Networking historical sources: A demand-side driven approach

Andrea Rosenbusch (*)

(*) Swiss Federal Archives, Bern, Switzerland
E-mail:andrea.rosenbusch@bar.admin.ch

ABSTRACT
Many archives have introduced online archival databases in the past few years. However, apart from being independent of place and time, hardly any of these databases make finding records any easier than it was before. This paper suggests to put more emphasis on the needs and behavior of users and to provide them with the context in which records were created, thus making it possible to link archives in distributed locations and cultural heritage in other organizations as well as enabling users to learn archival search strategies and make better use of the databases provided.

KEYWORDS: User focus, accessibility, network

INTRODUCTION

The Swiss Federal Archives is undertaking a project to create new spaces in the World Wide Web in order to make cultural heritage more accessible. Today, the structure of the Web largely reflects the organizations or individuals supplying information resources, not the contents they provide. Users of the Web looking for heritage related information and material (in our case, the focus is on modern history – political, social, economic) are not so much interested in the structure of the providing institutions as in certain topics. So far, there are very little subject-based entry points to the Web with the exception of search engines, and the possibilities of linking resources have hardly been fully exploited. The main aims of our project therefore are:

- To create entry points to our collection which are based on users' interests and abilities and which take different user groups into account;
- to provide the context of the sources: On the one hand, this makes it possible for users to ascertain the value of the information received; on the other, resources can be linked via their mutual context across institutional or other barriers.
- to initiate a database supplied by several partners independent of the type of resources provided on the basis of a common metadata set.
- to play an active role in building specialized networks which focus on subjects and thus create new entry points to several collections.

THE WEB AS AN INFORMATION RESOURCE

The complexity of the Web is an ideal prerequisite for the arrangement of information, and hyperlinks are ideal for exploring and learning. The complex structure seems to be ideal to represent archival search paths, which depend on vast knowledge of the context of records and connections between the elements in this context. Hyperlinks also to some

extent offset the necessity of creating and adhering to common standards which are so important when sharing databases. However, the Web demands information be largely self-explanatory to a growing and increasingly diverse number of people. The challenge, in our opinion, therefore lies less in the use of technology itself than in the arrangement of knowledge to make the maximum use of technology.

A FOCUS ON USERS

If we want our users to be able to search our resources independently on the Web, we need find out more about their needs and behavior and tailor our applications accordingly. Up to now, very little work has been done in the archival field to find out more about our users' needs. An exception is the recent study by Wendy Duff and Catherine A. Johnson. (1) Even less work seems to have been conducted on making Web applications user friendly. (2) The first step in our project thus is an empirical study of our users with the aim of taking their perspective more into account.

The Federal Archives holds written inquiries dating back over half a century as well as e-mail requests from the last few years. The latter seem to reflect the changing public addressing its questions to the archives as well as a change in strategies for finding information in the age of the Internet. In order to be able to perceive trends, we will analyze samples of such inquiries over several years. We aim to find anwers to questions on the following topics:

- The background of the inquirers (e.g. are they trained in or new to historical research),
- the search paths so far pursued (e.g. is this the first place they are looking for information),
- the way their attention has been directed towards this archives (e.g. have they found it via the Internet or been referred to it by someone else),
- the range or specification of knowledge they have on the topic they are doing research on (– this has a strong tie to the search paths, as a well-informed person with experience in research has many more possible access points to information),
- the kind of answer they have in mind (e.g. are they looking for a specific document, a piece of background information or a range of records).
- The degree to which the questions asked can be answered with the current finding aids.

Based on the findings, we will need to decide on how best to guide our users to the sources they are looking for: Is it by categorizing user groups according to experience or by arranging sources according to topics of interest? Can we assist users by making them specify exactly what the are looking for and prompting them to give details which might lead to that information? Does the selection of the approach(es) influence which user groups are best served?

As the last dot point above implies, we also expect the results of our survey to possibly have an influence on the finding aids we provide. It might be necessary to digitize or even create additional finding aids. In any case, we will have to enable users to make the most of the complexity of our finding aids without getting lost and without having to read too many explanations. All of these questions and tasks show us we are embarking on an ambitious and

lengthy project. In the meanwhile, however, we will try to simplify matters somewhat and carry out some projects which we believe from experience will enhance the usability of our archives.

CONTEXT

Most people who first use an archives do not know what it holds. The accessibility of information on the Internet adds to this problem. Enter a search term in an online archival database and see what you can make from the results! Users need to be able to judge the relevance and reliability of the information they are being given. Whilst the reliability must be assured by the provider of the information, i.e. the archival institution, the relevance of hits can only be estimated with much contextual knowledge, because archival finding aids are often very selective. This can lead to misleading results, showing many hits in irrelevant areas or no records at all because no detailed lists are available online. Therefore, users may have to pursue a variety of search strategies to find what they are looking for. For example, somebody is looking for biographical information on a Swiss school teacher who emigrated to South America in the 19th century. The Swiss finding aids will probably not list the name and maybe not even feature anything on schools, because these records were never shipped back to Switzerland. However, the diplomatic mission is bound to have some data like the date of arrival in South America, marriage, and death. The teacher might have been active in the local Swiss club which sent its records back home, and there might be some records relevant to the subject. To find documents, it is essential to know what kind of records might have been produced in the first place. Descriptions of record creators and their functions therefore are key access points for researchers, and even more so because they are not necessarily pertinent to an archival institution. Records which reflect the activities of a record creator or a function carried out by multiple record creators need not be held by the same institution, and custody might even change over time, but the context of the creation of the records will always stay the same. (3) Knowledge on record creators and their functions is also essential for the interpretation of records. No record creator is impartial, and the more records from different sources are scrutinized, the more differentiated research becomes. Finding distributed resources is the concern of many researchers. (4) To enhance the possibilities of the description of the context of records and to integrate descriptions of records themselves, we aim to participate in a shared database with other institutions of cultural heritage.

A SHARED DATABASE

Context is an essential element when a search is conducted by a machine on the Internet. Only if sufficient context is given can the user judge the usefulness and trustworthiness of information. The larger the range of resources used for the search, the bigger the problem. However, if only a sample of resources is searched through, is the information received complete or at least adequate? The compromise we suggest is that institutions, e.g. from the archives, museums and library sectors, create common databases to help users identify those resources. There is a difference between portals of institutions and common databases. A portal like that of UNESCO (5) is an improvement for researchers, but it still requires queries to be conducted in each resource separately. The common

database should be a possibility to overcome barriers between different kinds of providers, e.g. archives and libraries. We expect our users to mainly look for subject-based information, independent of it being a record, an object or a published resource. Let's assume our user is not looking for teacher X. in the shared database, but for the emigration from Switzerland to South America. With participants in the network from Switzerland alone, the user should be able to find records from a number of State and some private archives, objects in museum collections as well as some publications.

Creating a common database however poses some problems. Semantic and syntactical standards for the exchange of machine-readable archival data and the architecture for harvesting and storing the data have to be agreed on. Because archival description can be conducted on multiple levels (from aggregations of records to single documents), it is also necessary to agree on which level of detail description should be shared. This question still needs some investigation, especially into the user-friendliness of high level and contextual description. The question of common standards was important before the Internet was widely used. In today's World Wide Web, things have become a bit more flexible. On the one hand, the Dublin Core metadata initiative which allows to describe resources independently of their format or type is being increasingly embraced by all kinds of providers of information on the Web. On the other hand, the relatively general character of the Dublin Core metadata set can be offset by linking retrieved data to the database of the institution which originally created them. The user can then search the original database which might fulfill more sophisticated requirements. An example for sharing metadata is an award-winning portal to pictorial resources, PictureAustralia. (6) The contributing organizations each map metadata from own online-accessible databases to the Dublin Core metadata set. These metadata are harvested on a regular basis and stored in a central database which is then queried by the portal's visitors. The users receive thumbnails of images and the pertinent Dublin Core metadata in the shared area, but are directed to the providing institution's online database if they want to see an enlargened version of the picture and learn more about its context. Contributing to the network therefore adds only a small amount of work and costs to the institutions, does not influence their internal standards and practices, and is popular because the holdings are much more readily accessible. Though it might not have all the advantages of a highly standardized central database, the simplicity of the model guarantees maximum flexibility to the contributors and allows them to cater best to the needs of their patrons. It is here that the competitive element between heritage institutions battling for user frequencies and public and private funding takes effect: The joint database increases the chance of their resources being found in the first place, but it is the organization's own collection as well as its website and/or database(s) with its informational, navigational or visual qualities which can constitute its advantage over others.

BUILDING SPECIALIZED NETWORKS

The approach taken above is an important step to link cultural resources to one another. However, it still has an institutional starting point, and the kind of resources linked to each other will be determined by the participants in the partnership. For popular topics, we

believe a truly subject-based entry point would be preferable for users. This allows broader participation by giving everyone the opportunity to link their resources to the network. For example, individuals, families, libraries, archives and museums could contribute documents, photographs and objects (references to and/or digital representations of) Swiss emigrants to South America. Maybe some archives in Uruguay has records of the Swiss teacher we were looking for and is able to publish them through this network. This more informal approach might lead to less organized and maybe also less reliable information. But if it works, it can be much more than an information resource: It reflects what people are interested in, it is dynamic and diverse, giving records from non-institutional custodians equal weight to those of renowed organizations, and opening new access points to collections. Hopefully, it will animate debate and provide a wider perspective and more mutual respect for all resource providers involved.

CONCLUSIONS

In the case of networking historical sources, technology does not have such a great impact on *what* we are doing as on *how* we are doing it. The first generation of online archival databases has greatly simplified the previously tedious use of archival finding aids. Now we must begin to design applications which assist the users in actually finding what they are looking for. With the help of user studies, we hope to be able to guide our users in a better way according to their experience and the questions they ask, and to make archival finding paths more transparent. In order to achieve this, we must provide more information about the context of records and use this contextual information to link records in distributed locations.

ABOUT THE AUTHOR

Andrea Rosenbusch has been working as an archivist since 1993, mainly at the Swiss Federal Archives. In 2000, she spent six months at the National Archives of Australia to specialize in different aspects of electronic records. As an historian, she has published on innovation in the Swiss chemical industry in the 1920s and 30s and is regularly involved in making historical exhibitions.
E-mail:
andrea.rosenbusch@bar.admin.ch.

REFERENCES

1. Wendy Duff and Catherine A. Johnson: A Virtual Expression of Need: An Analysis of Email Reference Questions. Paper held at the Society of American Archivists' Annual Meeting in Denver, August 28 – September 3, 2000. In: *American Archivist* (forthcoming).

2. Andrea Rosenbusch, Are Our Users Being Served? A Report on Online Archival Databases. In: *Archives and Manuscripts*, May 2001.

3. Gavan McCarthy, Utilizing the Web to Build a Network of Archival Authority Records. In: *Janus* 1999.1, pp. 96 – 106, especially pp. 98f., and the exemplary site ‚Bright Sparcs' by the Australian Science and Technology Heritage Centre at *http://www.asap.unimelb.edu.au/bsparcs/.*

4. National Scholarly Communications Forum, Round Table No. 10, November 1999, Archives in the National Research Infrastructure, Communiqué and

Resolutions, resolution no. 6, *http://www.asap.unimelbedu.au/nscf/roundtables/r10/r10_resolutions.html.*

5. *http://www.unesco.org/webworld/portal_archives/.*
6. *http://www.pictureaustralia.org/.*