to interact with finding aids. XML documents, such as finding aids encoded in EAD, consist of many elements (parts of the document delimited by matching start- and end-tags), and each element of the finding aid can be a unit to return to the users. Furthermore, the logical structure of the document has the potential to assist the user in providing direct access to relevant information by pointing to the relevant element rather than the whole document.

We developed three different systems for searching in collections of digital finding aids corresponding to three fundamental choices of archival access. The first system provides access to the fonds or

archive as a whole; the second system provides direct access to individual archival material at any level of description; the third system retrieves archival material while preserving the original context. This paper reports on the results of an extensive user study with the three systems. Specifically, we examine the following research questions:

1. Is direct access to archival material in digital finding aids deemed useful in comparison to access to the whole fonds?
2. Can the use of structural presentation improve users' satisfaction in locating relevant information, or help them in deciding if a collection is relevant?
3. Can individual items be used as entry points for accessing an archival description? If so, does it affect the users' satisfaction in accessing relevant information?

The remainder of this paper is organized as follows. In Section 2 we discuss related work about archival access, and access to structured (XML) documents. In Section 3, we discuss the three systems that corresponded to three different design choices: access to the whole fonds, direct access to archival material, and access to archival material in context. Then, in Section 4, we detail the setup of our interactive experiment with the three systems. The results of our study are discussed in Section 5. Finally, in Section 6, we conclude by presenting and discussing our findings.

## 2. RELATED WORK

There are two broad strands of related work. On the one hand concerning archival access, and on the other hand concerning access to structured documents with XML Retrieval.

### 2.1 Archival Access

Archival descriptions are founded on two fundamental principles: the principle of "provenance" and the principle of "respect for original order" [6]. Provenance states that records of the same origin or creator must not be intermingled with those of any other creator. This is frequently referred to as "*respect des fonds*." The principle of respect for original order states that archives of a single provenance should retain the arrangement (including the reference numbers) established by the creator in order to preserve existing relationships, and to provide evidential significance of the creator's functions and activities [1].

When migrating the archival description from a physical to a digital environment, archivists expressed the importance to maintain the integrity of their collection. As Hedstrom [8] reminds us,

"Provenance and the relationship between context and the content of records were considered to be long-standing pillars of archival theory and practice. In the electronic era, they are vital to description, because they provide the key to distinguishing records from non-record material; to understanding why, when, and by whom a document was created; and to determining the context in which the records was created, and hence its value and meaning."

Turning to some of the current archival literature on user studies of finding aids systems, it is revealed that there is little known about the information searching behavior of users with digital finding aids. Coats [4] explained that the lack of user studies in EAD has been a possible determinant to the more widespread acceptance of EAD among archivists. Rosenbusch [19] emphasized the importance of the user needs, commenting that the development of online

archival resources has been almost exclusively supply-side driven instead of demand-side driven.

Rosenbusch [19] further argued that studying navigational features and contextual information is important, because these features better help users to understand the collection. This argument is advanced in the study of Yakel [25] which focuses on how users interact with finding aids systems. Her study suggested that finding aids interfaces need to provide a way to a navigational aid that supports users in providing local detail and global view of the finding aids. This suggestion emerged because she found out that the users were lost in the hierarchy and were unsure as to where they were, especially in the full text view. Another related finding is presented by Stocking [20] in his examination of EAD as implemented in the Access to Archives (A2A), a virtual national archival catalog for England. His result showed that data elements are required for component levels of description. His experience with A2A confirmed that users engage finding aids for archival material from the bottom up and that we need to provide them with the fullest description necessary at those levels.

### 2.2 XML Retrieval

Since finding aids are encoded in XML, it is worthwhile to learn from some of the current research from the field of XML retrieval [11]. The continuous growth of XML documents available in various repositories has attracted a lot of interests in research of focused access using structured documents. The Interactive Track at the Initiative for the Evaluation of XML retrieval is an organized annual meeting for researchers in this research topic [14, 18, 23].

Larsen et al. [15] investigated whether XML element retrieval is useful and helpful for users who are looking for relevant information. Their result suggest that the majority of users selected individual XML element as their entry points. However, XML retrieval might retrieve overlapping elements, i.e. components from the same document [17]. This overlapping issue can be dealt at the interface level as in [13] by presenting a hierarchical grouping of elements from the same document in the hitlist. The users expressed a clear preference for this approach as a means to access structured documents.

Furthermore, Malik et al. [16] studied the users' behaviors while interacting with the XML document. In their study, the users expressed appreciation for the presentation of the document structure as a way to provide context. The table of content seemed to provide sufficient context for users to navigate through the whole document and it gave a quick overview to assess potential relevant elements within that document. In this regard, a study by Szlávik et al. [21] on a feature- and query-biased table of contents is also relevant.

## 3. SYSTEMS AND RESULT DISPLAY

In this section we describe the systems that we used for our user study. These systems are based on MonetDB [2] with the XQuery front-end and the retrieval models of PF/Tijah [9]. The archival finding aids are indexed in a single index, but without filtering stopwords or applying any stemming. The same index is used for all the three systems. We used standard language models with Jelinek-Mercer smoothing to generate retrieval scores, which are used to rank the results in descending order according to relevance. The keywords are treated in this retrieval model as a 'bag of words', and we did not support the use of (Boolean) query operators.

We developed three versions of our Retrieving EADs More Effectively (README) system, which are depicted in Figure 1, where the hitlists of all the three systems show the results for the query "*studentenvereniging Amsterdam*" (in English: student society Amsterdam) that had been entered in the search box. Our system ranks

Retrieving Encoded Archival Descriptions More Effectively (README)

1000 results found in total

- Collectie SIMONE WALVISCH 1970-1978**  
...aankeekingen; stukken betreffende politiek, democratisering van de universiteit van Amsterdam, 1973-1974, de algemene Studentvereniging Amsterdam (ASVA), de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek aan de uva 1971-1972, ...
- Archief F. van der Goes (1854-) 1887-1939 (-1983)**  
... (Brouwer & Grootenans)
- Archief Democratisch-Socialistische Studentenvereniging Politieia. Afdeling Amsterdam. 1948-1956, 1967-1969.**  
...aankomsten; stukken betreffende politiek, democratisering van de universiteit van Amsterdam, 1973-1974, de algemene Studentvereniging Amsterdam (ASVA), de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek aan de uva 1971-1972, ...
- Collectie Igor Cornelissen 1936-1979**  
...militie bij het vijfde volk 1906-1906, later het parool en vanaf 1902, maar bijhoudend bevoorrecht, en actief voor de Studentvereniging politiek, de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek, ...
- Collectie Pieter Wiedijk (1867-1938) 1878-1938**  
...skolen, h. 1885, ...smt., a. en g. 1888, ...stube, j. 1903 (7), ...stovogel, ...
- Archief S.D.A.P. afdeling Groningen (1894) 1895-1920**  
...tot bestudering van het sociaal... 1919. 1. stuk. patrietische aktoeren verkiezingen 111 - 113 stukken betreffende de gemeenteraadsverkiezingen. 1919, 1920, ...
- Archief Ir. Charles G. Cramer Ca. (1879-) 1910-1975**  
...groot aantal organisaties, zoals de vredebeweging de Jarde weg, de socialistische discussiegroep de vâm in haaren, de democratisch-socialistische Studentvereniging politiek, het comité sud-afrika en de europese federatieve beweging... niet alleen socialemae speelde een belangrijke rol in het leven van cramer; mistene e...
- Documentatiecollectie TON REGTEN 1952 - 1989**

(a) Whole Fonds (WF) system

Retrieving Encoded Archival Descriptions More Effectively (README)

1000 results found in total

- Collectie SIMONE WALVISCH 1970-1978**  
...aankeekingen; stukken betreffende politiek, democratisering van de universiteit van Amsterdam, 1973-1974, de algemene Studentvereniging Amsterdam (ASVA), de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek aan de uva 1971-1972, ...
- Archief F. van der Goes (1854-) 1887-1939 (-1983)**  
... (Brouwer & Grootenans)
- Archief Democratisch-Socialistische Studentenvereniging Politieia. Afdeling Amsterdam. 1948-1956, 1967-1969.**  
...aankomsten; stukken betreffende politiek, democratisering van de universiteit van Amsterdam, 1973-1974, de algemene Studentvereniging Amsterdam (ASVA), de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek aan de uva 1971-1972, ...
- Collectie Igor Cornelissen 1936-1979**  
...militie bij het vijfde volk 1906-1906, later het parool en vanaf 1902, maar bijhoudend bevoorrecht, en actief voor de Studentvereniging politiek, de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek, ...
- Collectie Pieter Wiedijk (1867-1938) 1878-1938**  
...skolen, h. 1885, ...smt., a. en g. 1888, ...stube, j. 1903 (7), ...stovogel, ...
- Archief S.D.A.P. afdeling Groningen (1894) 1895-1920**  
...tot bestudering van het sociaal... 1919. 1. stuk. patrietische aktoeren verkiezingen 111 - 113 stukken betreffende de gemeenteraadsverkiezingen. 1919, 1920, ...
- Archief Ir. Charles G. Cramer Ca. (1879-) 1910-1975**  
...groot aantal organisaties, zoals de vredebeweging de Jarde weg, de socialistische discussiegroep de vâm in haaren, de democratisch-socialistische Studentvereniging politiek, het comité sud-afrika en de europese federatieve beweging... niet alleen socialemae speelde een belangrijke rol in het leven van cramer; mistene e...
- Documentatiecollectie TON REGTEN 1952 - 1989**

(b) Individual Archival Material (IAM) system

Retrieving Encoded Archival Descriptions More Effectively (README)

1000 results found in total

- Collectie SIMONE WALVISCH 1970-1978**  
...aankeekingen; stukken betreffende politiek, democratisering van de universiteit van Amsterdam, 1973-1974, de algemene Studentvereniging Amsterdam (ASVA), de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek aan de uva 1971-1972, ...
- Archief F. van der Goes (1854-) 1887-1939 (-1983)**  
... (Brouwer & Grootenans)
- Archief Democratisch-Socialistische Studentenvereniging Politieia. Afdeling Amsterdam. 1948-1956, 1967-1969.**  
...aankomsten; stukken betreffende politiek, democratisering van de universiteit van Amsterdam, 1973-1974, de algemene Studentvereniging Amsterdam (ASVA), de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek aan de uva 1971-1972, ...
- Collectie Igor Cornelissen 1936-1979**  
...militie bij het vijfde volk 1906-1906, later het parool en vanaf 1902, maar bijhoudend bevoorrecht, en actief voor de Studentvereniging politiek, de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek, ...
- Collectie Pieter Wiedijk (1867-1938) 1878-1938**  
...skolen, h. 1885, ...smt., a. en g. 1888, ...stube, j. 1903 (7), ...stovogel, ...
- Archief S.D.A.P. afdeling Groningen (1894) 1895-1920**  
...tot bestudering van het sociaal... 1919. 1. stuk. patrietische aktoeren verkiezingen 111 - 113 stukken betreffende de gemeenteraadsverkiezingen. 1919, 1920, ...
- Archief Ir. Charles G. Cramer Ca. (1879-) 1910-1975**  
...groot aantal organisaties, zoals de vredebeweging de Jarde weg, de socialistische discussiegroep de vâm in haaren, de democratisch-socialistische Studentvereniging politiek, het comité sud-afrika en de europese federatieve beweging... niet alleen socialemae speelde een belangrijke rol in het leven van cramer; mistene e...
- Documentatiecollectie TON REGTEN 1952 - 1989**

(c) Archival Material in Context (AMC) system

Retrieving Encoded Archival Descriptions More Effectively (README)

1000 results found in total

- Collectie SIMONE WALVISCH 1970-1978**  
...aankeekingen; stukken betreffende politiek, democratisering van de universiteit van Amsterdam, 1973-1974, de algemene Studentvereniging Amsterdam (ASVA), de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek aan de uva 1971-1972, ...
- Archief F. van der Goes (1854-) 1887-1939 (-1983)**  
... (Brouwer & Grootenans)
- Archief Democratisch-Socialistische Studentenvereniging Politieia. Afdeling Amsterdam. 1948-1956, 1967-1969.**  
...aankomsten; stukken betreffende politiek, democratisering van de universiteit van Amsterdam, 1973-1974, de algemene Studentvereniging Amsterdam (ASVA), de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek aan de uva 1971-1972, ...
- Collectie Igor Cornelissen 1936-1979**  
...militie bij het vijfde volk 1906-1906, later het parool en vanaf 1902, maar bijhoudend bevoorrecht, en actief voor de Studentvereniging politiek, de herstructurering van de letterenfakulteit van de uva, avsa-semester 1973, 1978, het bestuurconflict bij het Instituut voor Neerlandistiek, ...
- Collectie Pieter Wiedijk (1867-1938) 1878-1938**  
...skolen, h. 1885, ...smt., a. en g. 1888, ...stube, j. 1903 (7), ...stovogel, ...
- Archief S.D.A.P. afdeling Groningen (1894) 1895-1920**  
...tot bestudering van het sociaal... 1919. 1. stuk. patrietische aktoeren verkiezingen 111 - 113 stukken betreffende de gemeenteraadsverkiezingen. 1919, 1920, ...
- Archief Ir. Charles G. Cramer Ca. (1879-) 1910-1975**  
...groot aantal organisaties, zoals de vredebeweging de Jarde weg, de socialistische discussiegroep de vâm in haaren, de democratisch-socialistische Studentvereniging politiek, het comité sud-afrika en de europese federatieve beweging... niet alleen socialemae speelde een belangrijke rol in het leven van cramer; mistene e...
- Documentatiecollectie TON REGTEN 1952 - 1989**

(d) Result display of a fonds

Figure 1: Retrieving EADs More Effectively (README) systems

and retrieves the EAD XML elements separately as natural units by computing different weights for each unit. We directly post-process the results to generate the presentation in HTML/CSS.

**Whole Fonds (WF) system** The *Whole Fonds* system as shown in Figure 1(a) ranks and retrieves a full document, and is comparable to a conventional document retrieval system. For each result, a title and a snippet are presented.

**Individual Archival Material (IAM) system** Figure 1(b) shows the *Individual Archival Material* system, that retrieves XML element nodes as natural units, and it is therefore comparable to a standard XML element retrieval system. Besides the title and the snippet of the element, we also show its result path in XPath.

**Archival Material in Context (AMC) system** We add context to the *Individual Archival Material* system in Figure 1(c) with the *Archival Material in Context* system as described in [26], where we group results from the same archival description and present query-sensitive results ordered by document structure. Contextual structure is added in the hitlist by presenting the titles, headings, and other structural dependencies belonging to the relevant units.

The systems have a different *point of entry* to the finding aids that describes the archival material, which is where a user first enters a document when clicking on a result. For the *Whole fonds* system it is always the beginning of a document, which is the top of the file. For the *Individual Archival Material* and *Archival Material in Context* systems it depends on the XML element that was retrieved, which could be any level of the finding aid in the form of a result path. When a user clicks on a result in the latter two systems, the system will jump to this point using HTML anchors that were automatically generated in XSLT.

Figure 1(d) depicts the result display of an XML file that contains archival material (fonds). This display is generated dynamically on the fly with XSLT and fully presented in CSS, with on the left side the table of contents with the EAD headings `<heading>` and unit titles `<unittitle>`, and on the right side the presentation of the actual content of the fonds. The navigation within the finding aid is supported by clicking on an item in the table of contents or using the scrollbars. Since we deal with archival material in the form of inventories with different types of granularity, we also exactly preserve the original hierarchical structure in our presentation. This means that we keep the provenance that was compiled by the archival creator, and respect the fonds. The original keywords are highlighted in the fonds.

## 4. EXPERIMENTAL SETUP

In this section, we describe the setup of our interactive experiment. The three systems, as introduced in Section 3 above, provide proof-of-concept implementation of three fundamentally different ways of using digital finding aids. But what is the value of each of the systems for archival access? Thus, we conducted a user study to assess the merits of these three systems. In our experiment we address three main questions:

- How satisfied are test persons with the systems?
- How do test persons compare the three systems?
- What are the general opinions on the usefulness of taking structure and context into account?

We address the first question by analyzing data collected via post-task questionnaires where we asked the test persons how satisfied they were with each system. The second question by analyzing

data collected via a post-experiment questionnaire. The third question by gathering data on the general opinion as part of the post-experiment questionnaire.

**Subjects** The intended user population consists of both current visitors of archival institutions, and potential users of archival material within the public at large. Hence subjects with varying degrees of knowledge of archival practices were recruited, and there precise background was carefully registered in the pre-experiment questionnaire. Subjects were students from the university community and archivists or historians recruited via advertisements in mailing lists of archivists. Subjects' participation was voluntary.

**Document Collection** We used a data-set of 2,866 EAD documents from the International Institute of Social History [10] (IISG), ranging in length from 117 words to 166,000 words. The mean length of an EAD document was 2,179 words. The finding aids were mostly written in Dutch.

**Tasks** Three work task scenarios [3], with 3 different topics were used in the experiment, in addition to a task used for training. The tasks were similar in a way that they were neither biased for any system nor were in favor for a particular system. The tasks were designed to mimic situations where users were interested to find all finding aids that discuss various aspects or instances of a topic. Thus, all tasks were open-ended, required the subjects to navigate through more than a single finding aid in order to be answered, and had different sub topics. An example of a task:

*"As a student in Amsterdam, you are interested in the history of the local student societies. You decide to do research about this topic, with the purpose of writing an article for the student society's magazine. For example, you want to know more about which societies there were, when these were founded, for which students these societies were accessible.*

*Using the digital inventory of the IISG, you want to check out which archives contain interesting pieces for your research. Depending on these findings, you assess whether a visit to the IISG for your research is worthwhile."*

**Measures of Performance** We used subjective performance measure using five-point Likert scale and open-ended questions. Subjects were instructed that there were no correct answers and their assessment of their search result was completely subjective.

**Matrix** The three systems and three tasks were both rotated and counterbalanced in a Greco-Latin design [22]. This results in 9 rotations, the first rotation starting with System WF/Task 1, then System IAM/Task 2, and finally System AMC/Task 3. Over the nine rotations each system and each task appeared exactly three times in each column.

A within-subject laboratory experiment was conducted to compare the 3 systems. This study was conducted in a one-on-one setting. Before the experiment, subjects were introduced to the experiment, followed by a *pre-experiment questionnaire*, which collected demographics questions and the subjects' searching experience. This was followed by a *training session*, where the subjects were presented with a topic and given as long as necessary to get familiarized with each system. Subjects then used these systems to complete three search tasks. Each task was preceded with a *pre-task questionnaire*, collecting information regarding subjects' familiarity, level of interest and easiness of the search topic. Then the *simulated task* itself was performed using the assigned system. The *post-task questionnaire* gathered feedback on the completed task and usefulness of features of the interface. The experiment was closed with a *post-experiment questionnaire*, which asked comparative questions over the three search systems. For each question-

**Table 1: Questions in the pre-task questionnaire**

Q2.1	How familiar are you with the topic of the search task?
Q2.2	How interested are you in the topic of the search task?
Q2.3	How easy do you think it will be to find information for this task?

**Table 2: Responses on pre-task questionnaire: mean scores and standard deviations (in brackets)**

	Q2.1	Q2.2	Q2.3
Task 1	2.00 (1.22)	2.67 (1.00)	3.56 (1.33)
Task 2	2.11 (1.17)	2.67 (1.00)	3.44 (0.88)
Task 3	2.44 (1.24)	3.22 (0.83)	3.67 (0.87)

naire we differentiate subjects' opinion regarding the overall system, hitlist and result display. Subjects were allowed to use up to 15 minutes to complete each task. Subjects' participation in the whole experiment lasted roughly 1.5 hours.

## 5. EXPERIMENTAL RESULTS

### 5.1 Pre-Experiment Questionnaire

Nine subjects participated in this experiment (7=Male, 2=Female). Subjects had a mean age of 37.56 years. As part of the pre-experiment questionnaire, we asked subjects' experience with archives. Six subjects had received archival education or training. Accordingly the same subjects were familiar with archival terminologies in Dutch, but only four of them were familiar with archival terminologies in English. Seven subjects had ever conducted historical research and consulted archival finding aids. Eight subjects had ever visited an archive's website and consulted online finding aids. Since we were using a data set from International Institute of Social History (IISG), we asked subjects if they had ever visited the institution. We found out that five subjects had ever visited the IISG and its website. As for the Web searching experience, subjects' mean experience was 7.56 years.

### 5.2 Pre-Tasks Questionnaire

Each task was preceded with a pre-task questionnaire, collecting information regarding subjects' familiarity, level of interest and easiness of the search topic. Table 1 shows the items asked in the pre-task questionnaire. The answer categories used a 5-point scale (1=not at all and 5=extremely). Subjects' responses are presented in table 2. Although we tried to make the simulated tasks equal in many respects, task 3 was rated slightly higher compared to the other two tasks in terms of subject's familiarity (Q2.1) and level of interest (Q2.2). In terms of subjects' prediction on the easiness to find information for the task (Q2.3), the rates were comparable.

### 5.3 Post-Task Questionnaire

After each completed task, subjects filled in a questionnaire regarding their searching experience. The answer categories used a 5-point scale with 1=not at all and 5=extremely.

**Overall Satisfaction** Table 3 shows the first questions of the post-task questionnaire. Subjects' responses on their searching experiences are presented in Table 4 and 6.

Overall, all systems were rated positively. For some question (Q3.3, Q3.4, Q3.7, Q3.8) subjects rated WF more positive than IAM and AMC. Only in terms of easiness to get started on the search (Q3.1) AMC was rated the highest. As shown in Table 6, we got the same pattern of responses for ease of use (Q3.10) and

**Table 3: Post-task questions on overall system**

Q3.1	Was it easy to get started on this search?
Q3.2	Was it easy to do the search on this task?
Q3.3	Are you satisfied with your search results?
Q3.4	How relevant was the information you found?
Q3.5	Did you have enough time to do an effective search?
Q3.6	How much time did you spend on: a. Querying; b. Browsing; and c. Reading
Q3.7	How certain are you that you completed the task?
Q3.8	How well did the system support you in this task?
Q3.9	How easy was it to learn to use the system?
Q3.10	How easy was it to use the system?

**Table 4: Responses on subjects' searching experience: mean scores and standard deviations (in brackets)**

	Q3.1	Q3.2	Q3.3	Q3.4
WF	4.11 (0.78)	4.11 (0.78)	4.44 (0.73)	4.33 (0.71)
IAM	4.22 (0.83)	4.00 (1.00)	3.44 (1.13)	4.11 (0.78)
AMC	4.33 (0.71)	4.00 (1.32)	3.89 (1.45)	3.89 (1.45)

**Table 5: Responses on the efficiency of the system: mean scores and standard deviations (in brackets)**

	Q3.5	Q3.6.a	Q3.6.b	Q3.6.c
WF	4.56 (0.73)	2.33 (0.87)	3.11 (0.93)	3.56 (1.13)
IAM	4.11 (1.05)	2.44 (0.88)	3.00 (0.71)	3.44 (0.88)
AMC	4.00 (1.22)	2.11 (0.78)	3.56 (1.01)	2.67 (0.87)

**Table 6: Responses on subjects' searching experience: mean scores and standard deviations (in brackets)**

	Q3.7	Q3.8	Q3.9	Q3.10
WF	4.33 (1.00)	3.89 (0.93)	4.67 (0.50)	4.22 (0.67)
IAM	3.00 (1.32)	3.33 (0.50)	4.33 (0.71)	3.89 (1.05)
AMC	3.00 (1.66)	3.56 (0.73)	4.22 (0.67)	3.89 (0.78)

**Table 7: Post-task questions on hitlist**

Q3.13	How satisfied were you with the information provided in the hitlist?
Q3.14	Was the overview of results clear?
Q3.15	Was it easy to select the most promising result?

ease of learning (Q3.9). Although subjects rated all systems positively, WF was rated the highest. A possible explanation was that IAM and AMC were not resembling systems familiar to the test persons, and it took time to learn and to use the systems.

Table 5 shows subjects' responses on the efficiency of the systems. It shows that subjects had enough time to do an effective search for all three systems (Q3.5). Interestingly, subjects indicated that they required more time in querying using IAM system compared to the other two systems (Q3.6.a). Not surprisingly, subjects spent more time in reading using WF compared to the other two systems (Q3.6.c) because the system returned the whole fonds thus required subjects to read more before they found the relevant information. Subjects indicated that among all three systems, they required the least time in reading and the most time in browsing (Q3.6.b) using AMC system. It was clear that the systems affect the information seeking behavior of the test persons.

**Hitlist** Table 7 shows post-task questions toward hitlist features. As shown in Table 8, WF was rated highest in terms of subjects' satisfaction with the information provided in the hitlist (Q3.13) and

**Table 8: Responses on hitlist: mean scores and standard deviations (in brackets)**

	Q3.13	Q3.14	Q3.15
WF	3.78 (0.67)	3.67 (0.87)	3.44 (0.88)
IAM	3.11 (0.78)	2.89 (0.93)	3.11 (1.17)
AMC	3.33 (0.87)	3.22 (0.67)	4.11 (0.78)

**Table 9: Post-task questions on result display**

Q3.17	How satisfied were you with the display of found finding aids?
Q3.18	Was the result display clear?
Q3.19	How useful is the table of contents?
Q3.20	How useful is the highlighting of query words?
Q3.21	How often did you use search (ctrl+f) within the finding aids?
Q3.22	Did you get lost in the structure of the finding aids?

**Table 10: Responses on result display: mean scores and standard deviations (in brackets)**

	Q3.17	Q3.18	Q3.19
WF	3.44 (1.01)	3.78 (0.83)	3.56 (1.24)
IAM	3.56 (1.13)	3.44 (1.13)	3.22 (1.39)
AMC	3.67 (0.50)	3.89 (0.60)	3.78 (0.97)

**Table 11: Responses on result display: mean scores and standard deviations (in brackets)**

	Q3.20	Q3.21	Q3.22
WF	4.00 (1.12)	2.11 (1.45)	2.11 (1.05)
IAM	4.11 (1.36)	1.33 (0.71)	2.11 (1.17)
AMC	4.44 (0.73)	1.33 (0.71)	1.78 (0.97)

clearness of results overview (Q3.14). We also found that AMC was rated the highest in terms of supporting subjects to select the most promising result in the hitlist (Q3.15). This can be explained by the possibility in AMC to choose many relevant results per fonds and to present the context of where the results were found.

**Result Display** Table 9 shows post-task questions toward result display features. As shown in Table 10 and 11, result display of AMC was rated the most positive for all items we asked on the questionnaire (Q3.17–Q3.21). Since the same result display was used for all three systems, the only difference was the entry point (either at the start of the whole document, or at a specific element). Interestingly, the difference of entry point gave major difference to the test persons. It was found that when subjects used AMC, they were more satisfied with the display of found finding aids (Q3.17) and were less frequently lost in the structure of the finding aids (Q3.22) compared to the other two systems. We believe because the structure of the finding aids were already shown in the hitlist, thus subjects can predict where they would be in the finding aids before they were presented with the result display. Another interesting point was that subjects used less Ctrl+F within the finding aids (Q3.21) using IAM and AMC systems. This was plausibly related with the possibility to go directly to the relevant archival material without having to search within the fonds. This is also consistent with our finding in terms of efficiency in question Q3.6.

## 5.4 Post-Experiment Questionnaire

After all tasks were completed, subjects filled in a post-experiment questionnaire on the comparative evaluation of the three systems. We asked subjects to compare the three evaluated system in terms

of the overall impression of the systems, hitlist and result display. The answers categories used a 5-point scale (1=not at all and 5=extremely). As part of the post-task questionnaire, we asked subjects to choose which system, hitlist, and result display they like the most and why. Table 12 presents the main results.

**Overall** In terms of subjects' satisfaction on the overall system (Q4.3), AMC system was rated highest. When asked which system they liked the best (Q4.6), 7 subjects chose AMC, 1 subject chose both WF and AMC and 1 subject chose WF. Subjects who chose for AMC argued that the system provided them with the total overview of the finding aid, gave the most opportunity to assess the relevance of the information and took them directly to the relevant part of finding aids. Subject who chose for WF argued that it was easier to see the relevancy using WF. We attribute it to WF being clearer to the user due to its simplicity. Since the system only showed title and preview of the finding aids, it was easier for the users to use and to interpret. Another subject chose for WF because it was better for searching general information.

**Hitlist** For all the questions about subjects' satisfaction toward hitlist, subjects rated AMC the highest. The overall impression of the hitlist (Q4.7) was the highest for AMC. When asked which system they liked the best (Q4.11), 7 subjects chose AMC, and 2 subject chose WF. The reasons for subjects preferences toward hitlist were similar as the reasons toward the overall system.

**Result display** For all the questions about subjects' satisfaction towards result display, subjects rated AMC the highest, although less convincingly than for the hitlist. This may come as no surprise since the result display was identical for all systems, except for the entry point. The overall impression of the result display (Q4.12) was rated highest for the AMC system. When asked which system they liked the best (Q4.15), 4 subjects chose AMC, 2 subjects chose WF, 2 subjects chose both IAM and AMC, and 1 subject chose WF and AMC. The result displays of IAM and AMC were similar. The reasons for choosing IAM and AMC was because they provided direct access to the relevant information. Subjects chose WF because the system would be useful for searching for general information, where subjects did not need to be lead to deep structure of the finding aids. Furthermore, another user chose WF because the entry point helped in finding the first query match in the finding aid.

## 5.5 General Views

Four open-ended questions were asked in the post-experiment questionnaire to gain more qualitative insight in the feedback of the users. For each of those open questions, positive qualitative feedback was provided that valued aspects of presenting archival material in context. We elaborate here on these results.

The first open question was about the usefulness of the structure. The subjects highlighted that the structure provided context to the information and was very important for their research. They also mentioned that the structure gave possibility to navigate through the finding aids. The main criticism was that the meaning of the structure was not very clear to the users. It was because the structure was presented using EAD tags and this was not intuitive for the test persons, who were not intimately familiar with EAD.

The second open question was about direct access to relevant parts of an archival description. Almost all subjects appreciated the possibility to go directly to the relevant part of finding aids, mainly because it decreased their search effort. However, subjects mentioned and were observed that even when they directly accessed the relevant part of finding aids in the result display, they still wanted to know the global view of the finding aids. Hence, most users

**Table 12: Comparative evaluation: overall impression (Q4.3, Q4.7, Q4.12) and system preference (Q4.6, Q4.11, Q4.15).**

	System overall		Hitlist		Result display	
	Q4.3	Q4.6	Q4.7	Q4.11	Q4.12	Q4.15
WF	3.56 (0.88)	1.5	3.56 (0.88)	2	3.44 (0.88)	2.5
IAM	3.22 (1.48)	0	2.89 (1.54)	0	3.56 (0.53)	1
AMC	4.11 (0.93)	7.5	4.11 (0.93)	7	3.89 (0.93)	5.5

browsed to the beginning of the finding aids to know the context (title, biography, etc) of the fonds.

The third question elicited feedback about the usefulness of grouping results within the context – and again our users expressed clear appreciation for this feature. There were two main reasons for this. Firstly, it gave information about the origin of the fonds, which was essential for the users in their research. Secondly, it helped the users in their relevance assessment.

The last open question was about the future use of the system. Most of them expressed that they would use the type of search engine as AMC, especially for historical research. This finding also suggests that future work on the AMC variant of the README system is worthwhile. Complete responses from the users are presented in the Appendix 6.

## 6. DISCUSSION AND CONCLUSIONS

In this paper, we investigated how to provide access to archival material using digital archival finding aids. Archival finding aids are long and complexly structured documents describing archival material—the paper trails of the lives of corporate bodies, persons, and families—that are currently encoded in XML using the Encoded Archival Descriptions (EAD) standard. But how to provide access to such long and complexly structured documents? Using some of the insights from the field of XML retrieval, we developed three different systems for searching collections of digital finding aids corresponding to three fundamental choices about archival access. The first system provides access to the fonds or archive as a whole; the second system provides direct access to individual archival material at any level of description; and the third system retrieves archival material while preserving the original context. We conducted an extensive user study with the three systems to investigate the research questions that we have raised in this paper.

Firstly, is direct access to archival material in digital finding aids deemed useful in comparison to access to the whole fonds? The test persons were very positive toward the possibility of directly access archival material because it decreased their search effort. The preference was clear in the comparative evaluation where subjects rated the systems that provide direct access higher than the system that returns the whole fonds. Our open-ended questions also supported our quantitative results where most subjects appreciated the usefulness of direct access to relevant parts of long archival description. The lesson we can draw from this is that direct access in digital finding aids is promising. Almost unanimously our test persons responded that they would like to use the system for finding general information and specific archival material.

Secondly, can the use of structural presentation improve users’ satisfaction in locating relevant information, or help them in deciding if a collection is relevant? Our result showed that the structural presentation supports the test persons to locate relevant information and to select the most promising results in the hitlist. The test persons appreciated the possibility to see multiple results grouped per fonds and to see the context of where the results were found. This is in line with findings by Dumais et al. [5] where they compared list and category interfaces for search results interfaces. Their study showed that category interfaces, where items in the same

category were grouped together, allowed users to examine search results more efficient compared to list interfaces. Our result also showed that test persons were not lost in the structure of the finding aids. However, we found that there was a trade off between the presentation of structural overview and simplicity of the interface. We preserved document structure as a tree-like structure and presented the EAD structure in the hitlist, and test person complained about the unfamiliar codes of the EAD structure.

Thirdly, can individual items be used as an entry points for accessing an archival description? If so, does it affect the users’ satisfaction in accessing relevant information? Our experiment showed that individual items can be used as entry points for accessing an archival description. Users mentioned that the entry points help them in their navigational effort because the entry points reduced their effort in reading and scrolling. Test persons were more satisfied with the systems that allowed them to go directly to the relevant part of finding aids compare to the system that takes them to the beginning of the finding aids.

This paper focused on how archival material in context can be used effectively to access archival material. We evaluated three variant systems in a user study. The user study had an exploratory nature, and with nine test persons and 27 search tasks completed we have to be careful about drawing general conclusions. Still, our findings suggest that direct access to the archive was received favorably and that it helps test persons in navigation. Furthermore, presenting structural information in the hitlist provides valuable context of archival material. Test persons used this context both to know where they are in the hierarchy of the finding aids, and to support them in assessing relevancy of the results. Grouping results from the same archival description seems to be an effective way to deal with the structural dependencies between retrieved results. Finally, the study also gave concrete suggestions on improving the user interface by presenting the structural dependencies in a more intuitive way, which we will explore in future research.

**Acknowledgments** We thank the test persons for donating their time and the IISG for the use of archival descriptions in EAD. This research is supported by the Netherlands Organization for Scientific Research (NWO) under grant # 639.072.601.

## REFERENCES

- [1] L. J. Bellardo and L. L. Bellardo. *A Glossary for Archivists, Manuscript Curators, and Records Managers*. The Society of American Archivists, Chicago, 1992.
- [2] P. A. Boncz, T. Grust, M. van Keulen, S. Manegold, J. Rittinger, and J. Teubner. MonetDB/XQuery: A fast XQuery processor powered by a relational engine. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Chicago, IL, USA, 2006.
- [3] P. Borlund and P. Ingwersen. The development of a method for evaluating interactive information retrieval systems. *Journal of Documentation*, 53:225–250, 2000.
- [4] L. R. Coats. Users of EAD finding aids: Who are they and are they satisfied? *Journal of Archival Organization*, 2:25–39, 2004.
- [5] S. Dumais, E. Cutrell, and H. Chen. Optimizing search by showing results in context. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 277–284, Seattle, WA, April 2001.

- [6] L. Duranti. Origin and development of the concept of archival description. *Archivaria*, 35:47–54, 1993.
- [7] EAD. Encoded archival description version 2002, 2002. <http://www.loc.gov/ead/>.
- [8] M. Hedstrom. Descriptive practices for electronic records: Deciding what is essential and imagining what is possible. *Archivaria*, 36:53–63, 1993.
- [9] D. Hiemstra, H. Rode, R. van Os, and J. Flokstra. PF/Tijah: text search in an XML database system. In *Proceedings of the 2nd International Workshop on Open Source Information Retrieval (OSIR)*, pages 12–17, 2006.
- [10] IISG. International institute of social history, 2008. URL <http://www.iisg.nl/>.
- [11] INEX. INitiative for the Evaluation of XML retrieval, 2008. <http://www.inex.otago.ac.nz/>.
- [12] ISAD(G). *General International Standard Archival Description*. International Council on Archives, Ottawa, second edition, 1999.
- [13] J. Kamps and B. Sigurbjörnsson. What do users think of an XML element retrieval system? In *Advances in XML Information Retrieval and Evaluation (INEX 2005)*, volume 3977 of *LNCS*, pages 411–421. Springer, 2006.
- [14] B. Larsen, S. Malik, and T. Tombros. The interactive track at INEX 2005. In *Advances in XML Information Retrieval and Evaluation (INEX 2005)*, volume 3977. Springer Verlag, Heidelberg, 2006.
- [15] B. Larsen, A. Tombros, and S. Malik. Is XML retrieval meaningful to users? searcher preferences for full documents vs. elements. In *Proceedings of the 29th ACM SIGIR Conference*, pages 663–664, 2006.
- [16] S. Malik, C.-P. Klas, N. Fuhr, B. Larsen, and A. Tombros. Designing a user interface for interactive retrieval of structured documents: lessons learned from the INEX interactive track. In *10th European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*, pages 291–302, 2006.
- [17] S. Malik, B. Larsen, and A. Tombros. Report on the INEX 2005 interactive track. *SIGIR Forum*, 41:67–74, 2007.
- [18] S. Malik, A. Tombros, and B. Larsen. The interactive track at INEX 2006. In *Comparative Evaluation of XML Information Retrieval Systems (INEX 2006)*, volume 4518 of *LNCS*, pages 387–399. Springer Verlag, Heidelberg, 2007.
- [19] A. Rosenbusch. Are our users being served? a report on online archival databases. *Archives and Manuscripts*, 29:44–61, 2001.
- [20] B. Stockting. Time to settle down? EAD encoding principles in the access to archives programme (A2A) and the research library group’s best practice guidelines. *Journal of Archival Organization*, 3:7–23, 2004.
- [21] Z. Szlávik, A. Tombros, and M. Lalmas. Feature- and query-based table of contents generation for XML documents. In *29th European Conference on Information Retrieval (ECIR)*, pages 456–467, 2007.
- [22] J. Tague-Sutcliffe. The pragmatics of information retrieval experimentation, revisited. *Information Processing and Management*, 28: 467–490, 1992.
- [23] A. Tombros, B. Larsen, and S. Malik. The interactive track at INEX 2004. In *Advances in XML Information Retrieval*, volume 3493 of *LNCS*, pages 410–423. Springer Verlag, Heidelberg, 2005.
- [24] XML. Extensible markup language (xml) 1.0 (fourth edition), 2006. <http://www.w3.org/TR/xml/>.
- [25] E. Yakel. Encoded archival description: Are finding aids boundary spanners or barriers for users? *Journal of Archival Organization*, 2: 63–77, 2004.
- [26] J. Zhang, K. N. Fachry, and J. Kamps. Access to Archival Finding Aids: Context Matters. In *12th European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*, volume 5173 of *LNCS*. Springer, 2008.

## APPENDIX

### *Responses on the usefulness of direct access.*

Q4.17 Do you find it useful to be pointed to relevant parts of long archival description? Why?

Yes, it support quick glance.

Yes, I can navigate more easily.

Yes, it is easy to use

Yes, because I do not need to browse/search much.

Yes, The context of the information is important for research.

Yes, to prevent too much scrolling.

Sometimes where there is a whole section that is really relevant, most of the time I like to start reading at the beginning of the document and the highlighted terms are useful for finding the relevant parts.

Yes, but not always. In some topics, I know the information would be at the beginning of the finding aid, so I want to go to the top immediately.

Yes, it prevents long browsing.

### *Responses on the usefulness of structure.*

Q4.16 Did you like the idea that the search engine takes into account the structure of the archival descriptions? Why?

Yes, it can enhance navigational mobility. But the tree structure needs to be improved because it is not always clear from the hitlist.

Yes, that way you can decide what is relevant or not better and faster.

I like it that it brings me directly to the relevant parts of document, but I did not like the “code looking” part in the hitlist. The code takes a lot space and I have to scroll a lot.

Yes, the structure is useful.

I think it is a necessary instrument. The context of the information is important for research.

Yes, because archives contain hierarchical structured information.

I do not really care about the structure.

Yes, because it makes it easier to assess where in the fonds I can find information. However, the structure is based on EAD and not on the logical structure of the fonds itself.

Yes, the structure is helpful. But the way it is now (using the code) is difficult to understand.

### *Responses on the grouping of results within context.*

Q4.18 Do you find it useful that results are grouped within the context of the archival fonds? Why?

Yes, it is very intuitive. Because I want to be able to judge an archive at a glance, not judge it over and over again as in IAM.

Yes, I can navigate more easily.

Yes, useful. When opening a result, I browse a little through the document, so I see all the highlighted parts. When I go back to the hitlist, I look at the next result, I see that I already had a look at that document, and then I can conclude that I do not have to open it again.

Yes, to assess the relevance of the information.

Yes, it is important to group the results per archival fonds to know what the value of the information is.

Yes, it helps you to know the structure and to decide whether an archive is relevant or not.

Yes, it helps a lot to decide whether an archive is relevant or not.

Yes, but it should be grouped based on the logical structure of the finding aids (instead of EAD).

Yes, it gives an impression how often the query appears in the fonds. That says a lot about its relevance.

### *Responses on future use of the system.*

Q4.19 Would you use this type of search engine if it was available? If so, for what types of tasks?

Yes, for most tasks, any search that would involve sifting through the available documents.

Yes, for finding general information archival bodies or for finding archival material.

Yes

Yes

Yes, for historical research, making archive inventories, etc.

Yes, with the ratio of the results ordering.

No.

Yes, for many purposes. 1. To discover whether there are relevant fonds about the topic. 2. To discover whether there are relevant series/files in a fonds. 3. On the bases of this information, I can more easily start a search strategy that is not based on keywords, but on the basis of the logical structure of the fonds.

Yes, for searching archives to get information about a subject.