Is the Current Way of Constructing Corporate Authority Records Still Useful?
Qiang Jin

Catalogers have been establishing corporate body name headings (and other entities) in their original language in the official form, as they appear most frequently on the title pages of publications for print publications, for many years. A random sample of corporate headings from the Library of Congress Name Authority File created during 1998–2002 was searched on the Web via Google to find corporate Web pages. The purpose of this research is to begin to answer the question: Does the current way of constructing corporate authority records still help users find resources by and about corporate bodies in the online public access catalog in this Web-oriented environment?

For years, companies, businesses, associations, and scholarly societies have produced such valuable information as reports and research papers, which are essential to users’ research needs. Catalogers have worked hard for decades to provide users with resources by and about corporate bodies through library card catalogs and online library catalogs. It has proved to be very successful, as the name of the author and the title of the book were the most important clues by which the books were identified when cited and by which they were looked for in the catalog or called for in the library. However, the Web has changed the way users access information for their research needs. Does the current way of constructing corporate authority records still help users access resources by and about corporate bodies in the online public access catalog (OPAC) in this Web-oriented environment?

Context

A 2002 survey by the Digital Library Federation and the Council on Library and Information Resources indicated that most college students search the Internet before they search the OPAC for their research needs. A 2002 OCLC white paper, titled “How Academic Librarians Can Influence Students’ Web-Based Information Choices,” pointed out that “students continue to depend upon the library for information resources, in both electronic and print formats. The data strongly suggest that there are real opportunities for academic librarians to connect students with libraries’ high-quality resources.” Another report, “Web Growth Slows, but Time Online Rises” by Miller in USA Today, indicated that although Internet growth in the United States continues to slow, users are logging on more often and spending more time online. More than 457 million people worldwide had Internet access at home in 2001; the number of active global users hit 254 million.

Our users are searching the Web for their research needs, and the Internet is providing them with many alternative names for corporate entities. How important is OPAC authority control in the age of the Internet? What do users need to find the resources regardless of how they express the corporate name?

Authority control is important. Taylor defined the authority file in her article “Authority Files in Online Catalogs,” stating that “authority work [is] the process of determining the form of a name, title or subject concept that will be used as a heading on a bibliographic record; determining cross references needed to that form; and determining relationships of this heading to other authority headings.” Burger explained in his book Authority Work that corporate names, like personal names, come into existence when the body itself is created. Unlike personal names, however, corporate name changes occur with much more rapidity and for different reasons, and hence are treated differently by cataloging codes. Furthermore, corporate bodies can survive indefinitely. The longer lifespan of such entities provides a greater potential for change. Several recent studies indicate that the Web can also help with name authority work. Russell and Spillane explained in their article “Using the Web for Name Authority Work” that the Web could provide reference sources in finding information on personal and corporate authors. Jin proposed in the article “Creating Up-to-Date Corporate Name Authority Records by Using Official Corporate Home Web Pages” that catalogers could use current information on corporate Web pages to create up-to-date corporate authority records even if the form of a corporate name on the title pages of its publications matches the form in the Library of Congress Name Authority File (LC NAF).

Tillett pointed out in her article “Authority Control: State of the Art and New Perspectives” that:

the Web environment opens up new uses for authority records and new objectives to augment the traditional objectives. . . . We are moving from the stand-alone authority file of a single institution, or even from the shared online files towards a goal of sharing authority files among all communities. Shared authority information has the added benefit of reducing the global costs of doing authority work while enabling controlled access and better precision of searching."

Qiang Jin (qjiang@uiuc.edu) is Original Cataloging Librarian and NACO Coordinator at University of Illinois at Urbana-Champaign.
In his *Code of Cataloging Rules, Author and Title Entry: An Unfinished Draft* (1960), Lubetzky stated the general rule for authorship: “The work of a person, whatever its character or the medium by which it is presented, is entered under the name of the person as author of the work presented (Rule 1)”. Lubetzky continues:

Rule 21: A work of corporate authorship presents complexities not found in a work of personal authorship. These are treated in accordance with the following rules; otherwise a work of corporate authorship is treated the same as one of personal authorship. Rule 22: A work which, explicitly or implicitly, represents an act, communication, or product of the activity of a corporate body is entered under the name of that body.

According to Lubetzky “Corporate bodies, like individuals, are generally known and referred to by their names and should therefore be entered directly under their names as individuals are.”

Between October 9–18, 1961, an international conference on cataloging principles organized by the International Federation of Library Associations under the sponsorship of the Council on Library Information Resources and the United Nations Educational, Scientific, and Cultural Organization was held in Paris. Representatives from fifty-five nations around the world reached an agreement on the choice of the name of the corporate author in the original language in the official form, as it appeared most frequently on the title pages of publications. The reason for this agreement was “behind the form of heading for corporate authors was the form to be adopted for corporate headings aiming at following as closely as possible the same principles at those which regulated the choice of heading of personal authors.” Lubetzky’s rule for entering corporate body directly under the name of that body was adopted during the Paris Conference.

AACR1 and AACR2 followed the Paris Principles, and thus catalogers are instructed to “enter a corporate body directly under the name by which it is commonly identified. . . If variant forms of the name are found in items issued by the body, use the name as it appears on the chief sources of information as opposed to the forms found elsewhere in the items.” “The chief source of information for printed monographs is the title page.” These rules proved to be very successful for retrieval of materials by and about corporate bodies for card catalogs as well as for online library catalogs for many years because, as Lubetzky noted:

the prevailing type of materials which is found in a library and is recorded in its catalog is the book; and the most important characteristic of the book, for the purpose of cataloging, is the fact that it is provided with a prominent identification tag in the form of a title page. The cataloger can thus anticipate how a particular book will normally be cited and looked for and provide for it accordingly. The title page generally includes the name of the author and the title of the book, sometimes only the title of the book, and occasionally only the name of the author. The name of the author and the title of the book are therefore the most important clues by which the book will be identified when cited, and which it will be looked for in the catalog or called for in the library.

However, with the rapid development of Internet technology in recent years, users’ searching behavior has changed greatly. Local OPACs as well as corporate Web pages are now easily accessible over the Web through such search engines as Google and others. As users move between the Web and OPACs as they search for information for their research needs, they will find the forms of corporate names on corporate Web pages and in the LC NAF sometimes differ. This may create confusion and uncertainty that may inhibit users from accessing materials by and about those corporate bodies in the OPAC. Because there are now other equally important clues, such as corporate Web pages, available for users to find on the Web, catalogers may need to revise the cataloging rules that Lubetzky created decades ago in order to help today’s users find resources by and about corporate bodies in the OPAC. A question arises: If a user searches the OPAC after seeing a form of corporate name on a corporate Web page, will he or she be able to find any relevant items? If the form of corporate name on the Web doesn’t match the form of the corporate name used in the OPAC as either an authorized form or a cross reference, the answer is no. To more directly address the question and to determine the size of the problem, a random sample of AACR2 corporate headings that were entered into the LC NAF between 1998 and 2002 was searched on the Web via Google to find corporate Web pages.

### Method

In November 2003, the author extracted AACR2 corporate headings created between 1998 and 2002, excluding conference names (MARC21 bibliographic field 111) and geographic names (field 151), from LC NAF through the tape load of the University of Illinois at Urbana-Champaign Voyager system. These dates were chosen because corporate bodies in LC NAF between 1998 and 2002 were more likely to have corporate Web pages than corporate bodies whose names appeared prior to that time. The total number of corporate names in LC NAF is more than 100,000, which is a little more than 10 percent of total name headings, including all author, title, and subject access points created between 1998 and 2002.
After those corporate headings were extracted, a minimum random sample size of 385 corporate headings was required to get a confidence level of 95 percent and a confidence interval ±5 percent for this study. This means that, for example, if 70 percent of the form of authorized corporate headings in the LC NAF match the form of corporate headings on corporate body Web pages, the author is 95 percent certain that the actual percentage range of matched headings is between 65 and 75 percent. The author’s hypothesis was that the form of 70 percent of corporate headings entered between 1998 and 2002 in the LC NAF would match the form of corporate names on the Web, but 20 percent would not match and 10 percent would not have corporate Web pages.

The author identified 450 corporate names rather than 385 corporate names using a random number generator. A larger sample was chosen in case certain types of headings turned out to present problems.

Two categories of materials were excluded from this study. First, very small music groups were excluded from this study. Many headings for music groups were entered in the LC NAF in 1999, but few of these groups have corporate Web pages. There could be several reasons why music groups didn’t yield satisfactory results, including that the group might be made up of only two or three people. Perhaps they didn’t bother to create Web pages. It is also true that small music groups disband very easily. In addition, searching the music groups is likely to yield a lot of hits that advertise the related music products rather than the groups’ Web pages. Second, some foreign corporate names, including those from Thailand, Bangladesh, and Arabic-speaking countries, were not included in the random sample due to lack of access to appropriate linguistic expertise.

Some of the corporate names were not used for this study because the sample was successfully completed with the first 385 names that resulted from eliminating the special cases in order to maintain the integrity of the random sample.

Corporate names as established in the LC NAF were searched on the Web using the Google search engine by keying in the names of the corporate bodies exactly as they appeared in the 110 field, using at least one keyword search in both basic and advanced search features. The final sample searched in Google consisted of 385 corporate headings in English, European languages (including Slavic languages and Greek), Chinese, Japanese, and Korean.

**Results**

The results from searching corporate Web pages fall into three categories: matches, no Web pages, and no match. In the matches category, corporate names in the LC NAF match the form of corporate names on the corporate Web pages. Results listed as the no Web pages category indicates official corporate Web pages are not found even though corporate names may be mentioned on the Web. The no match category includes corporate names in the LC NAF that differ from the form of corporate names found on corporate Web pages. The findings are summarized in table 1.

**Matches**

Two hundred eighty one corporate names in the LC NAF matched corporate Web pages (73 percent). Users who have looked at the corporate home pages should be able to find materials by those corporate bodies in the OPAC. One example is Denmark Technical College. Its LC NAF record is given in figure 1, and its corporate Web page is shown in figure 2.

According to the NACO *Participants’ Manual*, published in 1996: “Give the name of the home page, gopher, etc. and the date if it was consulted. In subfield $b$, give a location, if appropriate, and the information found. Generally, don’t include the URL (Uniform Resource Locator) since the address often changes.” NACO rules allow catalogers to add references from corporate body Web pages, and it has become part of routine for many catalogers to check the Web for name authority work. This corporate authority record created in 2002 indicates a Web page for this corporate body. It tells us this cataloger checked the corporate Web page when creating this corporate authority record.

A second example in this category is provided for the Super 8 Motel (figures 3 and 4).

This random sample supports the author’s hypothesis that the form of 70 percent of corporate names in the LC NAF match the form of corporate names on corporate body Web pages. This shows that Lubetzky’s attempt to bring corporate authorship within the framework of personal authorship by entering the corporate body directly under the name of the body proved to be very successful, as more than 70 percent of the forms of corporate body match in the LC NAF and on corporate Web pages. This means if users search OPACs for corporate names found on corporate Web pages, they can find many (73 percent in this random sample) without any problems, assuming that materials from those corporate bodies are held in the collections.

**No Web Pages**

This category is for those corporate bodies that don’t have corporate body Web pages that can be easily located through a Google search. This category represents 9.4 percent (thirty-six corporate names) of this random sample. Reasons for having no Web pages include: (1) the body no
Table 1. Results of Searching LC NAF corporate names (n=385) against Google

<table>
<thead>
<tr>
<th>Type</th>
<th>No.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matches</td>
<td>281</td>
<td>73.0</td>
</tr>
<tr>
<td>No Web pages</td>
<td>36</td>
<td>9.4</td>
</tr>
<tr>
<td>No match</td>
<td>68</td>
<td>17.6</td>
</tr>
<tr>
<td>Total</td>
<td>385</td>
<td>100</td>
</tr>
</tbody>
</table>

010 n 2002073890
040 DLCSb engSc DLC
005 20020806102130.0
110 2 Denmark Technical College (Denmark, S.C.)
410 2 South Carolina.$b Denmark Technical College
670 NUCMC data from Benedict College for Davis, M.W. Marianna W. Davis papers, (197-)[198-]$b (Denmark Technical College)
670 Its WWW Home page, Aug. 6, 2002$b (Denmark Technical College is a comprehensive two-year college dedicated to achieving academic excellence; originally founded in 1947 as South Carolina Area Trade School; became affiliated with the State Board of South Carolina Technical and Comprehensive Education in 1969; Denmark, S.C.)

Figure 1. The LC NAF record for Denmark Technical College

010 n 2002010463
040 DLCSb engSc DLC
005 20020416120256.0
110 2 Super 8 Motel
410 2 Super Eight Motel

Figure 3. The LC NAF record for Super 8 Motel

Figure 2: Denmark Technical College Web site (www.denmarktech.edu/lrc.htm), viewed March 2005

longer exists; (2) the body is a small corporate body; and (3) the body is transitory. One example for each category will be listed below.

E. A. Leete and Son (figure 5) was used as a sample of a body that no longer exists. According to this authority record, E. A. Leete and Son (Guilford, Connecticut) existed in the late nineteenth and early twentieth century. The author didn’t expect to find a Web page for this company.

Palatka Urban Planning (figure 6) was used as an example of a small or local corporate body. According to the author’s experience, a lot of small corporate bodies do not bother to create Web pages. They use other ways to do marketing and outreach.

The American Hospital Association Centennial Planning Committee was used as a sample of a transitory body. The author didn’t find transitory bodies (in this random sample) to have Web pages even after checking
related Web pages. The LC NAF record (figure 7) supported the author’s hypothesis. However, some corporate authority headings seemed to fall into the no Web page category at first look, but corporate Web pages were found after further searching. One example, Palmerston North Teachers College, is shown in figure 8.

From the corporate body Web page, we understand that Palmerston North Teachers College was established in 1956. In 1990, the name changed to Palmerston North College of Education. In 1996, Massey University absorbed Palmerston North College of Education, now the Massey University of College Education. Massey University Faculty of Education and Palmerston North College of Education were merged to form Massey University College of Education in 1996. The corporate Web page has information that makes it very easy for users to understand the merger.

No Matches

Sixty-eight corporate names (17.6 percent) were found to have different forms of name in the LC NAF and on cor-

Figure 5. The LC NAF record for E. A. Leete and Son

Figure 6. The LC NAF record for Palatka Urban Planning

Figure 7. The LC NAF record for Palmerston North Teachers College

Figure 8. The LC NAF records for the Palmerston North Teachers College and the Massey University College of Education

porate Web pages. This includes checking variant forms used as cross references in the authority records. For example, the Acoustical Society of America record (figure 9) had the headings and subdivisions in reverse order (this was the case for thirty-three corporate names in this study). The society’s Web site is shown in figure 10.
A second example involves corporate bodies in Chinese, Japanese, and Korean (CJK) (there were thirty corporate names in this study). For the China Iron and Steel Association (figure 11), “Zhongguo gang tie gong ye xie hui” was keyed in Google (English version) and the corporate Web page was not found. However, when the Pinyin form of the corporate name was keyed in Google (Chinese version), the corporate Web page was found. According to AACR2: 24.1B1., “If the name of the body is in a language written in a nonroman script, romanize the name according to the table for that language adopted by the cataloging agency.” When users type the romanized Chinese in Google (English version), they will not be able to find anything by those corporate bodies. Corporate names in Chinese are searchable in Google (Chinese version) when users type Pinyin. If users don’t know Pinyin, or can’t identify Chinese characters in Pinyin, or don’t have Chinese software installed in their computers, they will not be able to find romanized Chinese corporate names on the Web. On the other hand, when users type Chinese characters in the OPAC, most of them will not be able to find anything at the present time. Catalogers are required to romanize Chinese corporate names according to AACR2 when creating Chinese corporate authority records at the present time due to computer-related issues. More discussion of nonroman script headings is needed but is beyond the scope of this paper. Questions that should be addressed include: What problems are specific to corporate headings? What problems are inherent in searching for materials in CJK languages? What problems are created by the decision made for LC romanization? What problems are created by the application of AACR2 rules to non-Indo-European languages?

This is an example of the direct form of name versus the name of a parent body followed by a subdivision. The rules in AACR2 24.13 instruct catalogers to form the authorized headings this way. But users will not retrieve anything in the OPAC if they type in the name as it appears on its corporate Web page—“Technical Committee on Architectural Acoustics (TCAA) of the Acoustical Society of America (ASA).” Only if they use a keyword search will it be found. Since users might be looking for materials by the form “Technical Committee on Architectural Acoustics (TCAA) of the Acoustical Society of America (ASA)” on the Web in the OPAC, catalogers need to add that form as a reference (410). Catalogers should also add the form of the corporate body on the title page “Technical Committee on Architectural Acoustics of the Acoustical Society of America” as another reference (410) so that users will be able to find the materials by and about that corporate body no matter what type of systems they use. In the meantime, libraries should negotiate with vendors, especially those libraries ready to buy new systems, to add more features. These features should include better displays, suggestions to users, and the ability to revert to keyword search if a direct search fails. Most vendors will not take the initiative to change or add any features after a system is purchased.

This example shows that SPRING Singapore is a completely new form of a corporate name on its corporate Web page. From the corporate authority record created by LC in 1998, we know that the old form of corporate name “Singapore Productivity and Standards Boards” is on the title page of its publication issued by the body in its own language. We know from the note (670 field) in the authority record that National Productivity Board and the Singapore Institute of Standards and Industrial Research merged to become Singapore Productivity and Standards Boards. But we don’t know when the name changed to SPRING Singapore or whether the body has issued publications since the change. If users search the OPAC...
using the form of corporate name on the Web, SPRING Singapore, they will not be able to find the resources by and about that corporate body. At the present time, NACO catalogers are only required to look in other sources (beyond the item-in-hand and the database in which one is cataloging) for variants, fuller forms of the heading or dates, and so on when the heading conflicts with another in the LC NAF and the item-in-hand does not provide enough information to break the conflict. In other words, if the form on the title page matches the form in the LC NAF, catalogers are not required to check corporate Web pages to see if there is a different form of the corporate name. (www.loc.gov/catdir/pcc/naco/670faq.html) In this case, catalogers are not instructed to create an authorized heading under the new name on its corporate Web page, SPRING Singapore. In the case of a name change such as this example, the author suggests that catalogers add the new form of corporate name from the corporate Web page in the 510 field of the authority record and create a new authority record for the new corporate name with a 510 reference back to the older form so that users can find the resources by and about that corporate body no matter which form of corporate name they type in the OPAC. In the future, we might find SPRING Singapore on the title pages of its publications, but in the meantime, users are not able to find materials by the form of corporate name on its corporate body Web page in the OPAC—assuming, of course, that libraries have cataloged resources by this body.

### Conclusions

During the Paris Conference in 1961, representatives from around the world agreed on the choice of the corporate author in the original language in the official form as it appeared most frequently on the title pages of publica-
tions. AACRI and AACR2 followed the Paris Principles on the standard form of heading for corporate bodies because the name of the author and the title of the book were the most important clues by which the book was identified when cited and by which it was looked for in the catalog or called for in the library.22 Catalogers have been constructing corporate authority records following the same rules for many years. The Internet has changed the way users access information for their research needs. Today we have a generation of library users who search for information for their research needs in Web search engines before they search the OPAC. There are now other equally important clues besides the title pages for print publications—corporate Web pages.

This random sample of corporate names from the LC NAF created between 1998 and 2002 indicates that 73 percent (±5 percent) of the form of corporate names in the LC NAF match the form of corporate names on the Web, as hypothesized by the author. Users will have no problems finding resources by and about those corporate bodies in the OPAC if they use the form of corporate names on the Web. The study also shows that 9.4 percent (±5 percent) of corporate bodies don’t have corporate Web pages. The most serious problem is that 17.6 percent (±5 percent) of percent corporate bodies in this random sample have different forms of names in the LC NAF than on their corporate Web pages. This percentage could be even higher, more than 20 percent if the foreign corporate names from Thailand, Bangladesh, and Arabic-speaking countries are included. This means users have a one-in-five chance of not finding the resources by and about corporate bodies in the OPAC if they use the form of corporate names on the Web. These corporate authority records would be enhanced if data from the corporate Web sites were used.

Since the highest principle for the construction of cataloging codes should be for the convenience of the catalog users, according to the Statement of International Cataloguing Principles, it is important that catalogers check corporate Web pages when creating corporate authority records.23 According to several recent surveys, users search the Web for their research needs before they search the OPAC. If catalogers continue the current way of constructing corporate authority records, users may not be able to find more than 20 percent of the resources by and about corporate bodies in the OPAC.

According to the author’s experience, it usually takes a couple of minutes to find most corporate Web pages, and it only takes a few more minutes to add a cross-reference (410 field) and a note (670 field) in an authority record. Most corporate Web pages are stable and are there for both users and catalogers to find, unlike ten years ago, when the Internet was still in flux. According to the NACO Participants’ Manual, catalogers are already checking corporate Web pages to add supplementary sources to be cited in a 670 note.24 Searching corporate Web pages can save catalogers from having to check print reference sources to determine the authorized corporate headings. This is especially true at the present time as many libraries do not have the print sources needed due to budget constraints.

It is hoped that for corporate names in CJK, Unicode will be implemented soon in most libraries. Unicode provides a unique number for every character, no matter what the platform, program, or language, so that users can find the resources by and about those corporate bodies by typing characters in the OPAC.

When the corporate Web page indicates a name change, catalogers do need to spend more time to create a new corporate authority record and to link the old and new authorized names. It is very important that catalogers do this work to help users find the resources by and about those 20 percent of corporate bodies in the OPAC.

However, catalogers need to do more research on how many users are looking for corporate bodies in the OPAC and not finding them. Catalogers need to know under what specific situations (acronym/full form, earlier/later form) or types of corporate bodies (government agencies, educational institution names, nongovernmental organizations) most often include data not captured by publications by the body. They need to discuss the cost and benefit of using the form of corporate names on the Web both as cross references and authorized headings. They also need to investigate the stability of the corporate Web pages.

## Acknowledgments

The author would like to thank Robert Burger for his advice and insightful comments on this paper. The author would also like to thank Tina Chrzastowski for her encouragement and helpful suggestions. The author wishes to acknowledge the Research and Publication Committee of the University of Illinois Library for its financial support and Ying Zhang for her assistance of this project.

## References


5. Ibid.
12. Ibid.
13. Ibid.
15. Ibid.
17. Ibid., 2-3.
20. The Joint Steering Committee for Revision of AACR, Anglo-American Cataloging Rules, 24-5.
22. Svenonius and McGarry, Seymour Lubetzky.

Index to Advertisers

Art & Tech
EBSCO

67
cover 2

LITA
InfoUSA

cover 3
cover 4