Information retrieval by users in archival information systems : search terms selection patterns / Masha Zolotarevsky-Bear

Abstract

Information retrieval in archives was traditionally based on the principal of provenance, in which the search is based on what is known about the information unit. Using the provenance principle means that users seeking information must "get into the head" of the information creators – they must consider who could have created the information they seek and under what circumstances. This method of searching is appropriate for information retrieval by archivists or researchers familiar with archival information structure, but is not easy for users who have little knowledge about the way information is organized in archives. Currently, many archives are producing web-based finding aids. Constructing these finding aids, which would meet the needs of inexperienced users as well as knowledgeable researchers and allow independent search outside the repository, has become one of the main concerns of archivists.

The current study deals with the less investigated aspects of information retrieval. The issues addressed are search subject definition, the formulation of search queries, and a comparison between the ways users and archivists search for archival information.

The study is based on theoretical literature from the fundamental works describing the development of the archival profession to the current state, and on research which deals specifically with archival users and search behavior. Since the early 1970s, the professional literature has been discussing the urgent need for a systematic and broad study of archive users and their needs. A recent surge of users' studies has enhanced the archival profession, adding the reformulation of definition of archive users. This reformulation is based on the understanding of the great variety of purposes for which people use archival materials, the numerous uses for archival finding aids, and various search techniques that could potentially help archivists to better employ their search tools. Nevertheless, as this field is still relatively new in

---

archival science, several significant lacunae exist, which this study attempts to fill, mostly regarding the search behavior of various types of archive users.

The study focuses on two main questions: (1) How do users define their search subjects and search terms? (2) Are there differences in the ways archivists and external users search for archival information? These questions can be elaborated in the following sub-questions:

1. How do different types of users define search subjects?
2. Which search terms do users choose when searching in electronic information systems?
3. In what way are these search terms linked to the users' search subject and to their general purpose for visiting the archive?
4. Are there any differences in the way users translate their subject into search terms between different archives?
5. Are there any differences between users and archivists in the choice of search terms?

The questions address an existing issue and the research will add to knowledge in an area of archive users search patterns. The study departs from the examination in previous research, where the interest was in historians only. The present study broadens the scope, and examines a variety of archive users. Furthermore, the analysis presented here refers to both subject definition and the search queries as they are constructed by users in real time, the issue that has not yet received appropriate research attention. Finally, based on the research studies, I constructed the information model which shows the differences between direct users and indirect users – the information system interaction.

In order to answer the research questions and the sub-questions, I chose the qualitative-constructivist approach, primarily because of way the constructivists refer to the researched phenomenon –understanding people's behavior in their natural environment. This approach meets the research purpose, which was to learn how the different types of users behave in the archival reading room during the entire information retrieval process, beginning with subject definition, building search query or queries, and deciding to end the search.

In order to gather overall information different data collection methods were used. Statistical reports provided necessary data to understand the scope of activity in archives, and search logs filled in by users and archivists for information about subject definitions, search purposes, and search queries. In addition to the written material, I conducted in-depth interviews with archivists, thus obtaining their point of view on the search process in archives and helping to complete the understanding of users search patterns. Each of these methods contributed it part to the research and also helped fill in some of the lacunae that other methods left.

The research was conducted in four Israeli archives: the Yad Tabenkin Archives, Moshe Sharett Israel Labor Party Archives, the Oded Yarkoni History Archives of Petach Tikva, and the Archives of the Jabotinsky Institute. The conditions for selecting specific archives for this study were first, choosing archives with different profiles and second, choosing archives that had an electronic information system (preferable different systems in different archives in order to avoid possible bios) and workstations in the reading room, allowing users to perform searches on their own.

The findings of the research met the research questions and allowed to reach several conclusions regarding the subject of this research. First the research shows that different kinds of users have specific search patterns which can be depend on the search purposes or on the archives itself. These search patterns were expressed in the definition of the search subjects, choice of search terms, constructing the query, and the satisfaction with the search results. The findings show that archive users can construct simple and complex queries, and often start their research in archives already equipped with alternative search terms, names, places, and common terms from related literature. Notwithstanding, archive users do appreciate the interaction with the archivists and do not hesitate to ask archivist's help and guidance regarding using archival systems and retrieval information.

Second, the research findings show that there are indeed slight differences between the ways users and archivist search the information. Archivists were more familiar with advanced search techniques and with archival information structure and principles. They were also more familiar with specific holdings and fonds, and often could easily retrieve the information because they "just know." However, very often the users were more familiar with the research subject and therefore had more search

terms. Thus, the interaction between the users and archivists, in order to combine their areas of knowledge and expertise, can be very efficient in the information-retrieval process, and archivists should make an effort to allow this interaction for users in reading rooms as well as for virtual users who search for archival information through archives websites and portals.

Third, the research shows that systematic examination of the search subjects defined by users can illustrate the ways users search through archival information systems. This knowledge could help archivists to develop appropriate search tools as well as to improve their finding aids.

Finally, the information model of interaction between users and the information system was constructed. Based on the research findings, the model reflects the influence of the users' knowledge of the research subject, the archival structure, and the search techniques on the search query the users construct. In addition the different schemes show the differences between direct user-system interaction and the interaction through archivist's assistance, benefitting both parties to the archive search.

System No.
002436231