# Library Resources & Technical Services

ISSN 0024-2527 October 2001 Volume 45, No. 4

## ARTICLES

**Form Subdivisions** 187
Their Identification and Use in LCSH
*Edward T. O'Neill, Lois Mai Chan, Eric Childress, Rebecca Dean, Lynn M. El-Hoshy, and Diane Vizine-Goetz*

**Computer Cataloging of Electronic Journals in Unstable Aggregator Databases** 198
The Hong Kong Baptist University Library Experience
*Yiu-On Li and Shirley W. Leung*

## FEATURE

**Book Reviews** 212
*Margaret Rohdy, Editor*
The Intellectual Foundation of Information Organization, *reviewed by Michael Gorman*
Managing Electronic Serials: Essays Based on the ALCTS Electronic Serials Institutes, 1997–1999; Developing and Managing Electronic Journal Collections; Selecting and Managing Electronic Resources; Successes and Failures of Digital Libraries; Collection Development in a Digital Environment; Electronic Collection Management; Collection Development in an Electronic Environment; Assessing Digital Library Services, *reviewed by Sheila S. Intner*

**Instructions for Authors** 220

**Index to Volume 45** 222
*Edward Swanson*

American Library Association

Association for Library Collections & Technical Services

Visit *LRTS* online at www.ala.org/alcts/lrts.

For current news and reports on ALCTS activities, see the ALCTS Newsletter Online at www.ala.org/alcts/alcts_news.

## EDITORIAL BOARD

### Editor and Chair of the Editorial Board

John Budd
*University of Missouri*

### Assistant Editors

Elizabeth E. Cramer, *Appalachian State University*

Prudence W. Dalrymple, *Dominican University*

Jeanne M. Drewes, *Michigan State University*

Martin M. Kurth, *Cornell University*

Margaret A. Rohdy, *University of Pennsylvania*

Jane Treadwell, *Emory University*

Barbara A. Winters, *University of Georgia*

### Ex-Officio Members

Elaine Yontz, *University of Southern Florida*
Delegate of the Chair, Council of Regional Groups

Charles Wilt
Executive Director, ALCTS

Miriam Palm, *Stanford University, emerita*
Editor, ALCTS Newsletter Online



*Library Resources & Technical Services* (ISSN 0024-2527) is published quarterly by the American Library Association, 50 E. Huron St., Chicago, IL 60611. It is the official publication of the Association for Library Collections & Technical Services, a division of the American Library Association. Subscription Price: to members of the Association for Library Collections & Technical Services, $27.50 per year, included in the membership dues; to nonmembers, $55 per year in U.S., Canada, and Mexico, and $65 per year in other foreign countries. Single copies, $15. Periodical postage paid at Chicago, IL, and at additional mailing offices. **POSTMASTER:** Send address changes to *Library Resources & Technical Services,* 50 E. Huron St., Chicago, IL 60611. Business Manager: Charles Wilt, Executive Director, Association for Library Collections & Technical Services, a division of the American Library Association. Send manuscripts to the Editorial Office: John Budd, Editor, *Library Resources & Technical Services,* University of Missouri, 221M Townsend Hall, Columbia, MO 65211; (573) 882-3258; fax: (573) 884-4944; e-mail: buddj@missouri.edu. ALA Production Services: Troy D. Linker, Kevin Heubusch; Ellie Barta-Moran, Angela Hanshaw, Kristen McKulski, and Karen Sheets. Members: Address changes and inquiries should be sent to Membership Department—*Library Resources & Technical Services,* 50 E. Huron St., Chicago, IL 60611. Nonmember subscribers: Subscriptions, orders, changes of address, and inquiries should be sent to *Library Resources & Technical Services,* Subscription Department, American Library Association, 50 E. Huron St., Chicago, IL 60611, 800-545-2433; fax: (312) 944-2641; subscription@ala.org.

*Library Resources & Technical Services* is indexed in *Library Literature, Library & Information Science Abstracts, Current Index to Journals in Education, Science Citation Index,* and *Information Science Abstracts.* Contents are listed in CALL (Current American—Library Literature). Its reviews are included in *Book Review Digest, Book Review Index,* and *Review of Reviews.*

Instructions for authors appear on p. 220–221 and on the LRTS Web page at www.ala.org/alcts/lrts. Copies of books for review should be addressed to Margaret Rohdy, Book Review Editor, *Library Resources & Technical Services,* Van Pelt-Dietrich Library Center, University of Pennsylvania Libraries, 3420 Walnut St., Philadelphia, PA 19104-6206 (e-mail: rohdy@pobox.upenn.edu).



© American Library Association 2001

All materials in this journal subject to copyright by the American Library Association may be photocopied for the noncommercial purpose of scientific or educational advancement granted by Sections 107 and 108 of the Copyright Revision Act of 1976. For other reprinting, photocopying, or translating, address requests to the ALA Office of Rights and Permissions, 50 E. Huron St., Chicago, IL 60611.



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1992. ∞



Publication in *Library Resources & Technical Services* does not imply official endorsement by the Association for Library Collections & Technical Services nor by ALA, and the assumption of editorial responsibility is not to be construed as endorsement of the opinions expressed by the editor or individual contributors.

# Form Subdivisions

## Their Identification and Use in LCSH

**Edward T. O'Neill, Lois Mai Chan, Eric Childress, Rebecca Dean, Lynn M. El-Hoshy, and Diane Vizine-Goetz**

*Form subdivisions have always been an important part of the Library of Congress Subject Headings. However, when the MARC format was developed, no separate subfield code to identify form subdivisions was defined. Form and topical subdivisions were both included within a general subdivision category. In 1995, the USMARC Advisory Group approved a proposal defining subfield $v for form subdivisions, and in 1999 the Library of Congress (LC) began identifying form subdivisions with the new code.*

*However, there are millions of older bibliographic records lacking the explicit form subdivision coding. Identifying form subdivisions retrospectively is not a simple task. An algorithmic method was developed to identify form subdivisions coded as general subdivisions. The algorithm was used to identify 2,563 unique form subdivisions or combinations of form subdivisions in OCLC's WorldCat. The algorithm proved to be highly accurate with an error rate estimated to be less than 0.1%. The observed usage of the form subdivisions was highly skewed with the 100 most used form subdivisions or combinations of subdivisions accounting for 90% of the assignments.*

Recent efforts to distinguish between topical and form data are moving Library of Congress Subject Headings (LCSH) closer to a truly faceted subject vocabulary. While form data in LCSH are represented in both form headings and form subdivisions, under the current LC application rules, form data appear in most cases as subdivisions under topical or name headings.

In implementing the $v subfield code for form subdivision in the MARC 21 (formerly USMARC) format, a number of issues have come to the fore:

- distinction between form and topical subdivisions
- combinations of two or more form subdivisions in the same heading string

In this article, a method is developed to algorithmically identify form subdivisions lacking explicit form subfield coding.

### Explicit Coding for Form Subdivisions

Form subdivisions have been a part of LCSH since its inception. Beginning in 1906, the Library of Congress issued auxiliary lists of subdivisions that included a section of "General form divisions under subjects." Guidelines on the use of subdivisions, such as those published in the introduction to the eighth edition of

**Edward T. O'Neill** (oneill@oclc.org) is Research Scientist, Office of Research, Online Computer Library Center (OCLC), Dublin, Ohio.

**Lois Mai Chan** (loischan@pop.uky.edu) is Professor, School of Library and Information Science, University of Kentucky, Lexington.

**Eric Childress** (eric_childress@oclc.org) is Consulting Product Support Specialist, Metadata Services Division, OCLC.

**Rebecca Dean** (rebecca_dean@oclc.org) is Manager, Metadata Analysis and Investigation Section, OCLC.

**Lynn M. El-Hoshy** (lelh@loc.gov) is Senior Cataloging Policy Specialist, Cataloging Policy and Support Office, Library of Congress, Washington, D.C.

**Diane Vizine-Goetz** (vizine@oclc.org) is Research Scientist, Office of Research, OCLC.

Manuscript received March 8, 2001; accepted for publication June 25, 2001.

*Library of Congress Subject Headings* (Library of Congress 1975), instructed catalogers to use individual subdivisions either "as a topical subdivision," "as a form subdivision," or "as a form or topical subdivision" under specified types of headings for particular types of materials. Yet when the MARC format for encoding and communicating bibliographic data was developed in the late 1960s, a separate subfield code to identify form subdivisions in subject heading strings was not defined. Form subdivisions were included along with topical subdivisions in a general subdivision category to be coded as $x.

In 1991, a conference was convened at Airlie, Va., to consider the role of subdivisions in LCSH. One of the conference's six recommendations was: "The question of whether subdivisions should be coded specifically to improve online displays for end users should be considered . . . In particular, the Library of Congress should investigate implementing a separate subfield code for form subdivisions" (O'Hara Conway 1992). In response, the Library of Congress requested that the ALA Association for Library Collections and Technical Services (ALCTS) Cataloging and Classification Section (CCS) Subject Analysis Committee (SAC) investigate form subdivision coding. Hemmasi, Miller, and Lasater (1999) report on the issues that SAC identified and studied, including "retrospective conversion, varying cataloging practices and user needs across disciplines, no distinct list of form headings, cataloger training, and the redundancy of content in USMARC record elements" (unnumbered). In 1993, SAC recommended that a separate subfield code for form subdivisions be implemented. Subsequently, two discussion papers defining a new subfield code and posing questions on retrospective conversion, the use of a form subdivision subfield by online systems, authority control, implementation options, and general user opinions were considered by the USMARC Advisory Group before it approved a proposal to define subfield $v for form subdivisions in 1995. The proponents argued that a separate subfield code would make it possible to retrieve form data more predictably, improve online displays for users, and separate LCSH elements into their facets of topic, place, chronology, and form.

## Guidelines for Assignment

In applying form subdivisions, the question is: Where does the cataloger look for guidance? There are several sources and methods of information:

- *Subject Cataloging Manual: Subject Headings (SCM)* (Library of Congress 1996)
- *Free-Floating Subdivisions: An Alphabetical Index (FFS)* (Library of Congress 2000)
- Patterns discerned in assigned heading strings in LC MARC records
- Subdivision authority records
- The test of what the work "is" versus what the work is "about" to determine the appropriate category of subdivision
- The "reading backwards" or "from right to left" test to determine the proper order of subdivisions within the string

When a question arises, the first place for a cataloger to look for an answer is *Subject Cataloging Manual: Subject Headings.* The manual gives numerous instructions and examples on the application of many of the subdivisions, although they are scattered throughout the publication. The publication *Free-Floating Subdivisions: An Alphabetical Index* provides a quick reference for precombined subdivisions. Nevertheless, there are still situations not fully covered; many multiple free-floating subdivisions that appear in LC MARC authority records are not shown in *SCM* or *FFS.* For these, one must rely on other means. One possible approach is to examine patterns in assigned heading strings in LC MARC bibliographic records, which can serve as examples but hardly provide definite answers. The test that a form subdivision "represents what the book is, rather than what it is about" (Haykin 1951) may also be used to help in the distinction between form and topic. Finally, another test that has been suggested is to read the heading string backwards, i.e., from right to left, to see if the string fits the context of the item being cataloged. For example:

**Art—Bibliography—Periodicals**
(a serially issued art bibliography)
**Art—Periodicals—Bibliography**
(a bibliography of journals on art)

## Distinction between Form and Topical Subdivisions

Virtually all efforts to revise or improve LCSH, including the Airlie Conference (O'Hara Conway 1992), ALCTS/SAC/Subcommittee on Metadata and Subject Analysis (*Subject data in the metadata record 1999*), and OCLC's FAST (Faceted Application of Subject Terminology) project (Chan et al. 2001), consider form subdivisions as a distinct type and treat form subdivisions differently from general ($x) subdivisions. All of these efforts assume that form subdivisions can be identified. However, until recently, the Library of Congress coded form subdivisions the same as general subdivisions ($x). Only in 1999 did the Library of Congress begin explicitly identifying forms with the $v subfield code.

In coding form subdivisions, the first issue to be resolved is how to determine whether a particular subdivision in a subject string represents a topic or form. Although many terms clearly belong to one or the other category, many others are ambiguous. While subdivisions such as **—Education** or **—Quality control** can only be considered topical, others are not so obvious. For example, subdivisions such as **—Texts** and **—Translations into French [German, etc.]** may be used as either a topical or form subdivision, depending on the context. Even subdivisions such as **—Periodicals** are sometimes used as topical subdivisions. For example, in the heading:

**Academic achievement $xPeriodicals $vIndexes**
(an index to a journal on academic achievement)

the subdivision **—Periodicals** is topical; but in the heading:

**Universities and colleges $xFinance $vPeriodicals**
(a journal on higher education finance)

it is a form since it is assigned to represent a publication issued in serial form.

Currently, the subfield code for each free-floating subdivision is shown in *SCM* and *FFS*. The Library of Congress is in the process of creating authority records for free-floating subdivisions with specific information regarding subfield codes. When completed, the specific instruction will contribute greatly to consistency in application.

### Combinations of Two or More Form Subdivisions

The use of two or more subdivisions involving form data within the same heading raises at least three problems:

- When can a form subdivision be further subdivided by another form, geographic, or topical subdivision?
- In what order should the subdivisions appear?
- How does one code each subdivision, that is, how does one choose between $v, $x, and $z?

To answer the first question, *SCM* and *FFS* list many precombined multiple form subdivisions as an aid to catalogers. Examples include:

**—Biography—Dictionaries** (v-v)
**—Biography—Sermons** (v-v)
**—Maps—Facsimiles** (v-v)

In many cases, a form subdivision may be further subdivided by a topical subdivision.

**—Concordances, English—Authorized, [Living Bible, Revised Standard, etc.]** (v-x)
**—Dictionaries—Polyglot** (v-x)

In limited cases, a form subdivision may also be further subdivided by a geographic subdivision as in:

**School buildings $vSpecifications $zIowa**

However, it is not practical to list all possible combinations in *SCM* or *FFS*, and many such combinations not enumerated in these publications have been assigned to bibliographic records. For example:

**—Biography—Sources** (v-v)
**—Catalogs—Periodicals** (v-v)
**—Indexes—Periodicals** (v-v)
**—Observations—Periodicals** (v-v)
**—Statistics—Periodicals** (v-v)

Again, in each case, the cataloger is called upon to exercise judgment.

There are situations where LC instructions specifically prohibit certain combinations of form subdivisions. For example, **—Abstracts** should not be used after **—Congresses** (cf. *SCM* H1460). H1927 in *SCM* contains a list of form subdivisions that cannot be further subdivided by the subdivision **—Periodicals**. It is important that the cataloger be aware of the prohibition when using these subdivisions.

The second question relating to the use of two or more form subdivisions is: In what order should the individual form subdivisions appear within the string? The first place to seek guidance is in *SCM* or *FFS*. The lists of free-floating subdivisions enumerate many precombined subdivisions, for example, **—Bibliography—Catalogs**.

For combinations not listed in *SCM* or *FFS*, other methods must be employed. In most subject headings, the form subdivision appears as the last element, following the general pattern of subdivision order, **Topic—Topic—Place—Time—Form.** However, there are exceptions such as: **—Conversation and phrase books—Polyglot** (v-x).

When the desired combination is not enumerated, the cataloger must exercise judgment based on the context. One suggestion made earlier is to "read backwards," or from right to left, to see if the string fits the context of the document.

**—Periodicals—Indexes**
(for an index to periodicals)
**—Indexes—Periodicals**
(for a serially issued index)

For guidance on the third question, how to code subdivisions in each case, the Library of Congress has provided a most valuable service in indicating subfield coding for each free-floating subdivision in recent updates of *SCM* and *FFS*. Newly created authority records of subdivisions also indicate the appropriate coding information. Nevertheless, lists in these publications are not exhaustive. For example, while **—Biography—Anecdotes** (v-v) and **—Biography —Dictionaries** (v-v) are enumerated, the combination **—Biography—Bibliography** is not, even though it has been used in bibliographic records. The difficulty lies in the fact that one cannot assume that in all cases, when two or more form subdivisions appear under the same heading, the coding is always v-v. When an apparent form subdivision is followed by another form subdivision or another topical subdivision, the subfield code can change. For example,

**—Bibliography** (v)
**—Bibliography—Exhibitions** (v-v)
**—Bibliography—Methodology** (x-x)

**—Hymns** (v)
**—Hymns—History and criticism** (x-x)
**—Hymns—Texts** (v-v)

**—Maps—Early works to 1800** (v-v)
**—Maps—Facsimiles** (v-v)
**—Maps—Symbols** (x-x)

The specific guidance given in *SCM* and *FFS* is of enormous help, but what if one combines **—Abstracts** with **—Periodicals,** a combination not listed in *FFS*?

The advice often given for distinguishing between form and topical subdivisions is to ask whether the subdivision in question represents what the document "is" or what it "is about." This test can usually resolve the question of content versus form.

In certain cases, a trailing form subdivision may affect the coding of the preceding form subdivision, for example:

**—Maps** (v)
(Map(s) of . . . )
**—Maps—Bibliography** (x-v)
(list(s) of maps of . . . )
**—Periodicals** (v)
(serial(s) or periodical(s) on . . . )
**—Periodicals—Abbreviations of titles** (x-v)
(abbreviations of titles of serials or periodicals on . . . )
**—Periodicals—Bibliography** (x-v)
(list(s) of serials or periodicals on . . . )
**—Periodicals—Bibliography—Catalogs** (x-v-v)
(list(s) of serials or periodicals held by one organization or library)
**—Periodicals—Bibliography—Union lists** (x-v-v)
(catalog(s) of serials or periodicals on those subjects held by two or more libraries)

In some cases, a subject heading may include two or more form subdivisions, which further compound the problem in order and in coding, for example:

**Alcoholism $xPrevention $xPeriodicals $vAbstracts $vDatabases**
**Jews $zPoland $zRadom (Voivodeship) $xHistory $xSources $vBibliography $vCatalogs**

### The Subdivision —History

The application of the subdivision **—History** is particularly problematic. Currently, it is coded as a topical ($x) subdivision in *SCM* and *FFS*. In effect, when it appears in a subject heading string, it usually represents what the document "is" rather than what it is "about." For example, the heading **Education—History** is assigned to a work that "is" a history of education, not a work "about" the history of education. The problem is compounded when the subdivision **—History** is combined with another form subdivision. For example:

**Science $xHistory $vPeriodicals**
(a serial or periodical on scientific history)
**Science $xPeriodicals $xHistory**
(a history of scientific serials or periodicals)

Here, the method of judging by what it "is" versus what it is "about" fails to work.

A similar subdivision is **—History and criticism,** which is also coded as a general ($x) subdivision. The heading **Literature—History and criticism** is normally assigned to a history of literature rather than a work about literary history. The use of **—History** and **—History and criticism** also results in combinations such as:

**—Biography** (v)
(biography of . . .)
**—Biography—History and criticism** (x-x)
(a history or criticism of biography of . . .)
**—Music** (v)
(music of an ethnic group)
**—Music—History and criticism** (x-x)
(a history or criticism of the music of an ethnic group)

## Algorithmic Identification of Form Subdivisions

Identifying and coding form subdivisions is not a simple task. OCLC's WorldCat contains more than eight million unique Library of Congress topical and geographic subject headings—less than 4% contain explicitly coded form subdivisions. The other headings either do not contain any forms or have forms coded as general subdivisions. Identifying forms is difficult due to the complexity of forms structure and the fact that many subdivisions can be either topical (general) or form depending on the context of the heading.

The sheer number of headings demands that an automated procedure be developed to identify and recode form subdivisions. For this purpose, research staff at OCLC developed an algorithmic method based on a table-driven procedure. After extended review and analysis, the approach adopted in this project for identification is first to deal with the special forms, that is, form subdivisions with special or unique application rules, and then to use a table-driven procedure to identify the remaining forms.

### Step One: Identifying Special Forms

The following subdivisions are governed by special rules when they are used as the last subdivision in a heading string: **—Periodicals, —Juvenile, —Juvenile literature, —Juvenile films, —Juvenile sound recordings, —Databases, —Early works to 1800,** and **—Facsimiles.** Any of these forms can be removed from the heading and the remainder of the heading can be treated as if these forms were never part of the heading. For the purpose of identifying form subdivisions, the heading:

**Land value taxation $zIreland $xTables $xEarly works to 1800**

can be reduced to:

**Land value taxation $zIreland $xTables.**

After removing **—Early works to 1800,** any remaining forms in the heading can be identified using the table-driven procedure.

There are some additional restrictions on removing these forms. The restrictions on what can precede **—Periodicals** are specified in *SCM* (H1927). To prevent invalid combinations of form subdivisions from being identified, if any of the subdivisions specified in H1927 or the subdivisions **—Exhibitions** or **—Newspapers** immediately precedes **—Periodicals**, the subdivision is not removed from the heading. The "Juvenile" forms are restricted to headings not otherwise identified as juvenile. These are not removed when they begin with the word "Juvenile" or "Children's."

In headings involving this group of forms, the last subdivision in the string would be recoded as $v, and the rest of the heading would be analyzed with the last subdivision removed from the heading. For example, the heading **Cities and towns $zUnited States $xMaps $xDatabases** (before recoding) would be treated as **Cities and towns $zUnited States $xMaps** in the remainder of the analysis. The following are some examples where the last (underlined) general subdivision would be removed:

**Medical care $zArab countries $xEarly works to 1800**
**Photography $xCatalogs $xPeriodicals**
**Fuelwood consumption $zPrince Edward Island $xStatistics $xPeriodicals**
**Lesbian teenagers $zUnited States $xCase studies $xJuvenile literature**

However, the following would not be removed since they are exceptions to the general rule:

**African Americans $zNew York (State) $xGenealogy $xPeriodicals**
**Christmas $xJuvenile fiction $xJuvenile sound recordings**
**Art, German $zGermany (East) $xExhibitions $xPeriodicals**

Note that the remaining general ($x) subdivisions are not necessarily correct—only that they are not valid form subdivisions. Regardless of whether or not any forms are removed, the headings continue to be analyzed.

The forms **—Bibliography, —Congresses,** and **—Indexes** are also given special treatment. Any heading that ends with either of these subdivisions is recoded with **—Bibliography, —Congresses,** or **—Indexes** as $v, but none of the other subdivisions will be considered to be forms. The following headings are shown with the revised subfield codes (assuming the $v were originally coded as $x):

**Scottish poetry $y20th century $vBibliography**
**Nahuas $xPeriodicals $vIndexes**
**Urbanization $zNigeria $xStatistics $vCongresses**

There are four form subdivisions that can be geographically subdivided: **—Catalogs and collections, —Job descriptions, —Specifications,** and **—Registers of dead.** The following are examples of recoded form subdivisions followed by geographic subdivisions:

**Medicinal plants $vCatalogs and collections $zThailand $zSala Ya**
**Sewage disposal plants $vSpecifications $zTexas $zEl Paso**

The subdivision **—Readers,** which is used for reading texts, is another special case that can be followed by any topic. **—Readers** should be recoded as a $v but the following topic is retained as $x. Some examples of recoding **—Readers** are:

**Spanish language $vReaders $xCivilization**
**German language $vReaders $xScience**
**Russian language $vReaders $xSoviet Union**

Note that even under subdivision **—Readers** when the *topic* is a geographic name such as the **Soviet Union,** the subdivision containing the geographic name is coded as a topical ($x) subdivision rather than a geographic ($z) subdivision, because in this case the place name represents a topic rather than a location.

### Step Two: Identifying Forms

A table of form subdivisions was created by supplementing the list of forms identified in *Free-Floating Subdivisions* with other forms identified through various sources. All headings not subject to the special treatment described above are then checked to determine if they have terminating subdivision(s) matching those in the augmented list. This expanded list contains form patterns and their preferred subdivision coding. Included in the list are 639 entries containing from one to three subdivisions.

Since there can be no more than three form subdivisions after removing the special forms, the list is searched in three steps. The first search is for the last three (if they exist) general subdivisions. If it matches an entry in the list, the heading is recoded using the preferred coding. If no match is found, the last two general subdivisions are searched. If still no match is found, a final search is made for the last subdivision. For example, in the heading:

**American literature $xAfrican American authors $xHistory and criticism $xTheory, etc.**

the last three general subdivisions, **—African American authors—History and criticism—Theory, etc.,** would be checked against the table. When no match was found, the last two subdivisions, **—History and criticism—Theory, etc.,** would be checked. If, again, no match was found, the final subdivision, **—Theory, etc.,** would be checked. If all matches failed, the conclusion would be that all of the subdivisions were topical and that the original coding was assumed to be correct.

The following form subdivisions from the list serve as patterns for other national, ethnic, or language terms:

**—Concordances, *English***
**—Films for *English* speakers**
**—Harmonies, *English***
**—Interlinear translations, *English***
**—Liturgical lessons, *English***
**—Parallel versions, *English***
**—Paraphrases, *English***
**—Personal narratives, *English***
**—Sound recordings for *English* speakers**
**—Textbooks for *English* speakers**
**—Translations into *English***
**—Video recordings for *English* speakers**
**—Catechisms—*English***
**—Conversation and phrase books—*English***
**—Dictionaries, Juvenile—*English***
**—Dictionaries—*English***
**—Prayer-books and devotions—*English***
**—Textbooks for foreign speakers—*English***
**—Bio-bibliography—Dictionaries—*English***
**—Biography—Dictionaries—*English***

In these combinations, *English* serves as the pattern and can be replaced by any national, ethnic, or language terms.

## Evaluation of the Identification Algorithm

For the algorithm to be usable, it had to be highly reliable. With complex processes of this type, developing an error-free process is an unrealistic goal. Neither manual recoding by skilled professionals nor machine algorithms can be expected to produce perfect results. Even highly skilled professionals make mistakes—typically errors of oversight. Such an error is illustrated in the following heading:

**English language $vDictionaries $vChinese**

The form subdivision **—Dictionaries** can be subdivided by language, but the language subdivision **—Chinese** should be coded as a general subdivision ($x). This type of error is relatively common in spite of the fact that most professionals understand that language should be coded as a general subdivision. It is not that the cataloger didn't know how to code it but rather that it was overlooked. The only way errors of this type can be eliminated, or at least dramatically reduced, is to have at least two people recode each heading and to recheck each heading where the cod-

ing differs. The use of multiple coders, however, is very expensive, and would be difficult to justify in a production environment.

By contrast, algorithms produce very consistent results: they do not overlook anything. Algorithms, however, have very limited ability to understand the context. For example, the heading

**Executives $vQuotations**

is valid since **—Quotations** is authorized in the pattern heading for *Classes of Persons*. Based on general knowledge of language, most catalogers understand that executives are a class of people and, therefore, the pattern heading is appropriate. Algorithmic procedures have a much more difficult time with this type of contextual information. Unless the algorithm has been explicitly *told* or has previously *learned* that executives are a class of people, it has no way to validate this heading. As a result, the type of errors resulting from the algorithmic coding tend to be different from those made by people.

While recognizing that comparing manual and algorithmic error rates is a little like comparing apples to oranges, it nevertheless seems to be the best approach to evaluating the algorithm. It was assumed that algorithmic error rates that were as good or better than those observed in manual assignment would be acceptable. A methodology to estimate the algorithm's accuracy was required. Fortunately, since the Library of Congress currently is explicitly coding form subdivisions, there are a large number of records in WorldCat with the form subdivisions explicitly identified with the $v subfield code. For testing, all topical (650) and geographic (651) subject headings with explicit form coding were extracted to create a test file. Presumably, all of these records follow current coding practice.

To test the algorithm, all $v subfield codes in the heading were replaced with $x codes and then the heading was algorithmically recoded to explicitly identify the form subdivisions. For example, the heading **Agriculture $vIndexes** was changed to **Agriculture $xIndexes** in the test file. That heading then was algorithmically recoded as **Agriculture $vIndexes.** The resulting heading was then compared to the original to identify any headings for which the algorithmic form coding was different from the original heading. Presumably, all of the recoded headings that matched the original were correct. All headings that did not match the original were manually reviewed. As an example, headings pairs from that list are shown below:

**Artists $zGermany $vInterviews $vBibliography**

**Artists $zGermany $xInterviews $vBibliography**

**France $xPolitics and government $y1789- $vHistoriography**
**France $xPolitics and government $y1789- $xHistoriography**

The first heading of each pair is the original heading as it appeared in the MARC record. The second heading is the same heading after being algorithmically recoded. Each of these heading pairs was reviewed by at least two of the authors to determine the correct coding. When the initial reviewers did not agree on the coding, the headings were reviewed by all of the authors to ensure that the results were as accurate as possible.

During the review, it was found that some headings contained errors that could not be corrected by changing the subfield coding only. For example, in the heading

**Onondaga Indians $vProtraits**

the subdivision **—Portraits** is misspelled. As a result, the heading can only be corrected by changing the text of the subdivision. All headings with errors that could not be fully corrected by changing the subfield coding were removed from the test file.

The resulting test file contained 20,970 headings: 17,208 topical and 3,762 geographic. Of these headings, 662 contained manual coding errors resulting in a manual error rate of 3.15%. The coding of **—Dictionaries—English** as v-v rather than v-x was typical of the manual miscoding observed. The algorithm miscoded 15 headings, resulting in an algorithmic error rate of 0.07%, significantly better than in the case of manually coded records.

Caution is required in interpreting these results. First, the subject headings used in the test were assigned or recoded in early 1999 and, therefore, include the first attempts to explicitly code form subdivisions. As the catalogers gain experience in assigning subfield code $v, the accuracy of the coding can be expected to improve significantly. Second, the test headings were created in a *production* environment at the Library of Congress. In such an environment, accuracy must be balanced with productivity.

However, even recognizing that the current manual error rate is likely to be significantly less than the 3.15% observed, it appears to be impossible to achieve in a production environment a manual error rate as low as the algorithmic rate. To obtain manual error rates less than 1% would probably require at least two people independently assigning the subfield codes. Certainly, when compared to the manual error rate, an algorithmic error rate of less than 0.1% appears to be very acceptable.

## Usage Patterns

A large number of valid forms were identified but many were rarely assigned; 2,412 unique form subdivisions or combinations of form subdivisions were identified in topical and geographic headings from WorldCat. The geographically subdivisible forms **—Catalogs and collections, —Job descriptions, —Registers of dead,** and **—Specifications** were considered without their geographic subdivisions. For example, the combination **—Catalogs and collections —Japan** was treated simply as **—Catalogs and collections.** Collectively, these four forms would have resulted in 1,674 additional unique form subdivisions if they had been included with their geographic subdivisions. General ($x) subdivisions were included so that the forms identified consisted of a combination of one to four $v and/or $x subdivisions. The 100 most frequently assigned form subdivisions are shown in table 1. The complete table is available on OCLC's Web site (http://wcp.oclc.org/fast).

As shown in table 1, the most frequently used form is **—Congresses,** which has been assigned a total of 1,109,724 times in WorldCat, including the 317,800 times it had been assigned by the Library of Congress. The "Relative use by the Library of Congress" column indicates the relative frequency that the form was assigned by the Library of Congress compared to its use in contributed records. For example, the relative use of 50% for **—Periodicals** means that it is assigned by the Library of Congress about half as often as it is in contributed records. By contrast, **—Biography** is assigned more than twice as frequently by the Library of Congress. The wide variation in relative use by the Library of Congress reflects both a difference in cataloging practice and in the types of materials cataloged.

The forms identified contain 1 to 4 subdivisions, excluding any geographic subdivisions. The majority (63%) of the forms contain 2 subdivisions each, and only a quarter of all forms contain a single subdivision; 11% contain 3 subdivisions each. The only combination identified with 4 subdivisions was **—Biography—Dictionaries—Arabic—Early works to 1800** (v-v-x-v)

However, forms with 2 or more subdivisions were rarely assigned. Forms consisting of a single subdivision accounted for almost 95% of all assignments. The longer forms tend to be very specific, greatly limiting their applicability. Forms with a single subdivision were assigned an average of 12,587 times, those with 2 subdivisions were assigned an average of 311 times and forms with 3 or more subdivisions were assigned only an average of 14 times.

In general, the use of forms is very skewed; the 10 most assigned forms, **—Congresses, —Periodicals, —Biography, —Bibliography, —Directories, —Statistics, —Maps, —Handbooks, manuals, etc., —Catalogs,** and **—Fiction,** account for more than half of all assignments. The 100 most used forms account for more than 90% of all assignments. The remaining 2,463 forms account for less than 10% of all uses. The complete usage distribution is shown in figure 1.

## Conclusion

Form subdivisions, which describe what the document "is" rather than what it is "about," represent an aspect of the subject distinct from the topical aspect. However, to effectively utilize the form information requires that the form subdivisions be explicitly identified. Until recently, when the Library of Congress started explicitly assigning the *$v* subfield code, form subdivisions were coded as general subdivisions making them indistinguishable from topical subdivisions. OCLC's WorldCat contains more than eight million unique subject headings that potentially could contain form subdivisions. Identifying these potential forms is difficult since many subdivisions can be either topical or form depending on the context. Manual efforts to recode form subdivisions are slow and error-prone due to the complexity of the coding guidelines.

An algorithm was developed to identify and recode form subdivisions in Library of Congress topical and geographic subject headings. The algorithm proved to be highly reliable with an error rate estimated to be less than 0.1%, significantly less than the observed error rate for manual coding. The algorithm identified 2,563 unique forms or combinations of form subdivisions in WorldCat. The usage of these forms was very uneven; the 10 most frequently assigned form subdivisions accounted for more than half of all assignments. Perhaps the greatest advantages of the algorithmic approach are the high accuracy rate and the ability to handle a large number of operations efficiently.

Table 1. Usage of Common Form Subdivisions

| Total WorldCat Usage | Library of Congress Usage | Relative Use by the Library of Congress | Subfield Coding | Form Subdivision |
|---|---|---|---|---|
| 1,109,724 | 317,800 | 161 | v | Congresses |
| 994,462 | 110,689 | 50 | v | Periodicals |
| 522,795 | 200,163 | 249 | v | Biography |
| 342,555 | 63,605 | 92 | v | Bibliography |
| 337,366 | 44,244 | 61 | v | Directories |
| 315,553 | 49,704 | 75 | v | Statistics |
| 315,532 | 91,673 | 165 | v | Maps |
| 300,303 | 44,642 | 70 | v | Handbooks, manuals, etc. |
| 280,777 | 45,163 | 77 | v | Catalogs |
| 275,552 | 81,549 | 169 | v | Fiction |
| 264,891 | 74,228 | 156 | v | Exhibitions |
| 229,330 | 67,552 | 168 | v | Juvenile literature |
| 189,742 | 11,682 | 26 | v | Scores |
| 181,334 | 18,542 | 46 | v | Early works to 1800 |
| 122,016 | 42,750 | 217 | v | Case studies |
| 119,529 | 27,044 | 118 | v—v | Statistics—Periodicals |
| 113,277 | 24,665 | 112 | v | Dictionaries |
| 109,822 | 4,712 | 18 | v | Scores and parts |
| 77,729 | 16,968 | 112 | v | Sources |
| 70,872 | 9,178 | 60 | v | Texts |
| 67,174 | 16,941 | 136 | v | Pictorial works |
| 65,201 | 2,395 | 15 | v | Excerpts |
| 59,779 | 5,521 | 41 | v | Juvenile fiction |
| 58,980 | 20,452 | 213 | v | Guidebooks |
| 58,002 | 1,564 | 11 | v | Vocal scores with piano |
| 57,301 | 11,229 | 98 | v—v | Bibliography—Catalogs |
| 56,722 | 7,465 | 61 | v | Problems, exercises, etc. |
| 55,947 | 9,926 | 87 | v | Indexes |
| 54,698 | 5,119 | 41 | v | Examinations, questions, etc. |
| 53,669 | 2,371 | 19 | v | Sermons |
| 52,035 | 10,590 | 103 | v | Poetry |
| 47,310 | 3,120 | 28 | v | Curricula |
| 46,842 | 3,952 | 37 | v | Software |
| 46,436 | 920 | 8 | v | Parts |
| 40,967 | 154 | 2 | v | Juvenile films |
| 39,323 | 7,966 | 102 | v | Drama |
| 35,738 | 10,143 | 159 | v | Cases |
| 35,218 | 11,377 | 192 | v | Popular works |
| 34,810 | 1,950 | 24 | v | Readers |
| 34,026 | 831 | 10 | v | Collections |
| 33,077 | 4,394 | 62 | v | Specimens |
| 31,685 | 4,342 | 64 | v | Tables |
| 31,644 | 7,797 | 131 | v | Classification |
| 31,471 | 5,236 | 80 | v—x | Dictionaries—English |
| 30,681 | 990 | 13 | v | Songs and music |
| 30,469 | 11,123 | 231 | v | Correspondence |
| 29,369 | 4,800 | 79 | v | Facsimiles |
| 28,831 | 8,892 | 179 | v | Interviews |
| 28,261 | 10,939 | 254 | v | Miscellanea |
| 25,650 | 1,443 | 24 | v | Librettos |
| 25,548 | 4,863 | 94 | v | Outlines, syllabi, etc. |
| 25,254 | 6,609 | 142 | v | Identification |
| 23,340 | 1,207 | 22 | v | Music |
| 23,062 | 1,667 | 31 | v | Designs and plans |
| 22,330 | 3,731 | 81 | v | Abstracts |
| 20,660 | 1,122 | 23 | v | Controversial literature |
| 20,463 | 46 | 1 | v | Slides |
| 20,259 | 530 | 11 | v | Excerpts, arranged |

Table 1. Usage of Common Form Subdivisions, continued

| Total WorldCat Usage | Library of Congress Usage | Relative Use by the Library of Congress | Subfield Coding | Form Subdivision |
|---|---|---|---|---|
| 20,191 | 6,311 | 183 | v | Translations into English |
| 20,049 | 7,083 | 220 | v | Folklore |
| 19,271 | 252 | 5 | v | Newspapers |
| 18,741 | 4,554 | 129 | v | Terminology |
| 18,468 | 9,976 | 472 | v—v | Biography—Juvenile literature |
| 17,856 | 4,639 | 141 | v | Forms |
| 17,535 | 56 | 1 | v | Photographs |
| 17,226 | 1,822 | 48 | v | Laboratory manuals |
| 16,982 | 265 | 6 | v | Studies and exercises |
| 15,848 | 2,139 | 63 | v | Personal narratives |
| 15,547 | 1,692 | 49 | v | Programmed instruction |
| 15,391 | 2,839 | 91 | v | Glossaries, vocabularies, etc. |
| 15,378 | 2,205 | 67 | v—v | Bibliography—Periodicals |
| 15,182 | 127 | 3 | v | Hymns |
| 14,712 | 3,643 | 132 | v—x | Dictionaries—German |
| 14,664 | 3,238 | 114 | v | Yearbooks |
| 13,628 | 3,074 | 117 | v—x | Dictionaries—French |
| 13,427 | 4,430 | 198 | v | Anecdotes |
| 13,303 | 696 | 22 | v | Solo with piano |
| 12,740 | 5,019 | 261 | v | Humor |
| 12,611 | 794 | 27 | v | Observations |
| 12,269 | 2,468 | 101 | v | Bio-bibliography |
| 12,011 | 1,720 | 67 | v | Textbooks for foreign speakers |
| 11,961 | 405 | 14 | v | Juvenile |
| 11,909 | 3,585 | 173 | v | Patterns |
| 11,303 | 2,681 | 125 | v | Atlases |
| 11,078 | 1,069 | 43 | v | Databases |
| 10,978 | 2,672 | 129 | v | Registers |
| 10,790 | 4,494 | 287 | v | Diaries |
| 10,567 | 3,450 | 195 | v | Amateurs' manuals |
| 10,541 | 300 | 12 | v | Textbooks |
| 10,360 | 3,727 | 226 | v | Digests |
| 10,313 | 412 | 17 | v | Methods |
| 10,242 | 1,902 | 92 | v | Study guides |
| 10,064 | 473 | 20 | v | Instrumental settings |
| 9,887 | 3,002 | 175 | v—x | Dictionaries—Chinese |
| 9,813 | 507 | 22 | v | Conversation and phrase books |
| 9,735 | 2,859 | 167 | v | Genealogy |
| 9,622 | 2,063 | 110 | v | Charts, diagrams, etc. |
| 9,514 | 2,499 | 143 | v—x | Dictionaries—Polyglot |
| 8,994 | 2,519 | 156 | v—x | Dictionaries—Japanese |
| 8,983 | 21,46 | 125 | v | Portraits |

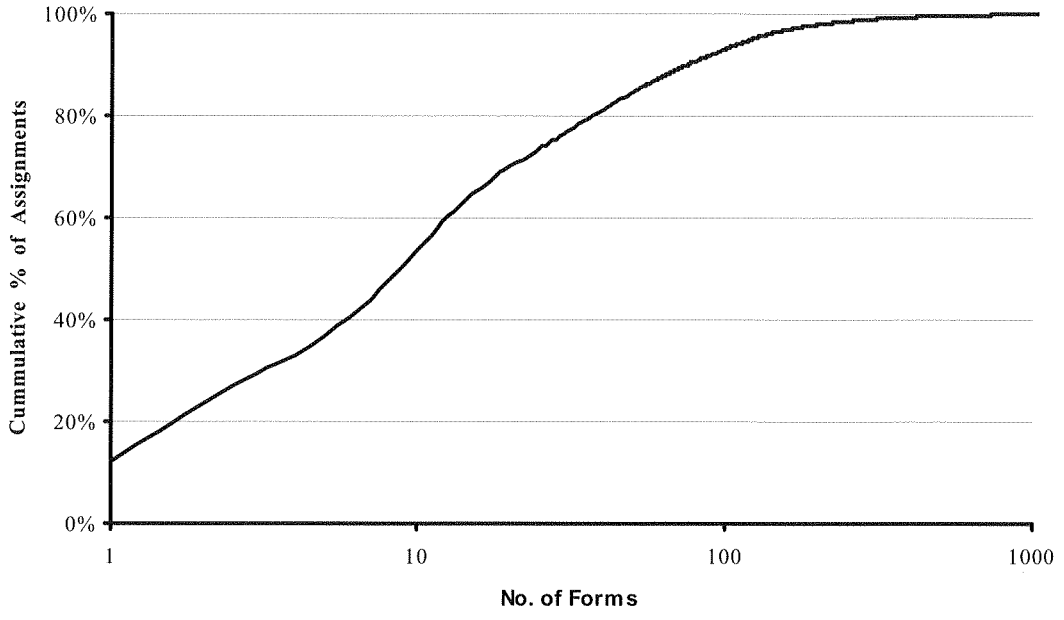



**Figure 1.** Form Subdivision Usage

## Works Cited

Chan, Lois Mai, Eric Childress, Rebecca Dean, Edward T. O'Neill, and Diane Vizine-Goetz. 2001. A faceted approach to subject data in the Dublin Core metadata record. *Journal of Internet Cataloging* 4, no. 1/2: 35–47.

Haykin, David Judson. 1951. *Subject headings: A practical guide.* Washington, D.C.: Govt. Print. Off.

Hemmasi, Harriette, David Miller, and Mary Charles Lasater. 1999. Access to form data in online catalogs. ALCTS Online Newsletter 10, no. 4. Accessed Aug. 31, 2001, www.ala.org/alcts/alcts_ news/v10n4/formdat2.html.

Library of Congress. 1906. *Preliminary list of subject subdivisions.* Washington, D.C.: Govt. Print. Off.

———. 1975. *Library of Congress subject headings.* Washington, D.C.: Library of Congress.

Library of Congress. Cataloging Policy and Support Office. 1996–. *Subject cataloging manual: Subject headings.* Washington, D.C.: Cataloging Distribution Service, Library of Congress.

———. 2000. *Free-floating subdivisions: An alphabetical index.* 12th ed. Washington, D.C.: Cataloging Distribution Service, Library of Congress.

OCLC Online Computer Library Center, Inc., Office of Research. *Usage of form subdivisions.* Forthcoming, http://wcp.oclc.org/fast/.

O'Hara Conway, Martha, ed. 1992. *The future of subdivisions in the Library of Congress subject headings system: Report from the Subject Subdivisions Conference, May 9–12, 1991.* Washington, D.C.: Library of Congress, Cataloging Distribution Service.

*Subject data in the metadata record recommendations and rationale: A report from the ALCTS/SAC/Subcommittee on Metadata and Subject Analysis.* 1999. Accessed Aug. 31, 2001. www.govst.edu/users/gddcasey/sac/MetadataReport.html.

# Computer Cataloging of Electronic Journals in Unstable Aggregator Databases

## The Hong Kong Baptist University Library Experience

Yiu-On Li and Shirley W. Leung

*The growth and use of aggregator databases have presented libraries with both new opportunities for collection enhancement and new challenges of bibliographic control. How to integrate full-text electronic journal titles in unstable aggregator databases into a library's OPAC has been an especially taxing matter for libraries. This article describes the Hong Kong Baptist University Library's effort to integrate full-text electronic journal titles from three large, unstable aggregator databases into its INNOPAC-based OPAC. The library's electronic journal computer program (EJCOP) does this in a simple, direct, consistent, and accurate manner and addresses some of the issues elaborated in the January 2000 Final Report of the Task Group on Journals in Aggregator Databases of the Standing Committee on Automation of the Library of Congress Program for Cooperative Cataloging.*

The rapid growth of aggregator databases in the past decade has allowed libraries to enrich their collections and offer unprecedented access to full-text journals for users. Subscription to large aggregator databases such as Academic Universe, Academic Search FullText Elite, and ProQuest Direct can expand a library's access to thousands of new journal titles in full-text form overnight—a phenomenon not possible in the environment of print subscriptions. At the same time, this new and exciting opportunity has brought some highly challenging bibliographic control problems for librarians in terms of how to identify, display, and maintain the currency of titles in aggregator databases in a manner that is user-oriented and cost-effective. Much has been written in the professional literature on the issues, problems, possible solutions, and recommendations related to the bibliographic control of aggregator databases. Two very useful recent publications on this topic are "Aggregation or Aggravation? Optimizing Access to Full-Text Journals" and the *Final Report* of the Library of Congress Program for Cooperative Cataloging Standing Committee on Automation's Task Group on Journals in Aggregator Databases (PCC Task Group). The former gives an excellent overview of the "benefits of aggregations, the challenges of describing aggregated full-text articles, and several alternatives

**Yiu-On Li** (yoli@hkbu.edu.hk) is Assistant Librarian (Systems) of Hong Kong Baptist University Library.

**Shirley W. Leung** (sleung@hkbu.edu.hk) is University Librarian of Hong Kong Baptist University Library.

We would be glad to share the EJCOP with interested libraries with the understanding that (1) the HKBU Library will be acknowledged as the developer of the EJCOP, and that (2) we would be informed of enhancements or other creative use made of the EJCOP in the spirit of continuous improvement to the program for the benefit of the library and information community. Please contact Yiu-On Li at yoli@hkbu.edu.hk for more information.

Manuscript received April 5, 2001; accepted for publication June 21, 2001.

for integrating print and electronic journal titles in library catalogs" (Calhoun and Kara 2000). The latter outlines its working assumptions and a set of recommendations on record creation and record maintenance, complete with proposed data elements for different methods of generating records for the titles in an aggregator database (PCC 2000).

This article describes the Hong Kong Baptist University (HKBU) Library's solution to integrating the full-text electronic journal titles in aggregator databases into its INNOPAC-based OPAC in a simple, direct, consistent, and accurate manner through its electronic journal computer program (EJCOP).[1]

## The HKBU Library

The HKBU Library (www.hkbu.edu.hk/lib) has a collection of over 670,000 volumes and approximately 4,940 printed journal subscriptions in a 60–40% split between English- and Chinese-language titles respectively. It has a staff of 14 full-time equivalent professional librarians and 46 full-time equivalent library assistants and part-time staff. The library began significantly expanding its access to electronic resources in 1999. In addition to acquiring aggregator databases such as JSTOR and Project Muse, it also acquired three aggregator databases made up of sizable collections of titles that may be available one month but unavailable a few months later—Lexis Nexis Academic Universe, EBSCO's Academic Search FullText Elite, and two subsets of Bell & Howell's ProQuest Direct: ABI/INFORM Global and Academic Research Library. (Whenever ProQuest Direct is referred to in the HKBU Library context of this article, it denotes these two subsets.)

While the selection and the acquisition of aggregator databases were rather simple steps, the issue of how to provide optimal access to full-text electronic journal titles was an entirely different and difficult matter. As we were keenly aware that even large academic libraries had found it taxing and laborious to follow the traditional approach of creating records for each full-text title in unstable aggregator databases and maintaining them on a consistent and current basis, we felt strongly that we needed to apply a new workflow to the bibliographic control of titles in these aggregator databases.

## Applying a New Systematic Workflow to Processing Aggregator Databases

We decided to take a two-pronged approach to handle the workflow of aggregator databases, depending on whether they are stable or unstable.

### Handling of Stable Aggregator Databases

We consider an aggregator database stable when we have a direct subscription to the journal titles either as a result of our print subscription (e.g., Cambridge Journals Online, Oxford Journals Online, SWETSNet) or the aggregator has some control over the journal titles in the database (e.g., Emerald Online, JSTOR, Project Muse). More specifically, if we were to cancel a journal title published by Cambridge University Press, we would know that the electronic version of the journal would also be deleted from our Cambridge Journals Online file. In the case of JSTOR, the list of titles we have access to is made known to us, and the titles therein remain stable. Calhoun and Kara call this type of aggregator database "vanilla" aggregations with the following attributes:

> All the titles have some common element, whether they are from the same publisher or cover the same broad subject, or both. For each of the journal titles available in the collection there are complete issues (or relatively complete issues with only minor, known differences from their print counterparts), which are accessible by both journal title and specific issue. . . . It is easy to identify the parts of the collection, relatively easy to maintain over time since the aggregator maintains a stable journal title list and would notify the subscriber of changes to the collection (2000).

The process we adopted to handle these "vanilla," or stable, aggregations is through a team effort by our acquisitions section and our cataloging section as follows:

1. If we already have a catalog record for the print version of the journal in our OPAC, our acquisitions staff responsible for serials will add the following MARC tags to the record:
   a. 599: coverage information note (usually the start date of the electronic full-text coverage)
   b. 740: database title
   c. 856: URL (direct URL to journal title page)
2. If we do not have a print version of an electronic journal title, the cataloging section will first create a MARC record by downloading it from OCLC and modify the record by adding the three above-mentioned tags—599, 740, and 856.

### Handling of Unstable Aggregator Databases

Calling them the "tutti-frutti surprise" aggregator databases, Calhoun and Kara described them this way:

> These can be quite large and amorphous. Titles available one month might not be available six months later, the user perhaps cannot select a specific journal title and issue, and the full issue of a journal might not be available, but only its articles pertaining to the broad subject category of the collection. Aggregation content might include monographs, reference books, and pamphlets (or parts, but not necessarily the complete full text, of these many publications). . . . If they're successful in providing a unique or value-enhanced resource to the collection, their content, whether clearly defined or broadly defined, can become an integral part of a library's online collections and services (2000).

The three most commonly cited examples of unstable aggregator databases are Academic Universe, Academic Search FullText Elite, and ProQuest Direct.

We recognized early on that it would be impossible for us to apply the same workflow for handling the unstable aggregator databases as we used for stable aggregator databases, so we decided to design a computer program to handle the processing. Before doing this, we set the goal of the project: to enable our users to access the full-text electronic journal titles in a simple, direct, consistent, and accurate manner. We also defined what we meant by these terms:

- *Simple.* We wanted to provide one-stop shopping for users to access electronic journals by having all our electronic journals fully searchable and integrated into our OPAC. In other words, if an electronic journal is included in different aggregator databases, the different URL links to the title will be contained in a single record rather than in several different catalog records.
- *Direct.* To save the time of our users, and our reference staff when they help users to retrieve desired electronic journal titles, we wanted to have the URL links of all electronic journals point directly to the journal title page and not the aggregator database Web site.
- *Consistent.* We wanted to avoid inconsistent handling of the records that often occurs when data elements are entered by staff in different sessions. Consequently, a unified and consistent method for creating records for electronic journal titles is required to facilitate data-sharing among different library sections.
- *Accurate.* Because of the sometimes-in-and-sometimes-out nature of records in unstable aggregator databases, we considered it important to develop an efficient and timely method of record maintenance to ensure the accuracy of the records of electronic journal titles that are integrated into our OPAC.

## Development of the Electronic Journal Computer Program

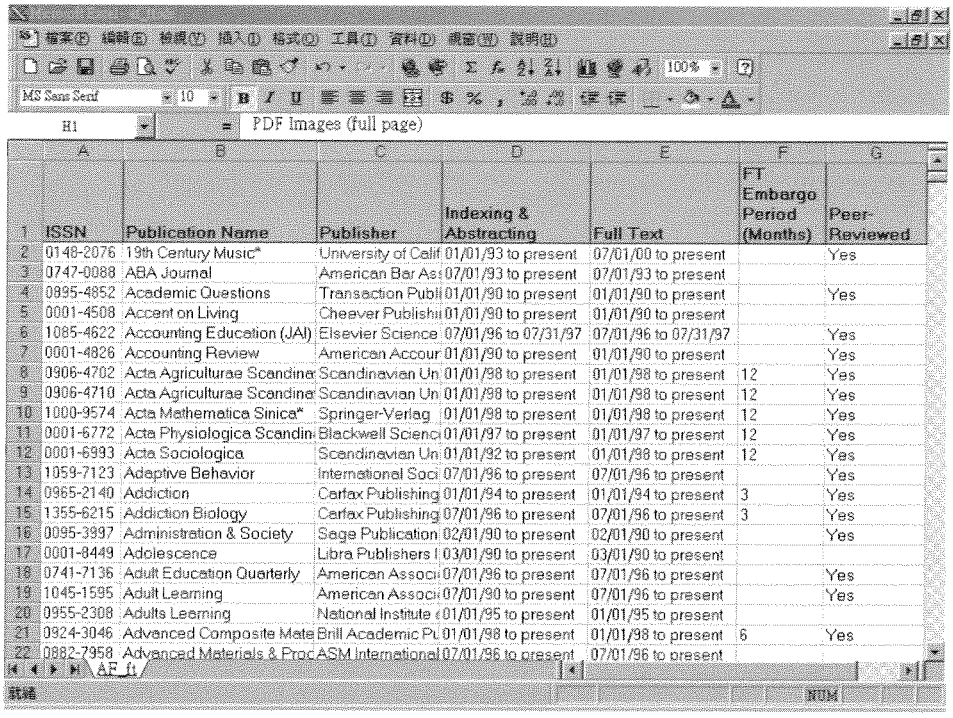
The library's systems office began the project to develop a computer program for the creation and maintenance of records for full-text electronic journal titles in unstable aggregator databases in January 2000 with support and consultation from staff in the acquisitions, cataloging, and reference and user education sections. The program, named Electronic Journal Computer Program (EJCOP), was completed by April 2000 to handle three main functions:

- To generate MARC records for electronic journals on a single-record approach with direct URL link to journal title pages, thus bypassing the aggregator database vendor's Web site
- To compile a complete HKBU Library electronic journal title-subscription list of stable and unstable aggregator databases by converting the MARC records for the titles in our OPAC into HTML files and exporting them directly to our Web site
- To devise an efficient and timely ongoing record-maintenance mechanism to ensure the accuracy and currency of information

In order to make available the full list of electronic journal titles for user access as quickly as possible, we initially created brief MARC records for the three unstable aggregator databases by using the title lists provided by the database vendors. These title lists are usually updated monthly or bimonthly by aggregators to reflect the latest changes in the databases. It is useful to note that these aggregator title lists have different formats as shown in figures 1–3. After the completion and successful testing of the three tasks in April, we then went back to upgrade the brief MARC records into full MARC records in early November 2000 by combining the vendor title lists with the MARC record sets supplied by aggregators. This effort resulted in 5,998 (90%) full MARC records with subject or corporate author access provided and 667 (10%) brief records having basic access information including ISSN and journal title. Given that there is no guarantee that they will appear in the next batch of record sets uploaded, we decided to leave the brief records just as they were.

## Issues Related to Vendor-Supplied MARC Record Sets for Unstable Aggregator Databases

Based on the data elements for MARC records recommended by the PCC Task Group, some vendors began creating MARC record sets for titles in their aggregator

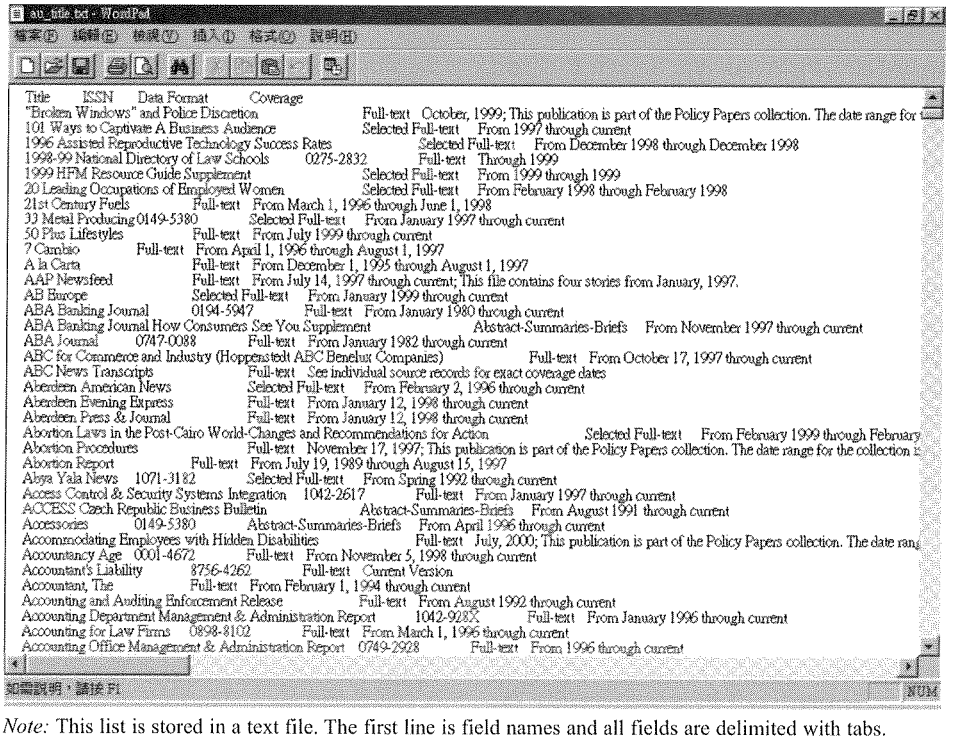


PDF Images (full page)

| | ISSN | Publication Name | Publisher | Indexing & Abstracting | Full Text | FT Embargo Period (Months) | Peer-Reviewed |
|---|---|---|---|---|---|---|---|
| 2 | 0148-2076 | 19th Century Music* | University of Calif | 01/01/93 to present | 07/01/00 to present | | Yes |
| 3 | 0747-0088 | ABA Journal | American Bar As | 07/01/93 to present | 07/01/93 to present | | |
| 4 | 0895-4852 | Academic Questions | Transaction Publi | 01/01/90 to present | 01/01/90 to present | | Yes |
| 5 | 0001-4508 | Accent on Living | Cheever Publishi | 01/01/90 to present | 01/01/90 to present | | |
| 6 | 1085-4622 | Accounting Education (JAI) | Elsevier Science | 07/01/96 to 07/31/97 | 07/01/96 to 07/31/97 | | Yes |
| 7 | 0001-4826 | Accounting Review | American Accour | 01/01/90 to present | 01/01/90 to present | | Yes |
| 8 | 0906-4702 | Acta Agriculturae Scandina | Scandinavian Un | 01/01/98 to present | 01/01/98 to present | 12 | Yes |
| 9 | 0906-4710 | Acta Agriculturae Scandina | Scandinavian Un | 01/01/98 to present | 01/01/98 to present | 12 | Yes |
| 10 | 1000-9574 | Acta Mathematica Sinica* | Springer-Verlag | 01/01/98 to present | 01/01/98 to present | 12 | Yes |
| 11 | 0001-6772 | Acta Physiologica Scandin | Blackwell Scienc | 01/01/97 to present | 01/01/97 to present | 12 | Yes |
| 12 | 0001-6993 | Acta Sociologica | Scandinavian Un | 01/01/92 to present | 01/01/98 to present | 12 | Yes |
| 13 | 1059-7123 | Adaptive Behavior | International Soci | 07/01/96 to present | 07/01/96 to present | | Yes |
| 14 | 0965-2140 | Addiction | Carfax Publishing | 01/01/94 to present | 01/01/94 to present | 3 | Yes |
| 15 | 1355-6215 | Addiction Biology | Carfax Publishing | 07/01/96 to present | 07/01/96 to present | 3 | Yes |
| 16 | 0095-3997 | Administration & Society | Sage Publication | 02/01/90 to present | 02/01/90 to present | | Yes |
| 17 | 0001-8449 | Adolescence | Libra Publishers I | 03/01/90 to present | 03/01/90 to present | | |
| 18 | 0741-7136 | Adult Education Quarterly | American Associ | 07/01/96 to present | 07/01/96 to present | | Yes |
| 19 | 1045-1595 | Adult Learning | American Associ | 07/01/90 to present | 07/01/96 to present | | Yes |
| 20 | 0955-2308 | Adults Learning | National Institute | 01/01/95 to present | 01/01/95 to present | | |
| 21 | 0924-3046 | Advanced Composite Mate | Brill Academic Pu | 01/01/98 to present | 01/01/98 to present | 6 | Yes |
| 22 | 0882-7958 | Advanced Materials & Proc | ASM International | 07/01/96 to present | 07/01/96 to present | | |

*Note:* This is stored in an Excel file.

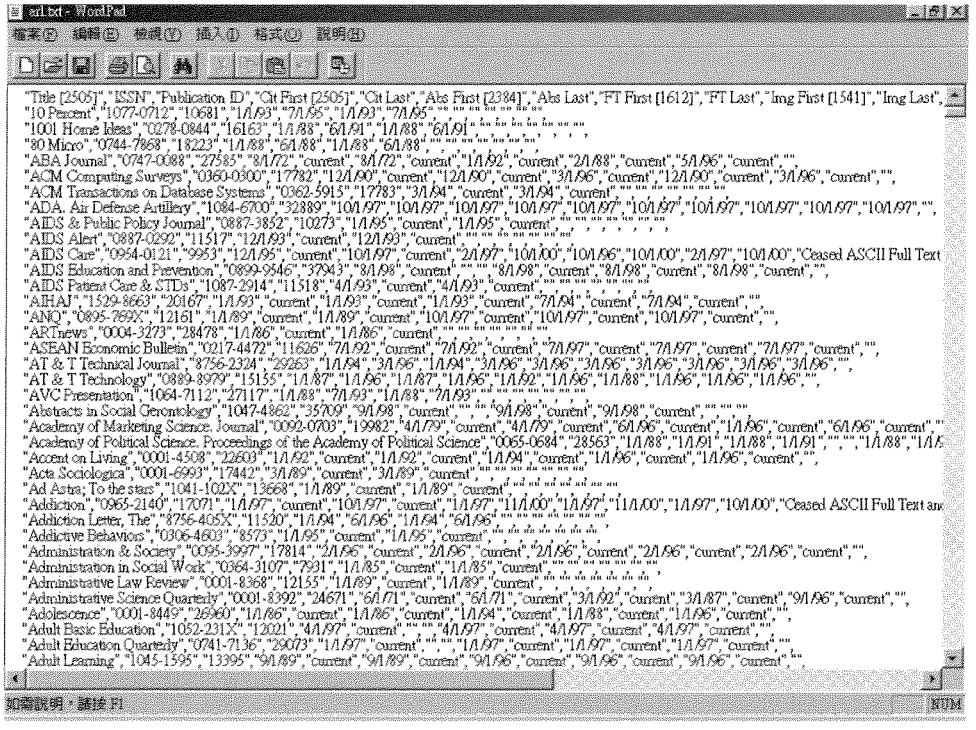
**Figure 1.** Academic Search FullText Elite Title List



```
Title	ISSN	Data Format	Coverage
"Broken Windows" and Police Discretion		Full-text	October, 1999; This publication is part of the Policy Papers collection. The date range for t
101 Ways to Captivate A Business Audience		Selected Full-text	From 1997 through current
1996 Assisted Reproductive Technology Success Rates		Selected Full-text	From December 1998 through December 1998
1998-99 National Directory of Law Schools	0275-2832	Full-text	Through 1999
1999 HFM Resource Guide Supplement		Selected Full-text	From 1999 through 1999
20 Leading Occupations of Employed Women		Selected Full-text	From February 1998 through February 1998
21st Century Fuels		Full-text	From March 1, 1996 through June 1, 1998
33 Metal Producing	0149-5380	Selected Full-text	From January 1997 through current
50 Plus Lifestyles		Full-text	From July 1999 through current
7 Cambio		Full-text	From April 1, 1996 through August 1, 1997
A la Carta		Full-text	From December 1, 1995 through August 1, 1997
AAP Newsfeed		Full-text	From July 14, 1997 through current; This file contains four stories from January, 1997.
AB Europe		Selected Full-text	From January 1999 through current
ABA Banking Journal	0194-5947	Full-text	From January 1980 through current
ABA Banking Journal How Consumers See You Supplement		Abstract-Summaries-Briefs	From November 1997 through current
ABA Journal	0747-0088	Full-text	From January 1982 through current
ABC for Commerce and Industry (Hoppenstedt ABC Benelux Companies)		Full-text	From October 17, 1997 through current
ABC News Transcripts		Full-text	See individual source records for exact coverage dates
Aberdeen American News		Selected Full-text	From February 2, 1996 through current
Aberdeen Evening Express		Full-text	From January 12, 1998 through current
Aberdeen Press & Journal		Full-text	From January 12, 1998 through current
Abortion Laws in the Post-Cairo World-Changes and Recommendations for Action		Selected Full-text	From February 1999 through February
Abortion Procedures		Full-text	November 17, 1997; This publication is part of the Policy Papers collection. The date range for the collection i
Abortion Report		Full-text	From July 19, 1989 through August 15, 1997
Abya Yala News	1071-3182	Selected Full-text	From Spring 1992 through current
Access Control & Security Systems Integration	1042-2617	Full-text	From January 1997 through current
ACCESS Czech Republic Business Bulletin		Abstract-Summaries-Briefs	From August 1991 through current
Accessories	0149-5380	Abstract-Summaries-Briefs	From April 1996 through current
Accommodating Employees with Hidden Disabilities		Full-text	July, 2000; This publication is part of the Policy Papers collection. The date rang
Accountancy Age	0001-4672	Full-text	From November 5, 1998 through current
Accountant's Liability	8756-4262	Full-text	Current Version
Accountant, The		Full-text	From February 1, 1994 through current
Accounting and Auditing Enforcement Release		Full-text	From August 1992 through current
Accounting Department Management & Administration Report	1042-928X	Full-text	From January 1996 through current
Accounting for Law Firms	0898-8102	Full-text	From March 1, 1996 through current
Accounting Office Management & Administration Report	0749-2928	Full-text	From 1996 through current
```

*Note:* This list is stored in a text file. The first line is field names and all fields are delimited with tabs.

**Figure 2.** Academic Universe Title List




*Note:* This list is stored in a text file. The first line is field names and each field is enclosed by double quotations.

**Figure 3.** ProQuest Direct Title List

databases in 1999 (PCC 2000). As of this writing, EBSCO provides MARC record sets for Academic Search FullText Elite for subscriber libraries free of charge; Bell & Howell charges $1 per MARC record for ProQuest databases; and MARC record sets are not yet available for Academic Universe.

The availability of vendor-provided MARC record sets has definitely made the creation and maintenance of electronic journal cataloging records much easier. In theory, libraries can simply upload the new complete record sets to their OPACs regularly to update existing records automatically. Records with a prior uploaded date can be identified as obsolete records and removed by a global delete function. In practice, as we discovered, there are some issues that libraries should be aware of and resolve before the MARC record sets provided by the aggregators can be used optimally.

### Duplicate Records

As there are duplicate titles among Academic Universe, Academic Search FullText Elite, and ProQuest Direct, the question was how to merge the duplicate records and display them in one record. An analysis we made in November 2000 showed that there was a 15% duplication rate among the three databases as shown in table 1.

Therefore, this duplication problem had to be dealt with before we could attain the goal of having different URL links for the title contained in a single record rather than in several records. As we also discovered, the MARC record sets provided by the two aggregator database vendors have different formats for Tag 035, which is used for record overlay by us, as illustrated in figure 4.

### Tag-Related Concerns

Some tags in the MARC record sets provided by aggregator database vendors may not be supported due to local parameters or a library's system design. For example, our local system parameters do not support Tag 773. While this problem can be solved by asking the library system vendor to change this field to an indexed field, another solution is to convert all Tag 773s to Tag 740s, which are indexed. On the other hand, the INNOPAC system does not accommodate display of the journal title coverage note in subfield $3 of Tag 856. A more ubiquitous problem for all libraries, however, is the treatment of subfield $u in Tag 856, which is commonly reserved for the URL of the individual electronic journal

title on a discrete basis. Because the IP range of the subscribing library is encrypted in the URL to the journal title page as is the case with ProQuest Direct and Academic Universe, the URL may be very long and exceed the acceptable length of Tag 856 allowed by some library systems. An in-house PERL script can be used to shorten this long URL. This, however, means that Tag 856 cannot straightforwardly be used as a default field for all libraries. More about this will be discussed in the section on treatment of Tag 856.

In short, although the MARC record sets provided by aggregator database vendors are very useful, they require some additional work before they can be uploaded into a library's OPAC.

### First Function of the EJCOP

The first function of the EJCOP is to:

- Generate MARC records for full-text electronic journals on a single record approach with direct URL links to journal title pages, thus bypassing the aggregator database vendor's Web site.
- Generate MARC records for full-text electronic journals on a single record approach. In carrying out this function, the EJCOP requires the complete title lists and the MARC record sets from the database vendors.2 If MARC records are not available (as is the case with Academic Universe), downloading from OCLC or creating them locally will be necessary. This is, unfortunately, a manual and title-by-title process. Figure 5 shows how the MARC record sets and the title lists are merged together to generate a complete list of electronic journal titles with MARC records.

Before exporting the final file of MARC records that represent only full-text titles to our INNOPAC system, the EJCOP also performs the following tasks:

1. Add basic or in-house Tag 001, 035, 599,[3] 740, and 856 to each matching MARC record (see figure 10 for a detailed description of these tags).
2. Differentiate titles with and without full text so that the final list of titles with MARC records exported to our INNOPAC system contains only full-text titles. Our decision to include solely full-text electronic journal titles from the unstable aggregator databases was to minimize user confusion that may result from encountering some titles with full-text articles, some with citations and abstracts, and some with citations only. By focusing only on journal titles with full-text articles, users are provided with the assurance of only accessing full-text titles through an OPAC search.
3. Combine duplicate electronic journal titles included in different aggregator databases to form a single record. The *ABA Banking Journal,* for instance, is included in all three of the aggregated databases. All three records were gathered together to form a single record as shown in public display format (figure 6) and in MARC record format (figure 7).

On the other hand, if no matching MARC record is found, a brief record with basic tags is generated automatically based on the information extracted from the vendor-

**Table 1.** Duplicate Titles in Three Aggregator Databases

| Database Title | AU | EB | PQD |
|---|---|---|---|
| Academic Universe | 4,376 | 162 | 275 |
| EBSCO's Academic Search FullText Elite | 162 | 1,342 | 731 |
| ProQuest Direct (ABI/Inform Global and Academic Research Library) | 275 | 731 | 2,012 |

*Note:* This comparison was made by matching the ISSN available in the vendor-supplied title lists. The total number of duplicate titles identified may be underestimated as many titles in Academic Universe did not provide any ISSN information.

| Database Title | Tag 035 |
|---|---|
| Academic Search FullText Elite | (EbpS)AFH0809741 |
| ProQuest Direct | (MiAaPQ)27493 and (OCoLC)5508974 |

**Figure 4.** Different Formats for Tag 035s in MARC Record Sets Provided by Aggregator Database Vendors

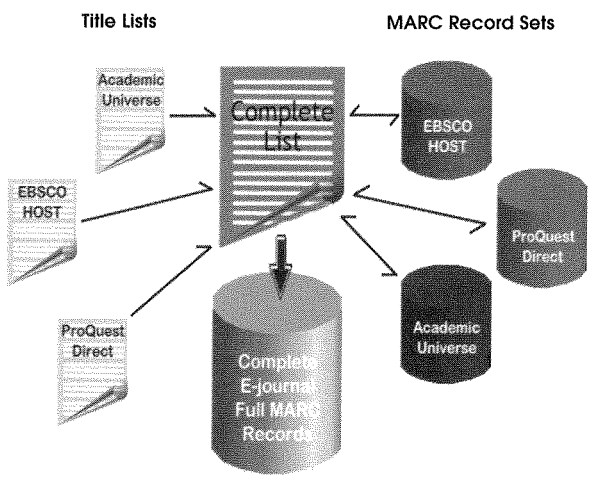



**Figure 5.** EJCOP MARC Record Conversion Diagram

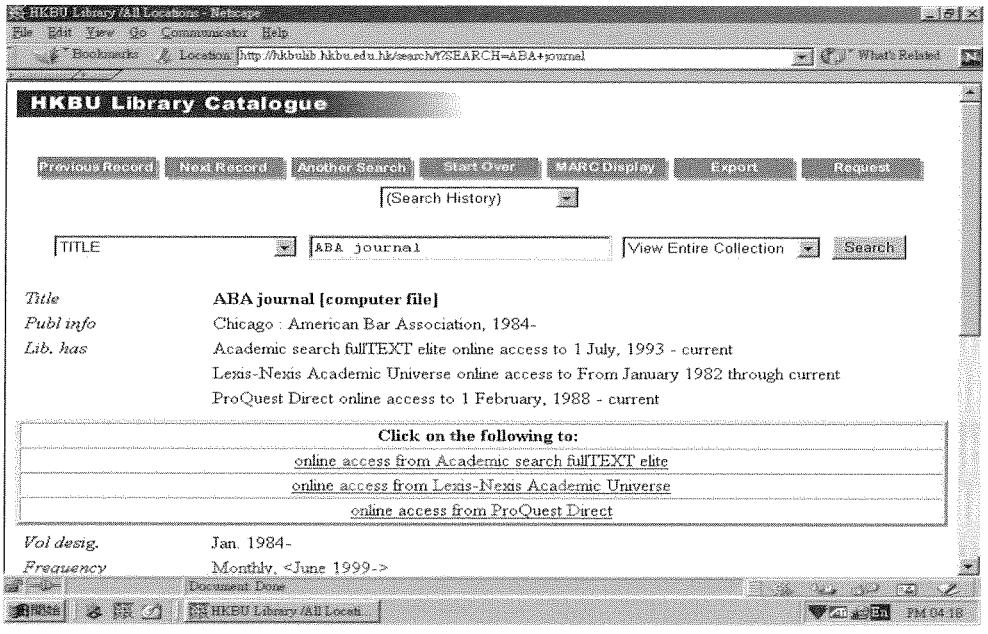



**Figure 6.** An Electronic Journal Title with Merged Full MARC Records in Public Display Format

```
001     hkbu_ejour01/13/01
006     m        d  ||||||
007     cr anu
008     840106c19849999ilumx1p   g   0   a0eng dnms a
022  0  0747-0088
035     ej0747-0088
245  00 ABA journal|h[computer file]
246  17 American Bar Association journal
260     Chicago :|bAmerican Bar Association,|c1984-
310     Monthly,|b<June 1999->
362  0  Jan. 1984-
500     Title from cover
515     Issues for Jan. 1984-     called also v. 70-
530     Online version of print publication
599     Academic search fullTEXT elite online access to 1 July, 1993 - current
599     Lexis-Nexis Academic Universe online access to  From January 1982
        through current
599     ProQuest Direct online access to 1 February, 1988 - current
610  20 American Bar Association|xPeriodicals
650   0 Law|zUnited States|xPeriodicals
710  2  American Bar Association
740  0  ABI/INFORM Global (ProQuest Direct)
740  0  Academic Research Library (ProQuest Direct)
740  0  Academic search fullTEXT elite
740  0  Lexis-Nexis Academic Universe
740  0  ProQuest Direct
780  00 American Bar Association.|tAmerican Bar Association journal|x0002
        -7596|w(OCoLC)1479565|w(DLC)   15017355
856     |uhttp://www.hkbu.edu.hk/lib-cgi/asfe.pl?t=07470088|zonline access
        from Academic search fullTEXT elite
856     |uhttp://www.hkbu.edu.hk/lib-cgi/au.pl?t=ABA+Journal|zonline access
        from Lexis-Nexis Academic Universe
856     |uhttp://www.hkbu.edu.hk/lib-cgi/pqd.pl?t=27585|zonline access from
        ProQuest Direct
```

**Figure 7.** An Electronic Journal Title with Merged Full MARC Records in MARC Format Display

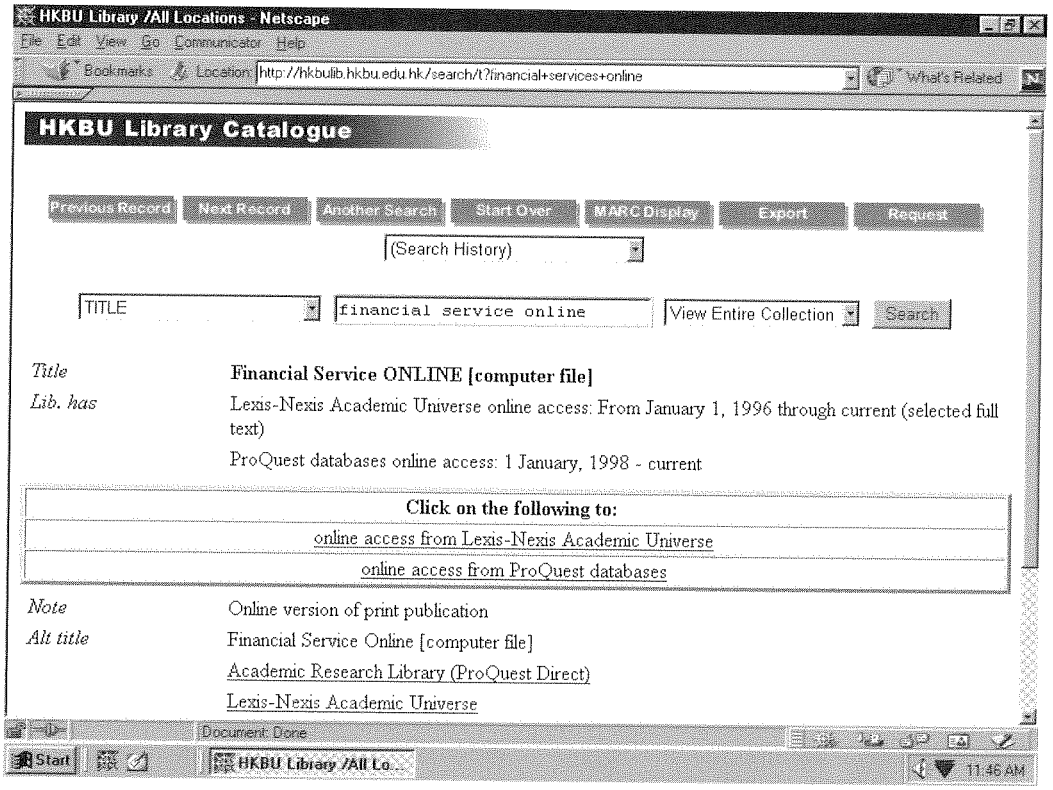



**Figure 8.** An Electronic Journal Title with Merged Brief MARC Records in Public Display Format

```
001     hkbu_ejour01/13/01
006     m        d  ||||||
007     cr anu
008                                    eng dnas a
022     1093-1244
035     ej1093-1244
245  00 Financial Service ONLINE|h[computer file]
246  3  Financial Service Online|h[computer file]
530     Online version of print publication
599     Lexis-Nexis Academic Universe online access:  From January 1, 1996
        through current (selected full text)
599     ProQuest databases online access: 1 January, 1998 - current
740  0  Academic Research Library (ProQuest Direct)
740  0  Lexis-Nexis Academic Universe
740  0  ProQuest Direct
856  40 |uhttp://www.hkbu.edu.hk/lib-cgi/au.pl?t=Financial+Service+ONLINE
        |zonline access from Lexis-Nexis Academic Universe
856  40 |uhttp://www.hkbu.edu.hk/lib-cgi/pqd.pl?t=37486|zonline access from
        ProQuest Direct
910     ej-brief
```

**Figure 9.** An Electronic Journal Title with Merged Brief MARC Records in MARC Format Display

**Table 2.** Basic Tags Supplied by EJCOP for All Electronic Journal MARC Records

| Tags | Descriptions | Field Value Generated by EJCOP |
|---|---|---|
| 001 | Control Number | Assigned by the organization creating, using, or distributing the record, library name with record created date, e.g., "hkbu_ejour11/12/00." |
| 022 | ISSN | ISSN |
| 035 | System Control Number | Dynamic unique record overlay point, see discussions below. |
| 245 | Title Statement | Journal title proper + GMD, i.e., title + "computer file]." |
| 599 | Note | Coverage note: Aggregator database title + "online access" + coverage date, e.g., "Academic search fullTEXT elite online access to 1 September, 2000-current." |
| 740 | Title Added Entry | Database titles, including titles for the name of aggregators and title of database subsets, for ABI electronic journal in ProQuest it will have two 740s, e.g., "ABI/INFORM Global (ProQuest Direct)" and "ProQuest Direct." These titles are used as links to the library OPAC for the electronic journal featured collections provided in the title page on the library Web site (see figure 12). |
| 856 | Electronic Location and Access | Direct URL to journal title page, may contain the default value of the in-house PERL scripts for some aggregator databases, e.g., ProQuest and Academic Universe. |

supplied title lists (see figures 8 and 9). These tags are included in the set of data elements recommended by the PCC Task Group as shown in table 2.

The obstacle encountered when creating direct URL link to journal title pages as a way of bypassing the aggregator database vendor's Web site is that Tag 856 in the INNOPAC system has a length limitation of 90 characters, but the direct URL links in Academic Universe and ProQuest Direct are often longer. For instance, the original direct URL to *21st Century Fuels*, a title in Academic Universe, is:

***http://cisweb.lexis-nexis.com/sourceselect/listSources.asp?srcpdn=academic&_session=cad47e44-14ef-11d4-b2b3-8a0c585aaa77.1.3133490923.73209.+.0.0&_state=&wchp=dGLkbb-lSlAl&_md5=a98aa4fef57ddd1310d948dad2a1eaf5&product=universe&unix=http%3A%2F%2Fweb.lexis-nexis.com%2Funiverse&extendRQ=Y&title***=21st+Century+Fuels

The above URL essentially contains two parts. The first is the fixed part (i.e., the bold and italicized portion from "http" to "title=") used as default value for all electronic journals in Academic Universe. The second is a variable part containing a special value for the particular journal title (i.e., the part after "title="). In order to overcome the INNOPAC system's Tag 856 limitation, an in-house Tag 856 for direct link to the electronic journal title is established by shortening the above URL by replacing the default value through the creation of an in-house PERL script:

http://www.hkbu.edu.hk/lib-cgi/au.pl?t=21st+Century+Fuels

It is important to note that the format of the variable parts is different among aggregator databases as shown in table 3. Furthermore, because Tag 856 containing the in-house PERL script in the aggregator-supplied MARC record sets is discrete for each subscribing library, this means each library must create its own in-house PERL script on its own computer server.

The EJCOP, however, has the capability of allowing different libraries to store their own default value of in-house PERL scripts for different aggregator databases. In addition, it has the capability of generating the variable part required by the PERL scripts as a generic program, thereby allowing different libraries to create their own MARC record sets. This is an important capability of the EJCOP because we think it has overcome the disadvantage as identified in the *Final Report* of the PCC Task Group. It considered local scripting by a single institution to create minimal-level record sets from vendor-supplied title/ISSN listings as the fourth best method for record creation, following the first best method of human-created analytics, as well as machine-derived analytics and machine-generated analytics from vendors as the second and third best method respectively. The disadvantage of local scripting by a single institution for record creation, as noted in the *Final Report*, is that:

> Individual libraries must do the work themselves, and the sets they produce are difficult or impossible to share with other libraries. No subject or corporate body access is included at all. This method is least likely to support deduping and record consolidation, especially if the ISSN is not available (PCC 2000).

In the case of the EJCOP, local scripting is applied, but the method used is automatic and sharable. Furthermore, 90% of the records in the machine-derived set by EJCOP

are full MARC records with subject or corporate body access provided. Lastly, the EJCOP has successfully designed an algorithm to solve the deduping and record consolidation problem by creating a relative Tag 035 as the basis of repetitive overlay for record maintenance (for details, see section on third function of the EJCOP).

### Second Function of the EJCOP

The EJCOP's second function is to compile a complete electronic title list by converting the MARC records in the OPAC into HTML files and exporting them to the library Web site.

After the cataloging records of all the electronic journal titles, for stable and unstable aggregator databases alike, are created and stored in the library's INNOPAC system, we then export the complete list from the INNOPAC system for EJCOP processing. The EJCOP can automatically reformat this complete list through a conversion program into HTML files for posting on the HKBU Library's Web site (www.hkbu.edu.hk/lib/ejours/index.html). This enables the users to browse and find a specific title under the heading of "Electronic Journals." In addition, users may also search for electronic journal titles by broad subject area via the aggregator database title or the collection subset title itself as illustrated by figure 10.

### Third Function of the EJCOP

The third function of the EJCOP is to devise an efficient record maintenance mechanism to ensure the accuracy and currency of information through special treatment of Tag 035.

As the aggregator journal titles are changing all the time, it is vital to devise a simple, efficient, and reliable mechanism for record maintenance. We think the EJCOP has successfully created an algorithm to solve this problem. We use Tag 035 as the basis of repetitive overlay for record maintenance. Because Tag 035 is used for record overlay

**Table 3.** Variations in Tag 856 among Three Unstable Aggregator Databases

| Aggregator Database | Tag 856 | Variable Part (bold and italicized portion) |
|---|---|---|
| Academic Universe | http://www.hkbu.edu.hk/lib-cgi/au.pl?t=33+Metal+Producing | Journal title, space between each word is replaced by a "+" |
| Academic Search FullText Elite | http://www.hkbu.edu.hk/lib-cgi/asfe.pl?t=07470088 | ISSN |
| ProQuest Direct | http://www.hkbu.edu.hk/lib-cgi/pqd.pl?t=19548 | ProQuest Direct Journal ID Number |

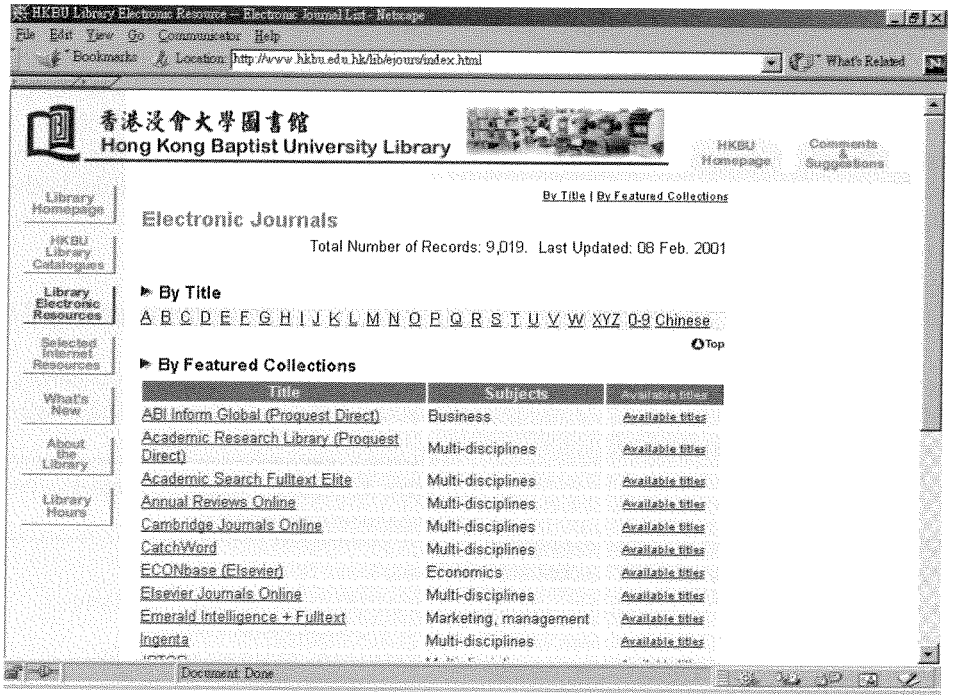



**Figure 10.** Complete Electronic Journal List on HKBU Library Web Site

when uploading MARC records of aggregator databases, this tag must be consistent and unique for each record in the entire record set so that the records can be maintained easily for however long they remain stable. On the other hand, the records for the unstable titles are added or deleted in synchronization with the reissuance of the entire record set by the aggregator database vendor. In order to provide a consistent and unique Tag 035 for each electronic title in an unstable aggregator database for overlay purpose, the EJCOP uses an algorithm that assigns the same value in Tag 035 to the same electronic journal title each time it appears in the vendor-supplied list.

This algorithm contains two calculations to build the relative Tag 035:

1. If ISSN is available. Example: The ISSN for *ABA Banking Journal* is 0194-5947, so the Tag 035 will become "ej0194-5947." As this title is covered by all three aggregator databases, their records will then merge and display as a single record for convenient access by our users as shown in figure 6.
2. If ISSN is not available. The EJCOP creates a unique overlay point by using the title to generate a string algorithmically according to this formula:

   "ej" +
   first 2 characters of each word of the title +
   last 2 characters of the first two words of the title +
   2d- and 3rd-to-last characters of the last word of the title +
   title string length (i.e., total number of all characters and spaces of the title)

For example, the Tag 035 for the journal title *Canadian Government Programs & Services Newsletter* is "ejcagopr&seneanntte50."

In the event that two titles have the same string value, the EJCOP will add a "#" to the end of the second string value and the number of duplicate records identified. For example, the following two titles:

Business Insurance 1995–1996 Directory of Managed Care Providers, and
Business Insurance 1996–1997 Directory of Managed Care Providers,

have the same string value as "ejbuin19diofmacaprssceer64". The EJCOP will automatically add "#1" to the second string value to make it the distinct string "ejbuin19diofmacaprssceer64#1."

Through this algorithmic method of generating matching overlay points, the EJCOP allows the HKBU library to provide an efficient, timely, and accurate ongoing record maintenance mechanism for unstable aggregator databases we subscribe to.

## Advantages Offered by the EJCOP

Using the EJCOP to generate MARC records ensures that the catalog records for the unstable aggregator databases are accurate and current, because they are based on the most recently updated title lists and available record sets supplied by the vendors. While more effort is needed when full record sets are not available, as is the case with Academic Universe, the EJCOP still offers many advantages. The result of the overall process optimizes the use of the title lists and MARC records (if available) as supplied by the aggregator database vendors because the key to the program is its heavy reliance on data conversion as a method of data creation. As a point of reference, the average task time required monthly by the new systematic approach through the EJCOP for Academic Universe, Academic Search FullText Elite, and the two subsets of ProQuest Direct—ABI/INFORM Global and Academic Research Library—is shown in table 4.

Following a monthly update schedule, the time required to create and maintain the unstable electronic journal records by using the EJCOP *annually* is thus estimated to be 22 hours in total (i.e., 110 minutes x 12/60 minutes = 22 hours).

After the most recent update in February 2001, the HKBU Library has 9,019 electronic journal record links (7,868 unique titles) with 87.6% of the links represented by titles in the three unstable aggregator databases—Academic Universe, Academic Search FullText Elite, and ProQuest Direct. Prior to the development of the EJCOP, we had approximately 200 electronic journal titles in our OPAC through a mish-mash of efforts with some records created by our cataloging staff and others by staff responsible for serials in our acquisitions section.

## Limitations of the EJCOP

Catalog records created or converted by the EJCOP for electronic journal titles in unstable aggregator databases are

**Table 4.** EJCOP Average Processing Time for Three Unstable Aggregator Databases

| Tasks | Time (min.) |
|---|---|
| Download records from aggregator databases | 15 |
| EJCOP generates MARC records | 40 |
| Upload MARC record sets to INNOPAC | 40 |
| Create complete electronic title on library homepage | 15 |
| Total | 110 |

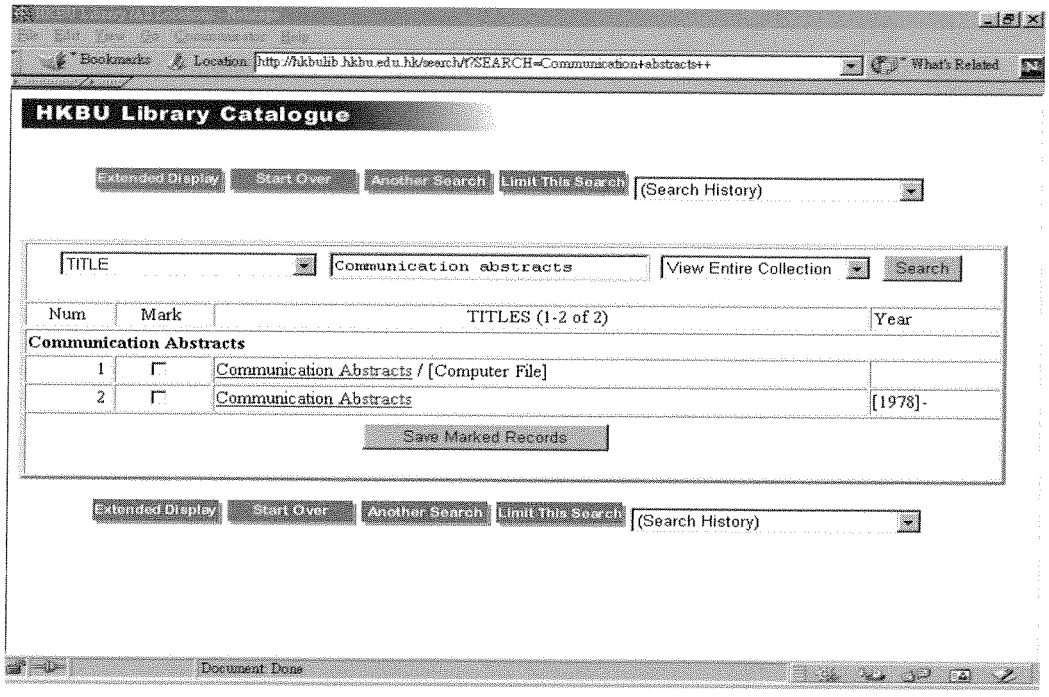



**Figure 11.** Duplicate Electronic Journal Title on HKBU Web OPAC

kept separate from the catalog records for the print versions of the journal titles. For example, there are two catalog records for the journal title *Communication Abstracts* as shown in figure 11. The one with the [Computer File] designation was generated by the EJCOP for a link to the Academic Search FullText Elite record while the second title represents the catalog record of the print version of the journal.

In principle, we can create a special add-on program for the EJCOP to merge the unstable catalog records with those for the print and online version of titles in our OPAC. In doing so, however, we would have to export all the journal catalog records from INNOPAC for processing. In addition to increasing the processing time, there is also the concern about making such frequent changes to stable journal records by a computer program unless there is very strong reason to do so. Given this concern and after taking into consideration that the present rate of extra catalog records created by the EJCOP from the three unstable aggregator databases came to be under 14.4% (702 titles) of English journal titles, we made the decision to keep the records for these aggregator electronic journal titles separate from the records for the print and online version of the titles.

It has also been pointed out to us that the title lists and MARC record sets provided by the unstable aggregator database vendors may not accurately reflect the actual titles in their databases. Therefore, it would be better for library staff to thoroughly check the completeness of the title lists and the record sets. While this may be feasible for very well-staffed libraries, it is impossible for the HKBU Library or other smaller libraries. We decided that we must rely on the aggregator database vendors, especially those involved with unstable aggregator databases, to provide quality service and make ongoing efforts to make improvements as needed.

## Conclusion

After using the EJCOP for slightly more than one year, we are very pleased with the outcome in terms of achieving our goal of having a consistent, accurate, and direct way of providing bibliographic access to the full-text electronic journal titles in unstable aggregator databases. By placing the emphasis on making optimal use of data provided by vendors and minimal dependence on manual processing, we have been able to greatly enhance our journal access and significantly expand the library's collection in an efficient and cost-effective manner. Furthermore, because the functionalities of the EJCOP were designed to apply to generic rather than specific unstable aggregator databases, the utility of the EJCOP goes beyond those databases we currently subscribe to.

We have shared the EJCOP with two other academic libraries in Hong Kong to date, the Lingnan University Library and the City University of Hong Kong.[4] Through their feedback and input, we have made several enhancements to the EJCOP. The first was to allow different libraries the option of using locally assigned field tags for bibliographic record overlay point, coverage information note, and database title added entry as shown in figure 10. The second was to add a new function to the PERL script to keep track of the electronic journal usage statistics by logging the date and time whenever an electronic journal title is clicked by a library user (see figure 12). The data captured could then be reformatted to produce various usage statistics reports.

Lastly, enhancement efforts are underway to deal with one of the current limitations in the EJCOP. We are working on a more cost-effective and systematic algorithm so that two separate records for the same electronic journal title (e.g., one appearing in a stable aggregator database and another in an unstable aggregator database) would be merged into one.

## Works Cited

Calhoun, Karen, and Bill Kara. 2000. "Aggregation or Aggravation? Optimizing Access to Full-Text Journals," ALCTS Newsletter Online 11, no. 4 (2000). Accessed February 27, 2001, www.ala.org/alcts/alcts_news/v11n1/gateway_pap15.html.

Library of Congress Program for Cooperative Cataloging (PCC) Standing Committee on Automation (SCA). 2000. Task Group on Journals in Aggregator Databases, *Final Report.* Accessed February 27, 2001, http://lcweb.loc.gov/catdir/pcc/aggfinal.html.

## Notes

1. The EJCOP was developed by Yiu-On Li, Systems Librarian of Hong Kong Baptist University, for his Master of Science in Computing degree at the City University of Hong Kong. This article is based on his dissertation entitled "A New Systematic Approach to Cataloging Full-Text Electronic Journals: The HKBU Library Experience" (January 2001).
2. Title list of Academic Search FullText Elite is available at www.epnet.com/maglists/maglist.htm. Academic Universe title

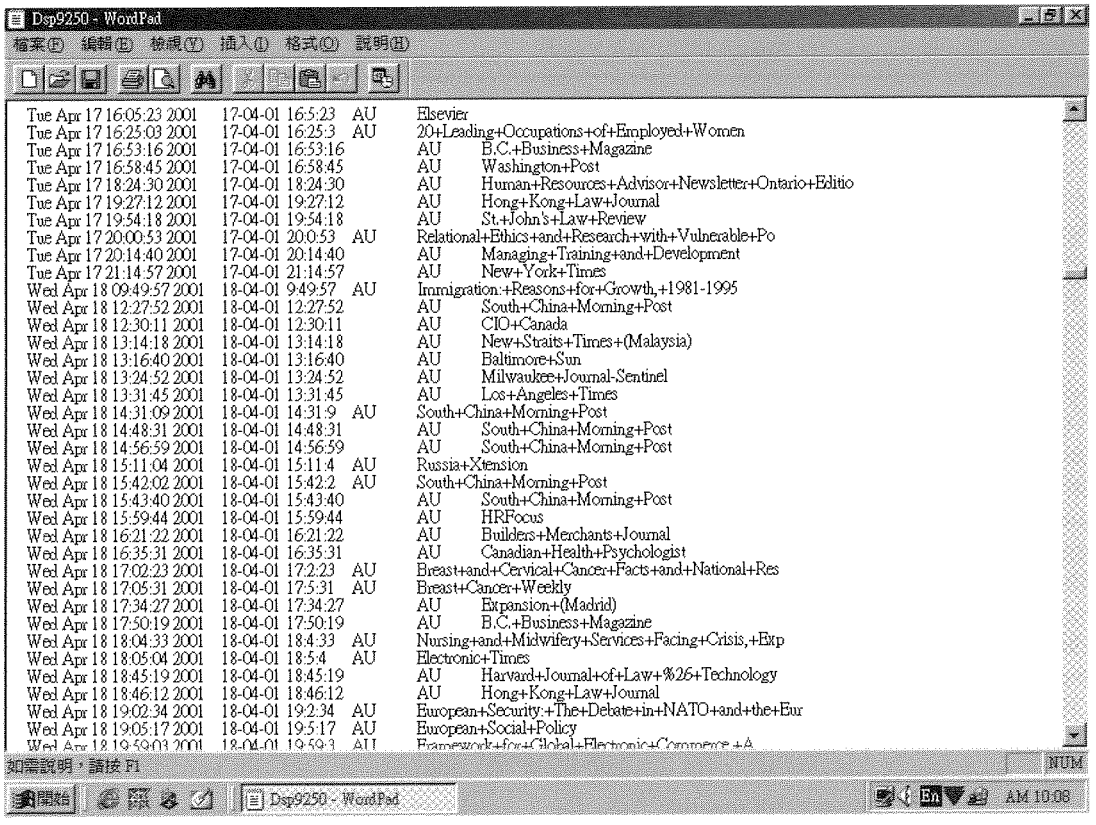



**Figure 12.** Usage Statistics for Academic Universe

list is found at http://cisweb.lexis-nexis.com/marketsource/reports.htm. Title list of ProQuest Direct can be retrieved at www.bellhowell.infolearning.com/hp/Support/Titles. A sample of a ProQuest MARC record is available at www.bellhowell.infolearning.com/hp/Features/Marc/MARCsample.html, and EBSCO's MARC record sets can be retrieved at www.epnet.com/maglists/maglist.htm. Accessed February 27, 2001.

3. We decided to place the coverage information of full-text electronic journal titles in Tag 599 instead of creating and attaching separate holdings or item records to the bibliographic records due to our wish to minimize ongoing maintenance work. Given that titles in unstable aggregator databases are intrinsically unstable, it would be a lot more extra work if we also had to remove the holdings or item records each time we removed obsolete electronic journal titles from our INNOPAC system. However, the EJCOP is a generic program, and it allows different libraries to assign their own local field tag to store the coverage information.
4. The Lingnan University Library reported in June 2001 that they had succeeded in creating 8,188 electronic journal MARC records (including 573 stable electronic journal titles) through the use of the EJCOP and had uploaded them into their INNOPAC system. The City University of Hong Kong Library is in the process of creating the MARC record sets and is planning to upload the records to their INNOPAC system during summer 2001.

This page intentionally blank

# Book Reviews

## Margaret Rohdy, Editor

***The Intellectual Foundation of Information Organization***. By Elaine Svenonius. Cambridge, Mass.: MIT Pr., 2000. 255p. $37 (ISBN 0-262-19433-3) LC99-041301.

A common complaint of bus travelers in London is "You wait for ages for a bus, then two come along together!" Consciously or unwittingly, the library profession has been waiting for a long time for a good book on the overall topic of the organization of information. Now, two have come along within a year—the book under review and Arlene Taylor's *The Organization of Information* (1999). It is, no doubt, due to the exigencies of the publishing process, but it is odd to find no mention of Taylor's book in the text or extensive bibliography of the Svenonius book. Be that as it may, all interested in this topic should read both books and enjoy their complementary virtues.

This is a book of unremitting intellectuality. That is not to say that it is difficult to read—Svenonius writes clearly and well (though is somewhat given to using five-dollar words like "hypostatization" and "disambiguation")—but the reader should understand that there are knotty and abstract issues being discussed here, and beach reading it ain't!

Svenonius begins with a somewhat challenging first chapter that wrestles with a number of definitional and philosophical questions—including my all-time favorite, "What is information?" The book is, after all, on the topic of organizing that elusive concept—or is it? The author, failing to arrive at an entirely satisfactory definition (as have many eminences before her), does make a clear distinction between "information" (however defined) and the embodiments of "information"—the latter being the true subjects of our attentions. When it comes down to it, in doing bibliographic control we have to have the idea of an incorporeal *Hamlet* in our minds, but what we describe, catalogue, and organize are the books, films, sound recordings, and so on, that are manifestations of the work and not the work itself. It is essential that the reader study and understand this first chapter as it lays out the assumptions on which the rest of the book is based. This should be true of any first chapter but, in this wicked world, often is not.

The first five chapters of the book are devoted to an analysis of the intellectual foundations of the organization of information. After the introductory chapter, Svenonius discusses the objectives of bibliographic systems (newspeak for "catalogs") and traces their evolution from Panizzi through Cutter, Lubetzky, and the Paris Principles to IFLA's *Functional Requirements for Bibliographic Records* (FRBR) in 1998. She commends the latter for distinguishing between finding, collocating, choosing, acquiring, and navigating as fundamental purposes of catalogs, and then goes on to discuss how those objectives may be realized in "full-featured bibliographic systems." The next chapter deals with the many types of bibliographic entities—documents, works, superworks, editions, and what the author calls "author sets" and "subject sets." This is a most interesting and comprehensive discussion, though, being of an amathematical mind, I cannot share the delight of others in statements such as "In set-theoretic notation, superwork may be defined as $SW_j$ = def [x: x is derived from $a_{wi}$]" (38). Such formulae aside, this is as good an analysis of these complicated matters as I have seen. The last two chapters of the first part of the book are concerned with bibliographic languages (codes, standards, classification denotations, subject heading lists, etc.) by means of which we create bibliographic records, and the principles that underlie bibliographic descriptions.

The second half of the book builds on the carefully laid intellectual foundations of the first to look at the "languages" used to construct bibliographic records and systems. The author divides these into the languages relating to works (chapter 6), documents (chapter 7), and subjects (chapters 8–10). Beginning this discussion, the author rather confusingly writes, "Work languages describe information entities . . ." (87). This latter term is not defined and following, as it does, the detailed discussion of the notion of work in chapter 3, introduces an unwelcome vagueness. Though "work" is an abstract notion, it is one that is understood at least intuitively by those who study and practice cataloguing. To such, "*Hamlet* is a work" is both understandable and useful, hardly something one could say about "*Hamlet* is an information entity." Leaving that aside, the discussion of vocabulary control, names (personal and corporate), titles (of works), the "main entry" question, and the many different relationships between works,

superworks, and editions is both clarifying and illuminating.

Chapter 7 deals with "document languages"—those formal denotations used to describe carriers of recorded knowledge and information. (It is a small point, but I prefer the logic of the organization of AACR2, which deals with the carriers before the works whose manifestations they embody, to the order of chapters here.) The author discusses what she calls physical (carrier), publication, and access attributes in detail but, again, at a somewhat abstract level, largely divorced from the framework and content standards (MARC, ISBD, etc.) that we use to contain and record those attributes.

The last three chapters deal with "subject languages"—alphabetic and classificatory. The former is sometimes referred to as "natural language" (as opposed to the artificial language of classifications), but this is erroneous because, though words from natural language are used, they are employed in a highly stylized and formal manner. In chapter 8, Svenonius demarcates these languages and discusses vocabulary selection and the semantics and syntax of alphabetic-subject languages. Chapter 9 takes all this into very deep waters with a discussion of relational and referential semantics in constructing alphabetic subject languages (subject headings to you and me). The last chapter deals with the topic of synthesis in subject languages and the application of techniques taken from classification theory. The book concludes with an afterword on the future of its topic, notes on the chapters, and a twenty-page bibliography.

This relatively short book (the text occupies 198 of its 255 pages) is worth the demands it makes on the reader and should be part of the library of anyone with a serious interest in bibliographic control.—*Michael Gorman (michaelg@csufresno.edu), California State University, Fresno*

### Work Cited

Taylor, Arlene G. 1999. *The Organization of Information*. Englewood, Colo.: Libraries Unlimited.

***Managing Electronic Serials: Essays Based on the ALCTS Electronic Serials Institutes, 1997–1999***. Ed. Pamela M. Bluh. ALCTS Papers on Library Technical Services & Collections, no. 9. Chicago: ALA, 2001. 189p. $38 (ISBN 0-8389-3510-9) LC00-67646.

***Developing and Managing Electronic Journal Collections***. By Donnelyn Curtis, Virginia M. Scheschy, and Adolfo R. Tarango. How-To-Do-It Manuals, no. 102. New York: Neal-Schuman, 2000. 267p. $55 (ISBN 1-55570-383-6) LC 00-55423.

***Selecting and Managing Electronic Resources***. By Vicki L. Gregory. How-To-Do-It Manuals, no. 101. New York: Neal-Schuman, 2000. 109p. $55 (ISBN 1-55570-382-8) LC 00-41569.

***Successes and Failures of Digital Libraries***. Eds. Susan Harum and Michael Twidale. 35th Annual Clinic on Library Applications of Data Processing, 1998. Champaign, Ill.: Graduate School of Library and Information Science, Univ. of Illinois, 2000. 134p. $30 (ISBN 0-87845-107-2; ISSN 0069-4789).

***Collection Development in a Digital Environment***. Ed. Sul H. Lee. Binghamton, N.Y.: Haworth, 1999. 118p. $49.95 (ISBN 0-7890-0794-0) $19.95 pbk. (ISBN 0-7890-0827-0) LC 99-39118.

***Electronic Collection Management***. Ed. Suzan D. McGinnis. Binghamton, N.Y.: Haworth, 2000. 159p. $69.95 (ISBN 0-7890-1308-8) $34.95 pbk. (ISBN 0-7890-1309-6) LC 00-53864.

***Collection Development in an Electronic Environment***. Ed. Thomas E. Nisonger. *Library Trends* 48, no. 4 (spring 2000): 639–941. Champaign, Ill.: Univ. of Illinois, 2000. Single issue, $18.50 (ISSN 0024-2594).

***Assessing Digital Library Services***. Ed. Thomas A. Peters. *Library Trends* 49, no. 2 (fall 2000): 221–390. Champaign, Ill.: Graduate School of Library and Information Science, Univ. of Illinois, 2000. Single issue, $18.50 (ISSN 0024-2594).

As the literature about electronic resources proliferates, it is increasingly difficult to examine the large number of new publications to sort out those likely to be of greatest interest and use to *LRTS* readers. As a strategy for dealing with the increased number of publications in a new area of growing interest, the review editor requested a "megareview" covering not one, two, or three titles, but eight recent books with seemingly interchangeable titles, all on the subject of managing electronic resources in libraries.

Careful examination of these titles reveals that four are dual publications, issued simultaneously as monographs and as journal issues, two from Haworth Press and two from the Graduate School of Library and Information Science at the University of Illinois at Urbana–Champaign. One is a set of conference papers bearing both an ISBN and an ISSN. The other three are ordinary monographs, although one, from ALCTS, consists of a set of multiauthored papers that originated as a conference program and subsequently grew into several well-attended regional institutes before publication in book form.

In this review, in order to give each title a fair share of attention and space, first each is described briefly; next they are examined as an integrated body of literature; and finally, they are evaluated in the broader context of available literature on the subject.

*Managing Electronic Serials* includes eleven chapters by different

authors, a bibliography, biographical sketches of the authors, and an index. Most papers include references, and some of the lists are quite extensive. The eleven chapters are "Seriality and the Web," by Regina Romano Reynolds; "Saddlepoints in Serials," by Ronald L. Larsen; "Electronic Serials into the Millennium," by Friedemann A. Weigel; "Managing Electronic Serials: An Overview," by Sharon Cline McKay; "Management and Technical Considerations for Acquiring and Accessing Electronic Serials," by George Machovec; "Seven Common Myths about Acquiring and Accessing E-journals," by Dan Tonkery; "How Intellectual Property Laws Affecting Libraries Are Changing," by Sarah E. Sully; "A License to Kill for . . . ," by Faye A. Chadwell; "Cataloging Electronic Resources: The Practicalities," by Norma J. Fair, Steven C. Shadle, and Beverley Geer; "Nuts and Bolts: Public Service in an Electronic Environment," by Raye Lynn Thomas; and "Another New Frontier: Trailblazing Electronically," by Thomas W. Leonhardt.

The best chapters in *Managing Electronic Serials* are Larsen's, the keynote address at the first regional institute, which provides an excellent overview of the issues to be addressed in collecting, managing, and utilizing electronic serials; and Chadwell's, which furnishes librarians with a guide to obtaining better custom-negotiated licenses instead of passively accepting the standard licenses that vendors offer for their products. As an introduction to electronic serials, this book is an admirable success in furnishing the initial plunge into the area of collecting and managing library resources. Each chapter has something of value for the novice. Each author tries to combine a discussion of principles and underlying theory with practical matters; however some, like Fair et al., are more heavily weighted toward the practical (to be expected given their topic, cataloging e-journals).

*Developing and Managing Electronic Journal Collections*, one of Neal-Schuman's valuable "How-to-Do-It Manuals for Librarians," consists of a preface, which briefly describes the authors' experience with e-journals at the University of Nevada–Reno, and eight chapters: "Introducing Electronic Journals"; "Selection"; "Technical Issues"; "Licensing and Legal Issues"; "The Order Process"; "Cataloging and Access"; "User Services"; and "Predictions for Electronic Journals." The appendixes include a glossary, sample collection policies, a method for comparing database source lists, a method for creating a full-text master list, and an electronic journal spreadsheet. Throughout the book, illustrations and sample screens augment the clearly written text. Theoretical principles are sometimes mentioned in passing, but this book is aimed at helping its readers do something concrete, thus its emphasis on systematic action. Each chapter opens with an overview and list of topics to be covered and, after the body of the chapter deals with each topic in turn, closes with a bibliography of additional sources to consult. Nothing is treated in depth, but everything is highly organized.

*Selecting and Managing Electronic Resources*, another of Neal-Schuman's well-written, carefully documented how-to manuals, approaches collection development for electronic resources from a more general viewpoint. Its chapters include "Collection Development Policies"; "Selection: Criteria and the Selection Process"; "Budgeting and Acquisitions"; "Organization and Access to Electronic Resources"; "Evaluation and Assessment"; "Copyright and Licensing Issues"; and "Preservation Issues." I assumed this manual would focus on monographs, complementing Curtis et al.'s manual on electronic journals from the same series, but this is not the case. Some material in the two manuals overlaps, but *Selecting and Managing Electronic Resources* provides a context in which the greater detail about serials in Curtis et al. can be understood more readily.

*Successes and Failures of Digital Libraries* consists of an introduction by the editors and eleven chapters: "NSF/DARPA/NASA Digital Libraries Initiative: A Program Manager's Perspective," by Stephen M. Griffin; "Update on the Networked Digital Library of Theses and Dissertations," by Edward A. Fox; "Processing and Access Issues for Full-Text Journals," by William H. Mischo and Timothy W. Cole; "Federated Search of Scientific Literature: A Retrospective on the Illinois Digital Library Project," by Bruce Schatz et al.; "From Usability to Use: Measuring Success of Testbeds in the Real World," by Laura J. Neumann and Ann Peterson Bishop; "Semantic Issues for Digital Libraries," by Hsinchun Chen; "Technology Transfer in a Public University," by Robert Wedgeworth; "A Publisher's Perspective: More Successes than Failures," by Tim Ingoldsby; "Lessons Learned from Full-Text Journals at OCLC," by Thomas B. Hickey; "The Future of Annotation in a Digital (Paper) World," by Catherine C. Marshall; and "Give Me Documents or Give Me Death: A Millennial Meditation on Documents and Libraries," by David M. Levy.

The focus here is on the processes of digital library development in a multilibrary context, not on acquiring titles for individual collections. Aiming to discuss not only successes but also failures, the authors look closely at many aspects of the Digital Libraries Initiative, the largest such project based at the University of Illinois at Urbana–Champaign, and far less deeply at other projects, such as OCLC's. After a close reading, one concludes that many issues and problems had to be addressed, but there were no failures. At most, author

Ingoldsby recognizes some disappointments (89).

Most chapters contain notes and references to sources, but two have none. The length of chapters varies from a few pages to twenty or more. Some are heavily illustrated; others not at all. Some are extremely detailed; others sketch only broad brushstrokes. Some are extraordinarily well written. In journal issues, such unevenness is expected; in a monograph, it draws criticism.

The papers in *Collection Development in a Digital Environment,* originally presented at the 1998 University of Oklahoma Libraries Conference, include: "Chaos Breeds Life: Finding Opportunities for Library Advancement during a Period of Collection Schizophrenia," by James G. Neal; "By the Dawn's Early Light," by Milton T. Wolf; "Collection Development and Professional Ethics," by Kenneth Frazier; "Making the Wild Wind Visible: Information Technology in a Brave New World," by Dennis Dillon; "Building the Global Collection—World Class Collection Development: A Chronicle of the AAU/ARL Global Resources Program," by Mary Case and Deborah Jakubs; "Scenario Planning and Collection Development," by Joan Giesecke; "Role of the Aggregator in the Emerging Electronic Environment," by Kathleen Born; and "Digital Information and the Subscription Agent as Information Coach," by Kit Kennedy.

The authors have produced well-written papers of similar length—all but two have references, and some are illustrated. One chapter worthy of special mention is Kenneth Frazier's piece on professional ethics in the electronic environment. In it, he explores two cases in which Gordon & Breach and Elsevier succeeded in using the law in ways that scholars and librarians might well find unethical. He also discusses "[t]he ongoing effort by publishers to use licensing contracts to abolish interlibrary loan services for information in electronic formats"—services that librarians are struggling to preserve (34). Very little that he has to say is good news for libraries. Still, when the potential for large profits is not involved, glimmers of sunlight filter in, as with the ambitious global collection development project of the Association of American Universities and the Association of Research Libraries.

The overall structure of *Electronic Collection Management* is almost identical to *Collection Development in a Digital Environment.* All but one of the eight papers are similar in length—Baldwin's paper on collection development in the new millennium is nearly twice the length of the others. All have notes and references; some have illustrations. The chapters are "Cultivating Our Garden: The Impact of Digital Full Text Periodicals on the Liberal Arts College Library," by Barbara Doyle-Wilch and Carla Tracy; "Information Technology and Collection Development Departments in the Academic Library," by David C. Fowler; "SPARC: An Alternative Lifestyle for Academic Libraries," by Marina Oliver; "Challenges in Electronic Collection Building in Interdisciplinary Studies," by Kristin H. Gerhard; "Collection Development in the New Millennium—Evaluating, Selecting, Annotating, Organizing for Ease of Access, Reevaluating, and Updating Electronic Resources," by Virginia Baldwin; "Collection Development and Organization of Electronic Resources," by Gerald L. Newman; "Distance Learning and the Opportunities and Challenges for Libraries," by William J. Gibbs; and "Some Issues for Collection Developers and Content Managers," by Thomas Peters. What is most interesting about this volume is the absence of focused discussions on legal issues—contracts, licensing, negotiation—or the role of vendors. Though practical matters are covered, as in the description of SPARC, for the most part these authors take a step back to view the overall landscape of collection development.

In Nisonger's book, *Collection Development in an Electronic Environment,* the chapters are "Electronic Resources and Academic Libraries, 1980–2000: A Historical Perspective," by Ruth H. Miller; "Are We There Yet? Online E-resources Ten Years After," by Ann Okerson; "Electronic Resources: Are Basic Criteria for the Selection of Materials Changing?" by Curt Holleman; "Principles of Selection for Electronic Resources," by Paul Metz; "Love's Labour's Lost: The Failure of Traditional Selection Practice in the Acquisition of Humanities Electronic Texts," by Beau David Case; "Reviewing the World Wide Web—Theory versus Reality," by James H. Sweetland; "Collecting Full-Text CD-ROMs in Literature: Theory, Format, and Selection," by Roberta Astroff; "Selecting Research Collections for Digitization: Applying the Harvard Model," by Kristine R. Brancolini; "The Impact of Electronic Resources on Serial Cancellations and Remote Storage Decisions in Academic Research Libraries," by Janice M. Jaguszewski and Laura K. Probst; "Twilight of the Gods? Bibliographers in the Electronic Age," by Dan Hazen; "Impact of Electronic Resources on Collection Development, the Roles of Librarians, and Library Consortia," by Glenda A. Thornton; "Hybridity, Mutability, Multiplicity: Theorizing Electronic Library Collections," by Marlene Manoff; "The Acquisitions Culture Wars," by Edward Shreeves; and "Navigating the Parallel Universe: Education for Collection Management in the Electronic Age," by Virgil L. P. Blake and Thomas T. Surprenant.

Many, though not all, of these chapters are longer than those in the other multiauthored titles described here, running twenty pages or more.

Each begins with an abstract and closes with references, most of which are long lists of sources. The emphasis here is on selection in principle and practice, often approached in general, but sometimes with regard to specific subjects or media. Miller's opening chapter is the kind of broadly conceived, carefully prepared historical survey for which *Library Trends* is well known and highly prized. Metz's paper gives lists of both traditional and newer criteria for selecting electronic materials that will be of great practical value to readers. Hazen's analysis of the role of bibliographers should assuage fears that they are becoming passé. Blake and Surprenant's closing paper discusses the important question of how educators should teach collection development in this brave new world of electronic resources.

The final volume in this array is *Assessing Digital Library Services.* Chapters include "Evaluating Digital Libraries for Teaching and Learning in Undergraduate Education: A Case Study of . . . ADEPT," by Chistine L. Borgman et al; "Unobtrusive Data Analysis of Digital Reference Questions and Service at the Internet Public Library," by David S. Carter and Joseph Janes; "Bundles in the Wild: Managing Information to Solve Problems and Maintain Situation Awareness," by Paul Gorman et al; "Digital Libraries and Their Challenges," by Daniel Greenstein; "Evaluating Digital Libraries: A Longitudinal and Multifaceted View," by Guy Marchionini; "Current Opportunities for the Effective Meta-Assessment of Online Reference Services," by Thomas A. Peters; "Digital Library Evaluation: Toward Evolution of Concepts," by Tefko Saracevic; and "Project Ethnography: An Anthropological Approach," by Michael Seadle.

The emphasis here is on assessment of digital library services; nevertheless, some articles focus on specific projects (e.g., Borgman et al., Carter and Janes, Marchionini, Seadle) while others furnish generalized models and methodologies (Greenstein, Peters, Saracevic). Gorman et al.'s interesting article examines how "bundle building," developing highly focused, well-organized minicollections for specific purposes, aids problem-solving. The articles are well written and include references and some illustrations. The issue offers a good balance of principles and practice on a topic that is often given less attention than it deserves.

What kind of body of literature do these eight items represent? First, there is less overlap than this reviewer anticipated, although there is some. Several require paying high prices for small amounts of content, but the fact is even the most costly item in terms of price-per-page includes important authors' writing on topics librarians should read about. Each offers a knowledgeable foray into an enormous topical area. Even where there is overlap, the approaches and treatments of different authors vary widely. For example, the chapter on cataloging by Fair et al. in *Managing Electronic Serials* couldn't be more different than its counterpart in *Selecting and Managing Electronic Resources*, in which Gregory lays out the possibilities, from no cataloging at all to providing full MARC records for each electronic item collected. Fair et al., after a quick obeisance to Cutter's "Objects," explain how to catalog electronic serials, giving field-by-field instructions.

The prices of these eight books vary widely, from $18.50 for Nisonger's 300-plus-page *Library Trends* issue to $55 for Gregory's monograph of slightly more than one hundred pages. Clearly, the length of an item does not govern its price, nor does the pricing appear to depend on monographic versus serial publication. For all the collection developers' complaints about price gouging by serials publishers and the outrageous prices of subscriptions, an $85 subscription to *Library Trends* offers the best value based on the amount of material on the subject of electronic resources.

Are these the best books available on the topics they cover? Yes and no. Yes, because they are excellent books, and no amount of quibbling over individual chapters or publication formats in any one of them diminishes that. No, because other books cover the individual subtopics contained within them in greater detail—for example, if one wants to read about cataloging Internet resources, several excellent manuals give better and more thorough instruction, and *The Journal of Internet Cataloging* offers generalized discussions of problems in each issue. But if one wants only to read a well-written overview of cataloging issues or the kind of instruction one can absorb in under an hour, these can meet one's need. What one discovers, however, is that Gregory does not give even a smattering of instruction in how to catalog an Internet journal despite its calling itself a "how-to" while Fair et al. do almost nothing to look at overall organizational issues despite the presence of their chapter on cataloging in a "theory plus practice" volume. So, a careful reader still needs to pay attention not only to the topic or subtopic being covered, but also how it is treated by its author. All of these eight books have good content, important authors, and authoritative publishers, however some offer more than others. Nisonger's *Library Trends* issue is the most "bang for the buck," and its subject—selection—is a critical need for librarians seeking to learn more about collecting electronic resources. Similarly, evaluation of those resources—Peters' subject—is sorely needed in the real world. Those two should get high priority on everyone's professional reading list.

The best introductory treatment of electronic resources is *Managing Electronic Serials*, edited by Pamela M. Bluh. It is not geared toward specific activities, but it is eminently read-

able, with topics treated in depth. Its multiauthor format is an advantage in furnishing a wide range of opinions on a great variety of topics, which single or joint authors writing on focused topics are hard-pressed to match. For practical help in beginning the acquisition of electronic resources, *Selecting and Managing Electronic Resources,* by Gregory, and *Developing and Managing Electronic Journal Collections,* by Curtis et al., are the best choices. If you can buy or read only one, then Curtis et al. is this reviewer's preference. While Gregory is more generic, and Curtis is confined to journals, the latter covers its topics a little more fully. Harum and Twidale's, Lee's, and McGinnis's books are good if they cover a topic in which readers have specific interest, for example, interdisciplinary studies (McGinnis), scenario planning (Lee), or the Digital Libraries Initiative (Harum and Twidale). Having said this, a caution is in order: In the rapidly evolving area of electronic publications and digital libraries, none of this content is truly current; the material in these eight books should be seen as contributing to readers' background and general knowledge. Printed books do not reveal late-breaking developments.—*Sheila S. Intner (intner@simmons.edu), Simmons College Graduate School of Library and Information Science, Boston, Massachusetts, and GSLIS at Mt. Holyoke College, South Hadley, Massachusetts.*

This page intentionally blank

# Instructions for Authors

## Manuscript Submission

Manuscripts of articles should be sent to the editor, John Budd, School of Info Science & Learning Tech., 221M Townsend Hall, Columbia MO 65211; (573) 882-3258; fax: (573) 884-4944; e-mail: buddj@missouri.edu.

In general, the editorial staff follows the Guidelines for Authors, Editors, and Publishers of Literature in the Library and Information Field adopted by the American Library Association Council in 1983 and available from the ALA Executive Offices. Information about copyright policies also is available from ALA headquarters.

## Manuscript Preparation

Please follow these procedures for preparing manuscripts for *Library Resources & Technical Services (LRTS)*:

1. Submit original, unpublished manuscripts only. Do not submit manuscripts that are being considered for publication in other venues. Authors are responsible for the accuracy of statements included. Papers presented at a conference should be identified with the conference name and date in the cover letter.
2. Manuscripts should be machine-printed and double-spaced. Three copies must be provided. Disk copy will be requested from authors for accepted articles.
3. Write the article in a grammatically correct, simple, readable style. Whenever possible avoid jargon, anthropomorphism, and acronyms. All acronyms must be accompanied by their full spelled-out form. For spelling and usage consult the *Random House Webster's College Dictionary* (New York: Random House, 1991). Verify the spelling and accuracy of all names in an appropriate source. Consult *The Chicago Manual of Style* 14th ed. (Chicago: Univ. of Chicago Pr., 1993) for capitalization, abbreviations, usage of numbers, etc.
4. Give the article a brief title; if the title does not fully describe the content of the article, add a brief subtitle. On the first page of the manuscript give the article title, the name(s) of the author(s), and the position title, institutional affiliation, and address of each author.
5. On the second page of the manuscript give the title followed by a brief, informative abstract. Do not identify the author(s) here or elsewhere in the manuscript. Number all pages throughout the manuscript.
6. Submit all references on separate pages at the end of the text, preceding any tables or illustrations.
7. *LRTS* follows *The Chicago Manual of Style* author-date system of references (see chapter 16). Verify each citation by sight, very carefully.
8. Follow the examples and suggestions in chapter 12 of *The Chicago Manual of Style* in designing tables. Submit each table on a separate page at the end of the manuscript. Indicate the preferred placement in the text with an instruction in square brackets. Provide each table with a brief, meaningful caption.
9. Be prepared to supply camera-ready copy for all illustrations. Accompany the manuscript with a photocopy of each, and a brief, meaningful caption noted on the verso.

## Editorial Policy

*Library Resources & Technical Services (LRTS)* is the official journal of the Association for Library Collections & Technical Services (ALCTS), a division of the American Library Association. The following statement of editorial policy was adopted by the ALCTS Board of Directors, April 1998.

### Purpose

The purpose of *LRTS* is to support the theoretical, intellectual, practical, and scholarly aspects of the profession of collection management and development, acquisitions, cataloging and classification, preservation and reformatting, and serials, by publishing articles (subject to double-blind peer review) and book reviews, and editorials and correspondence in response to the same.

### Audience

The audience for *LRTS* includes practitioners, students, researchers, and other scholars with an interest in collection development and technical services and related activities in all types of libraries.

### Frequency

*LRTS* is published quarterly, with the volume calendar corresponding to the calendar year. Numbers appear in January, April, July, and October.

### Scope

The editor of *LRTS*, with the assistance of an editorial board, strives to achieve a balance among the articles published in the journal so that the interests of each of the sections of ALCTS (Acquisitions, Cataloging and Classification, Collection Management and Development, Preservation and Reformatting, Serials) is represented in the journal. Articles on technology, management, and education, e.g., are appropriate to the journal when the application of these is to issues of interest to practitioners and researchers working in collection development and technical services. The scope of the articles published in *LRTS* is also guided by the Mission and Priorities Statement adopted by the ALCTS Board of Directors in 1990.

### Content

The content of *LRTS* is to include:

1. Articles that further the advancement of knowledge by reporting the results of research or other scholarly activity.
2. Periodic literature review essays that discuss issues and trends.
3. Notes that report unique or evolving technical processes.
4. Notes that report unique or evolving research methods.
5. Substantive book reviews of new publications.
6. A brief, factual, annual statement of the association's accomplishments.

*LRTS* is not an appropriate forum for brief reports on new products, new services, or other current news items.

# Index

## Volume 45, 2001

### Compiled by Edward Swanson

General Procedures Used in Compiling the Index

The following types of entries are included:

a. authors—of articles, reviews, and letters
b. titles—of articles and of articles about which letters were published
c. subjects—of articles and of books reviewed

Subject entries for individuals are identified by "(*about*)"; letters are identified by "(*c*)".

Reviews are indexed by name of reviewer and by subject of the work reviewed, identified by "(*r*)". They are also listed by title under the heading "Books reviewed".

Entries are arranged word by word following the "file-as-spelled" principle. Numbers are arranged before alphabetical characters; acronyms without internal punctuation are arranged as words.

Subject headings are based on: *ASIS Thesaurus of Information Science and Librarianship,* edited by Jessia L. Milstead (Medford, N.J.: Published for the American Society for Information Science by Learned Information, Inc., 1994).

Paging of Volume 45:
Pages 1–68 = Number 1 (January)
Pages 69–112 = Number 2 (April)
Pages 113–184 = Number 3 (July)
Pages 185–224 = Number 4 (October)


**A**

AACR 2
  revision of: 48–51
  seriality in: 49
Academic Book Center (Portland, Ore.): 104–9
academic libraries
  acquisitions in: 127–77
acquisition of library materials
  book reviews in: 127–77
aggregator databases
  cataloging of: 198–211
American Library Association, Committee on Cataloging: Description and Access: 49–51
*Anglo-American Cataloguing Rules,* Second edition, *see* AACR 2
approval plans: 127–77
authority control
  use of World Wide Web in: 73–79

**B**

Behrend, Linda: 180–81 (*r*)
bibliographic databases
  vendor records in: 10–19, 123–26
bibliographic displays
  design of: 20–46
bibliographic records: 178–79 (*r*)
  sources of: 10–19, 123–26
Black, Steve: 3–9
book reviews
  in acquistions: 127–77
"Book Reviews" 59–64, 110–12, 178–84, 213–18
"Book Vendor Records in the OCLC Database: Boon or Bane?" 10–19
Books reviewed
  *Assessing Digital Library Services* (Peters, ed.): 214–18
  *Collection Development in a Digital Environment* (Lee, ed.): 214–18
  *Collection Development in an Electronic Environment* (Nisonger, ed.): 214–18
  *Copyright Essentials for Librarians and Educators* (Crews): 110–11
  *Developing and Managing Electronic Journal Collections* (Curtis, Scheschy, and Tarango): 214–18
  *Digital Imaging: A Practical Handbook* (Lee): 179–80
  *Electronic Collection Management* (McGinnis, ed.): 214–18
  *FRBR Seminar: Functional Requirements for Bibliographic Records=Requisiti Funzionali per Record Bibliografici, Florence, 27–28 January 2000, Proceedings* (Guerrini, ed.): 178–79
  *From Carnegie to Internet2: Forging the Serials Future: Proceedings of the North American Serials Interest Group, Inc: 14th Annual Conference, June 10–13, 1999* (Fiander, Harmon, and Makepeace, eds.): 59–60
  *From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World* (Borgman): 60–61
  *The Future of Cataloging: Insights from the Lubetzky Symposium, April 18, 1998, University of*

*California, Los Angeles* (Connell and Maxwell, eds.): 61–62
*Handbook for Digital Projects: A Management Tool for Preservation and Access* (Sitts, ed.): 111–12
*The Intellectual Foundation of Information Organization* (Svenonius): 213–14
*Managing Electronic Serials: Essays Based on the ALCTS Electronic Serials Institutes, 1997–1999* (Bluh, ed.): 214–18
*Moving Theory into Practice: Digital Imaging for Libraries and Archives* (Kenney and Rieger): 179–80
*Organizing Audiovisual and Electronic Resources for Access: A Cataloging Guide* (Hsieh-Yee): 180–81
*Progress in Visual Information Access and Retrieval* (Sandore, ed.): 62–63
*Saving the Time of the Library User Through Subject Access Innovation: Papers in Honor of Pauline Atherton Cochrane* (Wheeler, ed.): 181–84
*Selecting and Managing Electronic Resources* (Gregory): 214–18
*The Social Life of Information* (Brown and Duguid): 63–64
*Successes and Failures of Digital Libraries* (Harum and Twidale, eds.): 214–18
Budd, John: 71
Burk, Bartley A.: 10–19

C

Calhoun, John C.: 127–77
California State University System: 127–77
Carrico, Steven: 104–9, 111–12 (*r*), 179–80 (*r*)
catalogers
education of: 52–53
cataloging: 61–62 (*r*)
bibliography: 54–58
of electronic journals: 198–211
history: 61–62 (*r*)
philosophy: 61–62 (*r*)
year's work: 47–58
CC:DA, *see* American Library Association, Committee on Cataloging: Description and Access
Chan, Lois Mai: 187–97
Childress, Eric: 187–97
*Choice:* 127–77
citation analysis: 3–9
classification
difference as concept in: 115–22
philosophy of: 115–22
sameness as concept in: 115–22
Cochrane, Pauline Atherton: 181–84 (*about*) (*r*)
communication disorders
core literature: 3–9
"Computer Cataloging of Electronic Journals in Unstable Aggregator Databases: The Hong Kong Baptist University Library Experience" 198–211
continuing resources
cataloging rules: 49
copyright law: 110–11 (*r*)
core literature collections
analysis: 3–9, 127–77
Crump, Michele: 104–9

D

Darby, Sharla: 95–103
Dean, Rebecca: 187–97
Delsey, Tom: 178–79 (*r*)
Dewey Decimal Classification: 115–22
difference
as concept in classification: 115–22
digital libraries: 60–61 (*r*), 214–18 (*r*)
digitization: 179–80 (*r*), 111–12 (*r*)
digitized images: 179–80 (*r*)
digitized library materials: 179–80 (*r*)
document delivery
unmediated: 80–89

E

Edwardy, Jeffrey: 90–94
"The Effect of Interface Design on Item Selection in an Online Catalog" 20–46
electronic journals
cataloging of: 198–211, 214–18 (*r*)
collection development: 214–18 (*r*)
electronic resources
cataloging of: 180–81(*r*), 214–18 (*r*)
collection development: 214–18 (*r*)
El-Hoshy, Lynn M.: 187–97

F

"Form Subdivisions: Their Identification and Use in LCSH" 187–97
FRBR, *see* Functional Requirements for Bibliographic Records
"From the Editor" 71
Functional Requirements for Bibliographic Records: 178–79 (*r*)

G

Gorman, Michael: 63–64 (*r*), 213–14 (*r*)
Gregory, Vicki L.: 110–11 (*r*)

H

Haslam, Michaelyn: 80–89
Hong Kong Baptist University: 198–211

I

information: 63–64 (*r*)
information organization, *see* organization of information
"Instruction for Authors" 65–66, 220–21
"Instructions for Book Reviewers" 67–68
interfaces
in OPACs: 20–46
Intner, Sheila S.: 60–61 (*r*), 61–62 (*r*), 214–18 (*r*)
"It's Academic: Shelf-Ready Standing Orders at the University of Florida's Smathers Library" 104–9
Italian language materials
vendor records for: 123–26

K

Kellsey, Charlene: 123–26

L

Layne, Sara Shatford: 62–63 (*r*)
"Letter to the Editor" 71–72
Leung, Shirley W.: 198–211
Li, Yiu-On: 198–211
*Library Journal:* 127–77
library materials
shelving of, *see* shelving of library materials
Library of Congress Subject Headings
form subdivisions in: 187–97
"Library-Subsidized Unmediated Document Delivery" 80–89
Lubetzky, Seymour: 61–62 (*about*) (*r*)

M

metadata: 51–52
modifiers, *see also* subject headings—form subdivisions in
"Monitoring Book Reshelving in Libraries Using Statistical Sampling and Control Charts" 90–94

N

NACO Project: 73–79
Name Authority Cooperative Project, *see* NACO Project
name authority records: 73–79
nonprint materials
  cataloging of: 180–81 (*r*)

O

O'Neill, Edward T.: 187–97
OCLC database
  vendor records in: 10–19, 123–26
Olson, Hope A.: 115–22
online public access catalogs, *see* OPACs
OPACs
  displays in, design of: 20–46
organization of information
  philosophy: 213–14 (*r*)

P

Persing, Bob: 59–60 (*r*)
physical processing of library materials: 104–9
Pontius, Jeffrey: 90–94
preservation of library materials
  digitization in: 111–12 (*r*)
  workshops: 95–103
"Preservation Workshop Evaluation" 95–103
publishers
  in relation to approval plans: 127–77

R

relevance
  in relation to OPAC displays: 20–46
retrieval effectiveness
  in relation to OPAC displays: 20–46
"Reviews, Holdings, and Presses and Publishers in Academic Library Book Acquisitions" 127–77
Rohdy, Margaret: 59–64, 110–12, 178–84, 213–18
Russell, Beth M.: 73–79

S

sameness
  as concept in classification: 115–22
"Sameness and Difference: A Cultural Foundation of Classification" 115–22
serial publications: 59–60, (*r*) *see also* electronic journals
  analysis: 3–9
Shedenhelm, Laura D.: 10–19
shelf order: 90–94
shelf-ready materials: 104–9
shelving of library materials
  statistical analysis: 90–94
SOLINET, Preservation Services: 95–103
Southeastern Library Network, *see* SOLINET
Spanish language materials
  vendor records for: 10–19
Spillane, Jody Lynn: 73–79
standing orders
  physical processing of: 104–9
Stowers, Eva: 80–89
subdivisions in subject headings, *see* subject headings—form subdivisions in
subject access: 181–84 (*r*)
subject cataloging
  year's work in: 53
subject headings
  form subdivisions in: 187–97
Swanson, Edward: 222–24

T

Thomas, David H.: 20–46
"Trends in Source of Catalog Records for European Monographs, 1996–2000: A Preliminary Study of Italian Monographs" 123–26

U

"Unicode Standard: Its Scope, Design Principles, and Prospects for International Cataloging" (July 2000): 71–72 (*c*)
University of Florida: 104–9
University of Nevada, Las Vegas: 80–89
unmediated document delivery, *see* document delivery—unmediated
"Using Citation Analysis to Pursue a Core Collection of Journals for Communication Disorders" 3–9
"Using the Web for Name Authority Work" 73–79

V

vendor records
  in bibliographic databases: 10–19, 123–26
visual materials
  cataloging of: 62–63 (*r*)
Vizine-Goetz, Diane: 187–97

W

Weinberg, Bella Hass: 71–72 (*c*)
Wiseman, Christine S.: 95–103
workshops
  evaluation: 95–103
World Wide Web
  use in authority control: 73–79

Y

Yankee Book Peddler: 104–9
"The Year's Work in Cataloging, 1999" 47–58
Young, J. Bradford: 181–84 (*r*)