

(2001): 73-79. Publisher's official version: <http://acts.metapress.com/content/lm15251/Ovr3mr36/>. Open Access version: <http://kuscholarworks.ku.edu/dspace/>.

[This document contains the author's accepted manuscript. For the publisher's version, see the link in the header of this document.]

Paper citation:

Russell, Beth M. with Jodi L. Spillane. "Using the Web for Name Authority Work." *Library Resources & Technical Services* 45 (2001): 73-79.

Abstract:

While many catalogers are using the Web to find the information they need to perform authority work quickly and accurately, the full potential of the Web to assist catalogers in name authority work has yet to be realized. The ever-growing nature of the Web means that available information for creating personal name, corporate name, and other types of headings will increase. In this article, we examine ways in which simple and effective Web searching can save catalogers time and money in the process of authority work. In addition, questions involving evaluating authority information found on the Web are explored.

Text of paper:

Using the Web for Name Authority Work

Beth M. Russell and Jodi Lynn Spillane

Beth M. Russell is Assistant Professor and Head, Special Collections Cataloging at the Ohio State University Libraries in Columbus.

It would be an understatement to say that the technology of the Web has affected all library operations. Name authority work is no exception. The vast resources accessible on the Web allow catalogers to verify the identity of authors, clarify corporate body relationships, and contact people for further information much more quickly and accurately than ever before possible. It would be foolish for catalogers not to exploit this resource in the time-consuming process of authority work. No doubt many catalogers are already using some of these techniques in their daily work, but the full potential of the Web to assist catalogers in authority work has yet to be realized. Businesses and organizations create Web pages that provide valuable information about their identity that can be referred to when cataloging the publications of these bodies. Personal Web pages, directories, and other types of information are also becoming much more common, allowing catalogers access to vast stores of information. As search engines become more sophisticated, and as catalogers become more sophisticated searchers, the time spent locating information needed to perform authority work will hopefully decrease even further. However, with the ease and convenience of Web searching comes the possibility for further confusion and a need to redefine acceptable sources for authority information.

Definitions

What do we mean when we talk about "using the Web (or authority work)?" A conservative definition for authority control is "the process of ensuring that every entry—name, uniform title, series, or subject—that is selected as an access point for the public catalog is unique and does not conflict, by being identical, with any other entry that is already in the catalog or that may be included at a later date" (Clack 1990, 2). Typically, authority control is thought of as a process related to cataloging, performed either at the point of cataloging, when headings are constructed for use in a

bibliographic record, or afterward, when headings already in the catalog are checked against an authority file. The process of authority control, however constructed, enhances the finding and collocation functions of the catalog (Talmacs 1998). Recent debates about the importance of authority control in an automated environment notwithstanding, most libraries still face the reality of relying on an authorized catalog in order to provide access to all their materials (Talmacs 1998). In order to create or edit existing name authority records that are accurate, complete, and useful, several types of information are needed. First, the identity of the person or corporate body must be verified. For example, if a piece being cataloged states the author is named "John Smith," the cataloger would need not only to figure out who *this* particular John Smith is, but to discover enough information to create a distinct heading that will set this author apart from other John Smiths. For this reason, it is often necessary to find other sources of information in order to resolve conflicts, such as birth and death years, middle initials or names, etc. Name changes and pseudonyms are another problem area in which catalogers often have to track down more information than is presented on the piece being cataloged. For corporate body names, it is sometimes necessary to determine what the heading should be by finding additional information. For example, a piece to be cataloged may present several forms of the name, and only through additional investigation can the cataloger determine "the name by which it is commonly identified" as required by current cataloging codes (AACR2R 1988, 441). Finally, while some types of information might not be necessary in a heading, they might be useful in supplementary fields in the authority record, such as 670 field (Source Data Found) or 678 (Biographical or Historical Data). For example, the place of business for a corporate body might not be necessary to resolve a conflict, but by including this in a note, future conflicts can be avoided.

In this article, we focus on authority work for the cooperative NACO project (the name authority program component of the Program for Cooperative Cataloging), although the same techniques could be used to locate information for local authority files or other types of files. Certainly information gleaned from any number of Web pages might be used to clarify subject-heading proposals as well. Finally, although the phrase "establishing a heading" is usually used to indicate the process of determining a heading for a specific authority file, it should be remembered that every time a cataloger constructs a new heading for a bibliographic record, a heading has, in some sense, been "established."

Role of Authority Control

It can be argued that authority work, and controlled vocabulary in general, has no place in the world of keyword retrieval and Web search engines. Even proponents of a modified concept of authority control state that keyword searches can lead to retrieval of relevant and useful records without complete recall and that some names, such as those used as subjects, might be more important than others to have under authority control (Younger 1995). However, the promise of the Web as a searching tool offers many examples of how misleading completely uncontrolled vocabulary can be. For example, entering the name "Beth Russell" in the popular Internet search engine Altavista.com retrieves several different Beth Russells. Most common by far is a creator of needlepoint designs, but countless other Beth Russells also appear, ranging from several affiliated with various universities to an "S. Beth Russell" of the Licensed Professional Counselors Association of Georgia. The problem is even further complicated when database searches related to cataloging, such as the keyword search in the OCLC database, regard "Beth" and "Russell" as separate entities, and search for the words individually rather than together. For most people doing basic searches on the Web, this strategy might be employed by default, often without the user knowing how their search is being conducted. People searching any database by keyword may be unaware of the way in which their results are selected and organized.

This means, in real terms, that even if someone is looking for a particular work, rather than works by a given author, results can be so numerous as to be confusing. Authority control (or controlled vocabulary in general) allows for the results of a search to be narrowed down to a relevant set. This is particularly useful in the case of pseudonyms and name changes.

NACO and USMARC

NACO participation has always been rather high-tech; the program depends on technology to allow submission and review of records, as well as searching of various databases to verify usage, headings, etc. It allows a certain degree of flexibility, however, in using other sources, such as in-house files and nonbibliographic sources.

The NACO Participants Manual (NPM) explicitly mentions citing Web home pages in the 670 field (Source Data Found):

Give the name of the home page, gopher, etc. and the date it was consulted. In subfield 1b, give a location, if appropriate, and the information found. Generally, don't include the URL (Uniform Resource Locator) since the address often changes (NPM 1996, 61).

One objective of the NACO project is "to increase the timeliness of cataloging copy" (NPM 1996. I). This is certainly a major strength of the Web; compared to earlier means of verifying names or resolving conflicts, the Web offers unprecedented speed. Since authority work has been proven to be one of the most costly parts of cataloging, greater speed translates into greater cost-effectiveness.

The USMARC format for authority records is also itself undergoing changes in response to the growing prominence of Web pages. In 1999, *MARC 21 - Authority* was updated with a definition of the 856 field to allow the inclusion of URLs for corporate body name authority records. This change was designed as "a potential means of improving access to information about the entity covered by the authority record" (MARBI Proposal No. 98-13). The basis for this decision is to allow easy access to more substantive information than can realistically be included in the authority record. The potential exists to expand the use of the field to other types of authority records, such as those for personal names, conferences, or even uniform titles. Certainly this new definition of the 856 points to greater acceptance of the role of Web pages in authority work, as well as signals new possibilities in integrating access to Web pages within library catalogs.

Locating Information on the Web

It is no secret that authority control is an expensive process, incorporating time spent in the operations necessary to achieve it (Clack 1990). The cost of authority control can be reduced through using the Web by decreasing reliance on expensive and out-of-date print sources. In addition, the use of macros and cut-and-paste functions, common across numerous software applications, can greatly reduce the time spent searching and manually entering data, as well as the possibility for error. Perhaps the simplest and most traditional way of using the Web for authority work is to provide an easy way to contact individuals directly. Online directories often list e-mail addresses or phone numbers, and locating the person for further information can be just a click away. The NACO manual offers examples of how a citation to a phone call or e-mail can be constructed in an authority record.

At least one cataloging department has embraced this concept in its documentation. The Princeton University Cataloging Department Web page on resolving personal name conflicts explicitly suggests trying to locate authors on the Internet and e-mailing them (or their publishers) to resolve conflicts (Princeton University Library 2000a).

Figure 1 illustrates how easily contact information can be located. A report originating from Great Britain's Transport Research Laboratory lists the author as "C Corbett" and lists an institutional affiliation with Brunei University. From the Web site of the university, it was easy to locate the authors e-mail address and send an e-mail asking for verification that "C. Corbett" was the "Claire Corbett" represented in the authority file. Also, since this piece had usage of "C Corbett," a cross-reference for that name was added, in addition to changes requested by the author.

Figure 2 illustrates this process further. In this case, the author had published a work under an earlier form of her name, which appeared in the OCLC database. A new publication with a different form of name prompted the cataloger to contact the author, whose e-mail was located through the Web site of her institution. In this case, the communication with the author not only allowed the preferred form of the name to be constructed as the authorized heading, but also assured the accuracy of the very useful cross-reference.

In establishing headings for personal names found in works published in Canada, or for corporate and geographic names within Canada, a cooperative agreement between the Library of Congress and the National Library of Canada requires verification of headings in the National Library of Canada's catalog.

Catalogs for other national libraries, while not strictly required to be consulted before establishing a name heading, can prove very useful, and Web interfaces for these catalogs make the process even easier. In Figure 3, for example, an authority record was constructed based on information found on the Web site interface of the National Library of Australia. Although this site lacks the "authority" inherent in headings from the National Library of Canada, it can prove very useful for providing corroborative information, such as verifying and distinguishing names. Catalogers working with materials from foreign countries ought to familiarize themselves with national library catalogs within their realm of expertise; these catalogs can be cited just like any other Web site.

The proliferation of Web pages for companies, organizations, and other corporate bodies makes it quite easy to consult these sources in doing authority work. First of all, corporate body Web pages often have clear contact information, including addresses or phone numbers, that can either be used to identify corporate bodies and resolve conflicts, or to contact the body for more information. Also, businesses frequently have histories or other public relations information on their pages that can be useful in constructing the 670 field.

An example of the usefulness of consulting corporate body Web pages can be found in figure 4 (Educational Human Resource Development). The piece in hand did not provide enough information to create a meaningful heading. The name on the piece, "Department of Educational Human Resource Development," contains "a term that by definition implies that the body is part of another" and clearly needs more information (AACR 1988, 24.13A Type 1). By consulting the home page for the body, its relationship to the College of Education and Texas A&M University is clarified. The banner along the side of the page indicates that Educational Human Resource Development is one department within Texas A&M University's College of Education.

A more abstract way in which Web pages can present their identities involves the use of clickable links to other bodies often found at the bottom of Web pages, sometimes even in graphics or banners. For example, it is common practice in many environments to provide a standard set of links to the larger organization at the bottom of each page on the university's server page, or, as was the case with the Educational Human Resource Department, in frames along the side. This format makes it easier for users to navigate among sites and prevents them from "getting lost" while clicking further and further into a complex hierarchy of pages. While the NACO manual does not expressly allow for citing this kind of information as proof of relationship among corporate bodies, it can be used to clarify what might be suspected from the piece in hand or from some other source.

Corroborative evidence can also be gleaned from URLs, since the naming structure of servers will often indicate a subordinate relationship.

Just as Web pages can be used to clarify information about individual and corporate bodies, series information can often be located and verified on the Web. Figures 5 and 6 illustrate a site with a huge amount of information about the publications of the company, and pieces within the Luftwaffe Profile Series are clearly indicated. While these lists would not take precedence over the actual piece as published, they do allow a quick way to verify the existence of a series and its component parts.

Even more interesting are figures 7 and 8. Here the series name appears to vary even among pages from the same university. "Qedem, Monographs of the Institute of Archaeology of the Hebrew University of Jerusalem" on the Web page of the Israel Exploration Society is cited as "Qedem Monographs" on the Web page of the Publications of the Institute of Archaeology, Hebrew University of Jerusalem. In this case, further information does not necessarily resolve a conflict, but provides catalogers with more information to use in constructing a logical and useful heading, using their own judgment. Interestingly, both sites were easily located simply by searching the phrase "Qedem monograph" in a major Internet search engine. With more common names, such retrieval may be more difficult.

In addition to personal and institutional Web pages, other Web-based sources can be used to provide information useful to authority work. For example, major commercial databases such as Amazon.com can verify an authors publishing history by quickly linking to other published works. This is most often necessary with common names; it is often difficult to tell if the author of a work in hand is the same as the author in a catalog, until it is discovered that the person has written several works in the same subject area or genre. Especially if a cataloger is already working in a Web browser, going to a site such as Amazon might be quicker, and certainly cheaper, than searching OCLC for preliminary research.

The *NACO Participants Manual* requires that the national authority file be searched, either through OCLC, RUN, or the Library of Congress database. If no heading is found in the NAF, bibliographic records in OCLC or RLIN are searched to determine the predominant form. The LC records must also be identified to determine if bibliographic file maintenance is required (NPM 1996). The judicious use of other resources in conjunction with OCLC or RLIN can speed up the process immensely for certain types of searches. A final category of Web-based sites that can be useful for authority work includes those that duplicate or update paper reference sources. Most catalogers will have their own set of sources that are used most frequently based on the type of material being cataloged. For example, The Handbook of Texas Online (www.tsha.utexas.edu/handbook/online/) incorporates the complete text of the first edition of this comprehensive work, changes from the second edition, and articles that were not included because of space limitations in the print editions. Corrections, updates, illustrations, and additions will be added, and the site includes both browsing and searching functions.

Figure 9 illustrates how a general-interest, Web-based fee site can provide useful information in constructing authority records. Here, Gale Group's Contemporary Authors database provided the biographical information needed to correctly identify Joseph Foster and O'Kane Foster as the same person. While this information could have been found in other sources, or in older (print) editions of the title, a cataloger frequently working with modern fiction would benefit from using online sources such as Contemporary Authors.

For people with fast and reliable Internet access and free access to a wide variety of online sources, using Web-based versions of reference sources can be much quicker and easier than relying on often out-of-date paper versions that may be located in other areas or even other buildings. The sources used will vary based on the subject matter, geographic area, or language focus of the work in

hand, but increasingly, catalogers can find electronic versions of their standard favorite tools, as well as other, "born digital" sources of the same information.

Of course, catalogers will consult some Web pages only once, to resolve a question about a piece in hand. Others may prove more useful and catalogers will wish to consult them frequently, especially if dealing with similar material. In this case, bookmarking Web pages is probably the best way to have these sites easily accessible (Wisniewski 1998).

Evaluating Authority Information Found on the Web

As suggested by some of the above examples, finding forms of names on the Web is not necessarily the end of the process in performing authority work. Two major issues emerge. First of all, are Web sources reference sources, and are they equivalent to print reference sources? Princeton University's Cataloging Department has answered "yes" to the first question, stating that Web sites other than a corporate body's own home page should be treated "as a reference source, with respect to the form of name of a corporate body found there" (Princeton University Library 2000b). This is consistent with the NACO Participants Manual and with common practice for a wide range of Web sites.

The breadth of sources that could be consulted by a working cataloger precludes a definitive statement of what is and is not a reliable reference source on the Web, but analogies to print publications continue to be useful. Works clearly presented as reference sources, whether free or fee-based, would correspond most closely to traditional references sources, while pages created by a fan of an author, for example, should be treated with a healthy degree of skepticism.

The second question regarding validity of information stems from the fact that Web sites can function as publications by themselves. Princeton University's Cataloging Department asserts in its internal documentation, "The formal presentation of the name of a corporate body on the body's own home page (i.e., at a Web site created or sponsored by that body), is tantamount to formal presentation on the chief source of one of the body's own publications. That is, it has special status as a source for the AACR2 heading" (Princeton University Library 2000). Not all catalogers may be comfortable with the idea of "special status," given the long history of privileging the chief source of information of print publications in questions of authority work.

The validity of information on the Web is certainly part of a much larger debate than can be explored in the current discussion. Nonetheless, in the authors' experience, common sense and cataloger judgment resolve most conflicts, with the benefits of locating name information on the Web far outweighing the intellectual issues such a practice sometimes raises.

Conclusion

Anyone who has turned to the Web for more information about a person or corporate body has, at some time or another, been frustrated. Search engines often return false hits, even when relatively sophisticated searches are constructed. Since corporate body names often change, and Web pages often change servers, matching up the name found on the piece in hand with a particular Web site can be difficult, especially when cataloging noncurrent material. For much older cataloging, the Web might not be as useful, although the increasing conversion of "classic" reference sources to Web interfaces can offset this shortcoming.

The authors' experience has shown that authors affiliated with educational institutions are the easiest to locate and corporate bodies that are governmental in nature tend to have the most helpful and reliable sites. Also, commercial sites, once located, can be very useful in providing more information about businesses.

Often the Web provides no help beyond a contact, but when it is easy to search for the contact information, this can be quicker than sorting through out-of-date directories or phone books. When the site provides more information, a great deal of time can be saved by simply citing the Web page itself, and not having to contact a "real" person in order to get the necessary information.

While this survey cannot be comprehensive, we hope that it has provided food for thought for catalogers, whether they are simply establishing a heading for use in a bibliographic record or contributing a heading to the NACO file. The Web is no doubt changing the way we do authority work, and by keeping our minds open to new strategies, we can use this to our advantage to be more productive and more efficient, without sacrificing the quality of our online catalogs and other bibliographic resources.

Figures

010 n 86100083
040 DLC \$c DLC \$d TxCM \$d DLC
005 19980504143231.5
100 1 Corbett, Claire
400 1 Corbett, C. \$q (Claire)
670 Maguire, M. The effects of crime and the work ... c1987: \$b CIP t.p. (Claire Corbett; research fellow, Centre for Criminological Research, Univ. of Oxford)
670 Deterrence of high speed driving, 1998: \$b t.p. (C. Corbett, Brunel University)
670 Email from author, April 29, 1998 \$b (Full name is Claire Louise Corbett. criminologist: does not wish her birthdate used)

Figure 1. Authority Record for Claire Corbett

1 010 no 00053569
2 040 OU \$b eng \$c OU
3 005 20000728064040.0
4 100 1 Conklin, Nikki Lynn Schafer, \$d 1951-
5 400 1 Schafer, Nikki L., \$q (Nikki Lynn, \$d 1951-)
6 670 Perspectives on issues programming in the Ohio Cooperative Extension Service, 1989: \$b cover (Nikki Lynn Schafer Conklin)
7 670 OU files \$b (Conklin, Nikki L.; Assoc. Prof., Agricultural Education, Ohio State University)
8 670 Email from author, Aug 12, 1999 \$b (b. 1951; maiden name is Schafer, used as undergraduate)
9 670 OCLC database, July 27, 2000 \$b (hdg.: Conklin, Nikki Lynn Schafer, 1961- ; usages: Nikki L. Conklin; Nikki Lynn Schafer Conklin)

Figure 2. Authority Record for Nikki Lynn Schafer Conklin

1 010 no 98074197
2 040 TxCM \$c TxCM
3 005 19980326051225.4
4 100 1 Phillips, Marc, \$d 1969-
5 670 International mixed drinks, 1995; \$b t.p. verso (Marc Phillips)
6 670 Natl. Lib. of Australia web catalogue, Mar. 13, 1998 \$b (hdg.: Phillips, Marc, 1969- ; usage: Marc Phillips)

Figure 3. Authority Record for Marc Phillips

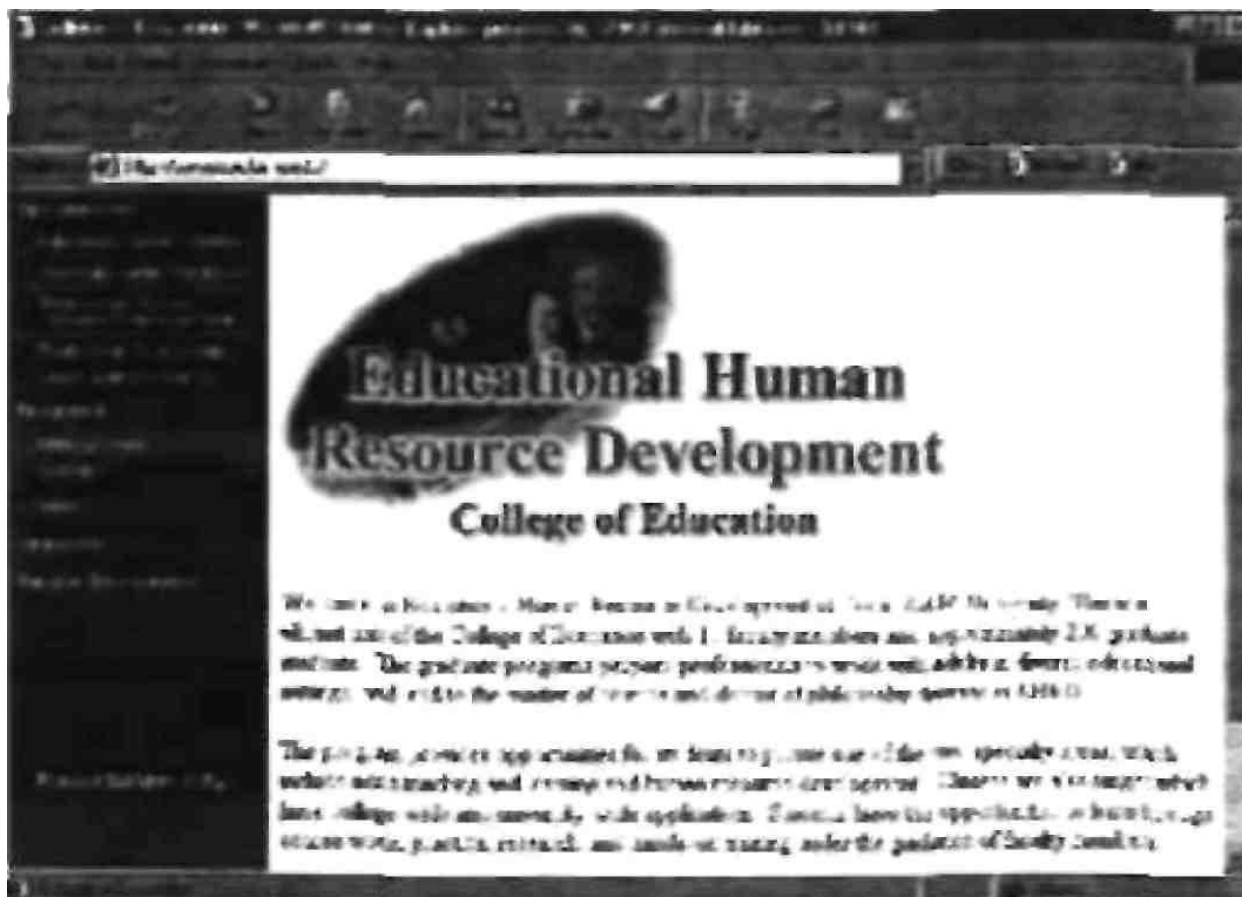


Figure 4. The Texas A&M University Educational Human Resource Development Home Page

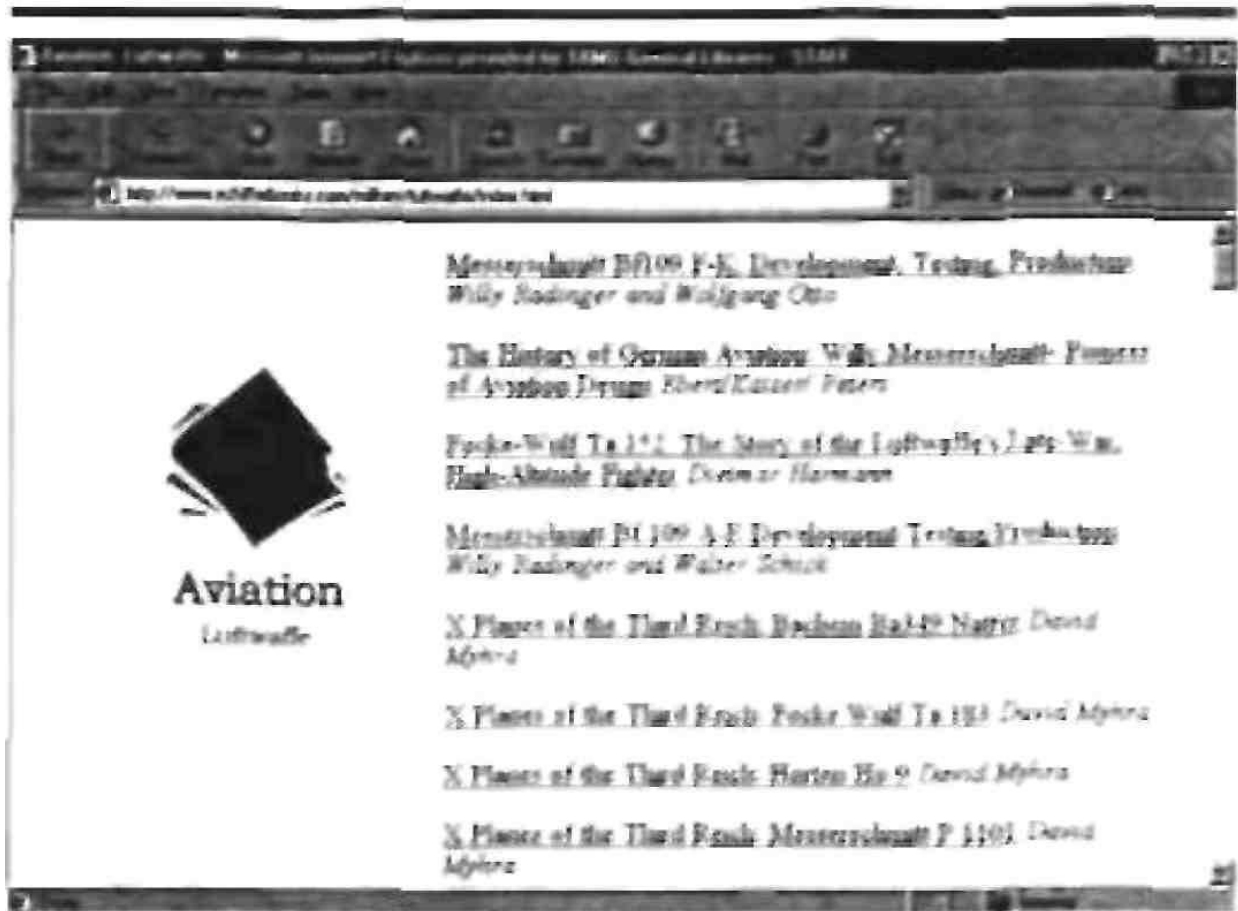


Figure 5. Schifter Military History Aviation Web Site

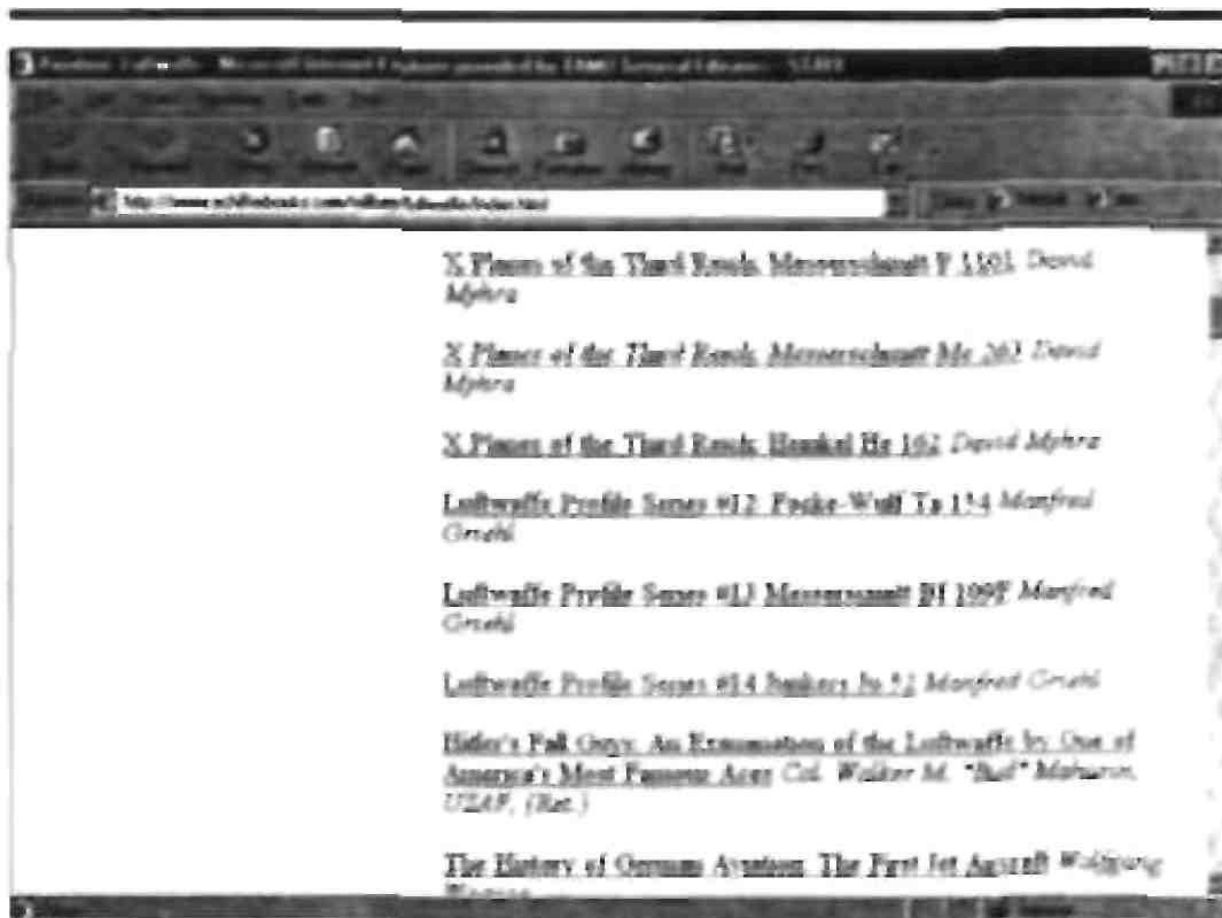


Figure 6. Pieces within the Luftwaffe Profile Series in the Schiffer Military History Aviation Web Site

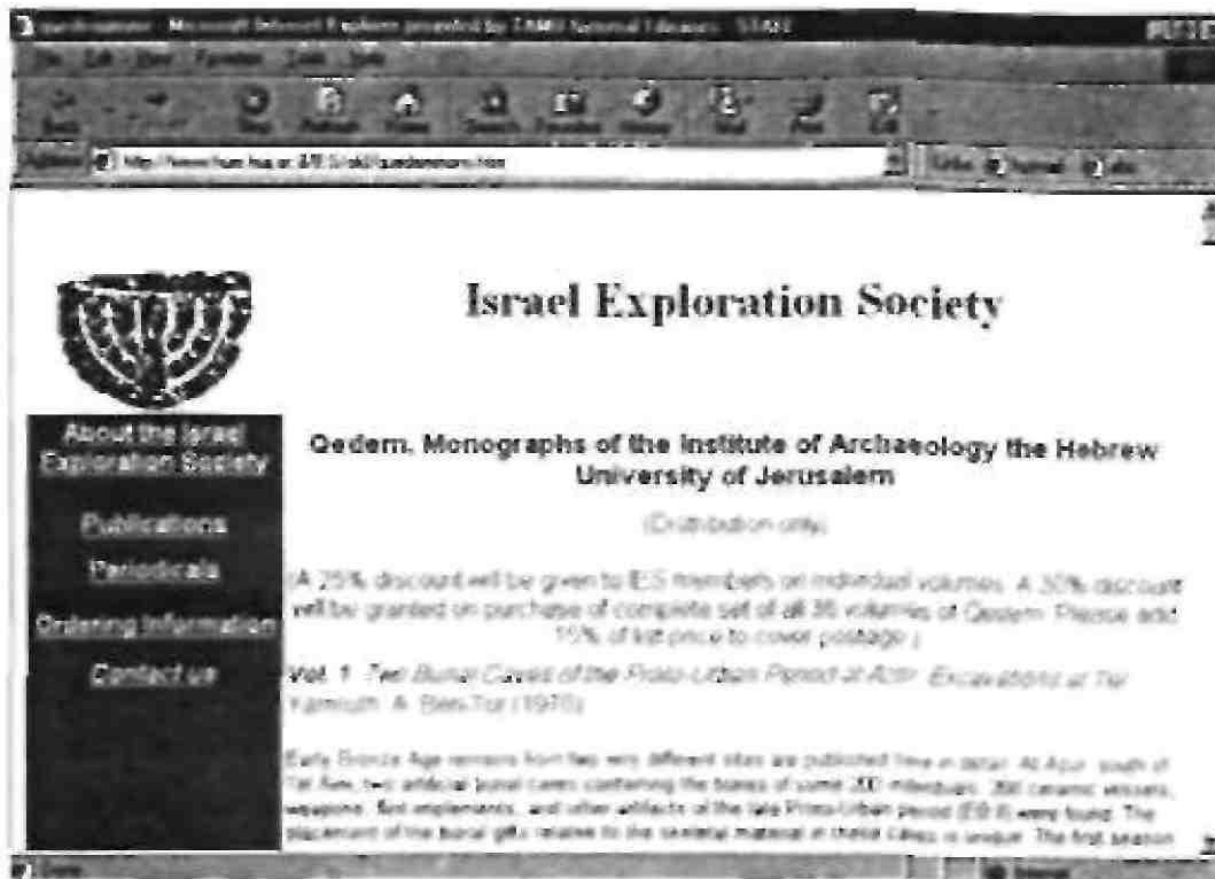


Figure 7. Israel Exploration Society, Qedem Home Page

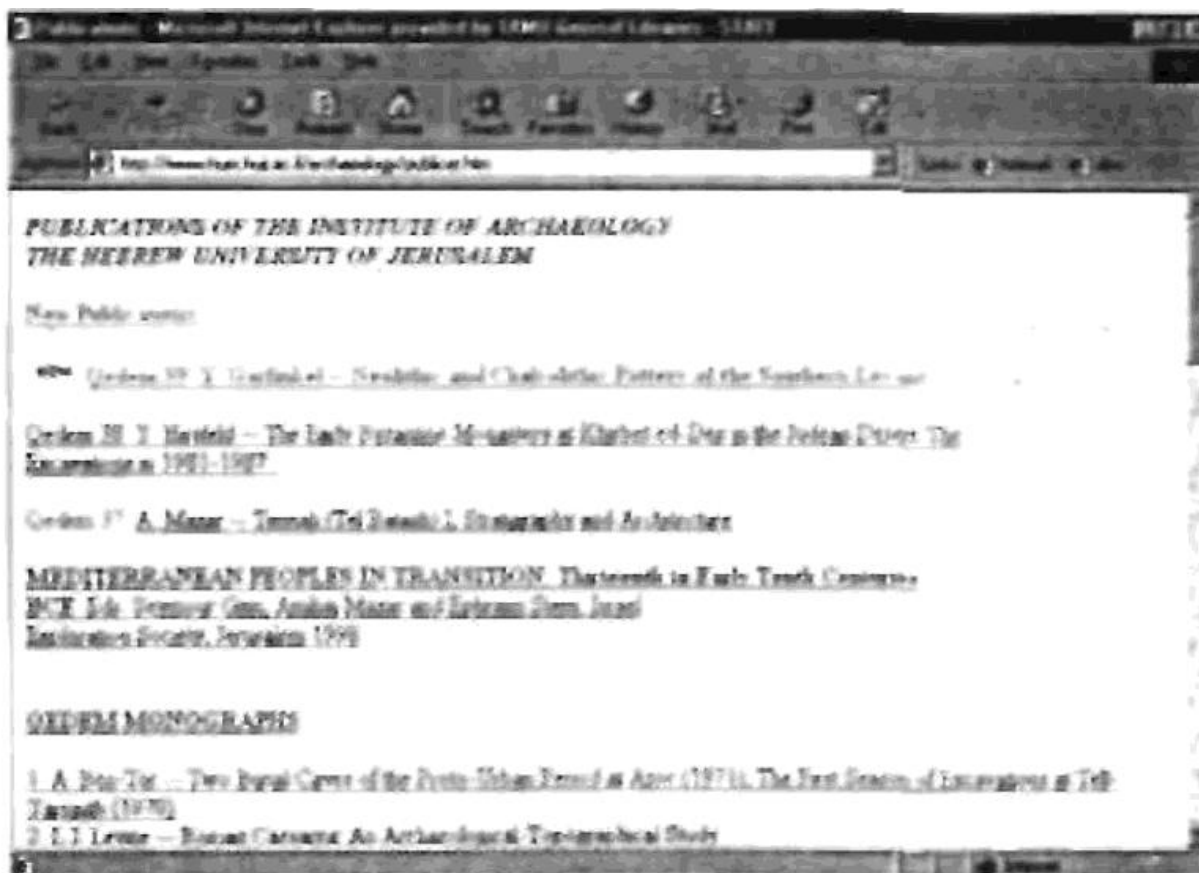


Figure 8. The Hebrew University of Jerusalem, Institute of Archaeology's Publication Home Page

- 1 010 nr 89017413
- 2 040 NjP \$c NjP \$d OU
- 3 005 20000804064840.0
- 4 100 1 Foster, Joseph
- 5 400 1 Foster, Joseph O'Kane
- 6 400 1 Foster, O'Kane
- 7 670 His D.H. Lawrence in Taos, c1972; \$b t.p. (Joseph Foster)
- 8 670 In the night did I sing. 1942: \$b t.p. (O'Kane Foster)
- 9 670 OCLC database. July 27, 2000 \$b (hdg.: Foster, Joseph; usages: Joseph Foster; O'Kane Foster) (hdg.: Foster, Joseph, b. 1898 ; usage not given) (hdg.: Foster, Joseph O'Kane; usage: Joseph O'Kane Foster) (hdg.: Foster, Joseph O'Kane, 1898-; usage not given)
- 10 670 Contemp. authors, July 27, 2000 \$b Galenet: contemp. authors (Foster, Joseph O'Kane, 1898-; also known as O'Kane Foster, b. June 27. 1898 in Chicago, IL.; reporter, author)

Figure 9. Authority record for Joseph Foster

Works Cited

- Anglo-American cataloguing rules*. 1988. 2d ed., 1988 rev. Ottawa: Canadian Library Assn.; London: Library Assn. Publ.; Chicago: ALA.
- Clack, Doris Hargrett 1990. *Authority control: principles, applications, and instructions*. Chicago: ALA.
- Institute of Archaeology. The Hebrew University of Jerusalem. 2000. Publications of the Institute of Archaeology, the Hebrew University of Jerusalem. Accessed June 6, 2000. www.hum.huji.ac.il/archaeology/publicat.htm.
- Israel Exploration Society. 2000. Qedem, Monographs of the Institute of Archaeology, the Hebrew University of Jerusalem. Accessed June 6, 2000. www.hum.huji.ac.il/IES/old/quedemmono.htm.
- Library of Congress. 2000. MARC 21 concise authority for authority data: Other variable fields (8XX). 856 Electronic location and access. Accessed May 30, 2000. www.loc.gov/marc/authority/ecadothr.html#mrca856.
- Library of Congress Network Development and MARC Standards Office. 2000. MARC Standards, Proposal no.: 98-13. Accessed May 30, 2000, <http://lcWeb.loc.gov/marc/marbi/1998/98-13.html>.
- NACO participant's manual*. 1996. 2d ed. Washington, D.C.: Library of Congress, Cataloging Distribution Service.
- Princeton University Library. 2000a. Princeton University Library's Cataloging Documentation. Name Authorities: Conflict Resolution. Accessed May 30, 2000, <http://libweb.princeton.edu/katmandu/cp20/confperl.html>.
- Princeton University Library. 2000b. Princeton University Library's Cataloging Documentation. Name Authorities: NACO Information, Types of Reference Services: Internet Resources. Accessed May 30, 2000. <http://libweb.princeton.edu/katmandu/cp20/refinternet.html>
- Schifferbooks.com. Military History. 2000. Aviation, Luftwaffe. Accessed May 30, 2000, www.schifferbooks.com/military/luftwaffe/index.html.
- Talmacs, Kerrie. 1998. Authority control. In *Technical services today and tomorrow*. Englewood, Colo.: Libraries Unlimited.
- Texas A&M University, College of Education. Educational Human Resource Development. 2000. Educational Human Resources Development home page. Accessed May 30, 2000. www.coe.tamu.edu.
- Wisniewski, Julia. 1998. Authority work. Internet resources, and a cataloger's home pages. *Art Documentation* 17 (1): 45-46.
- Younger, Jennifer A. 1995. After Cutter: Authority control in the twenty-first century. *Library Resources & Technical Services* 39: 133-41.