

# 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).<sup>1</sup>

### The HKBU Library

The HKBU Library ([www.hkbu.edu.hk/lib](http://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

	A	B	C	D	E	F	G
	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	Cheaver 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 Scandina	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	Cartax Publishing	01/01/94 to present	01/01/94 to present	3	Yes
15	1355-6215	Addiction Biology	Cartax 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	0324-3046	Advanced Composite Meta	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		

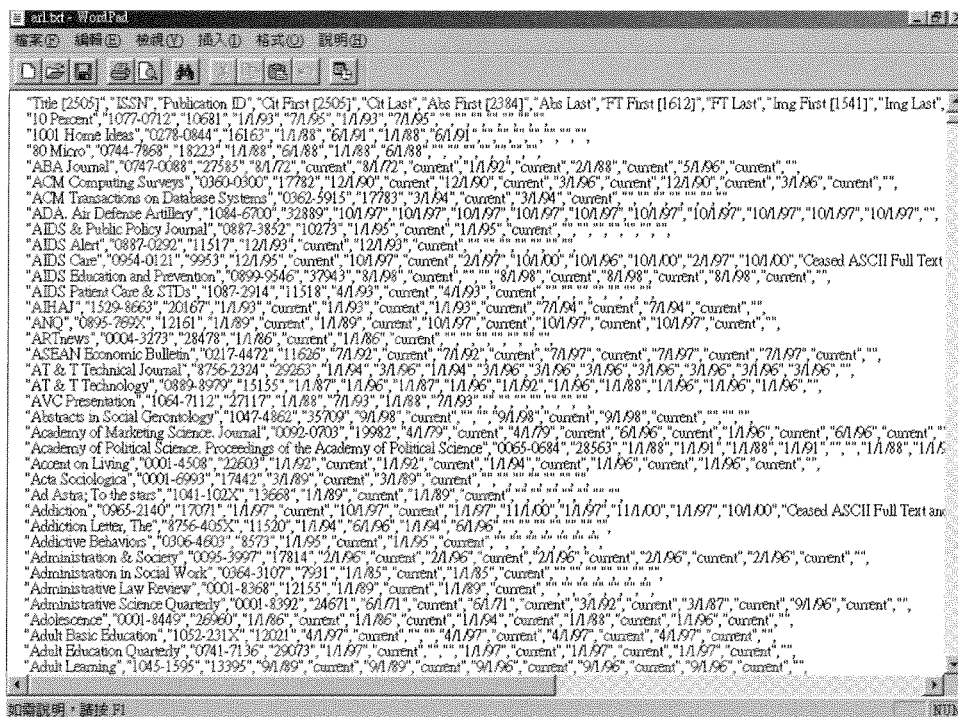
Note: This is stored in an Excel file.

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
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 RFM 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 is
Abortion Report		Full-text	From July 19, 1989 through August 15, 1997
Alya 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	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	Content 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	1042-9233	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.<sup>2</sup> 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,<sup>3</sup> 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

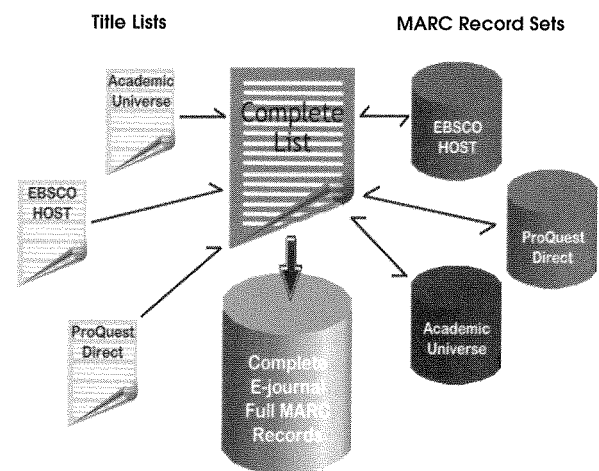
**Table 1.** Duplicate Titles in Three Aggregator Databases

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

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-

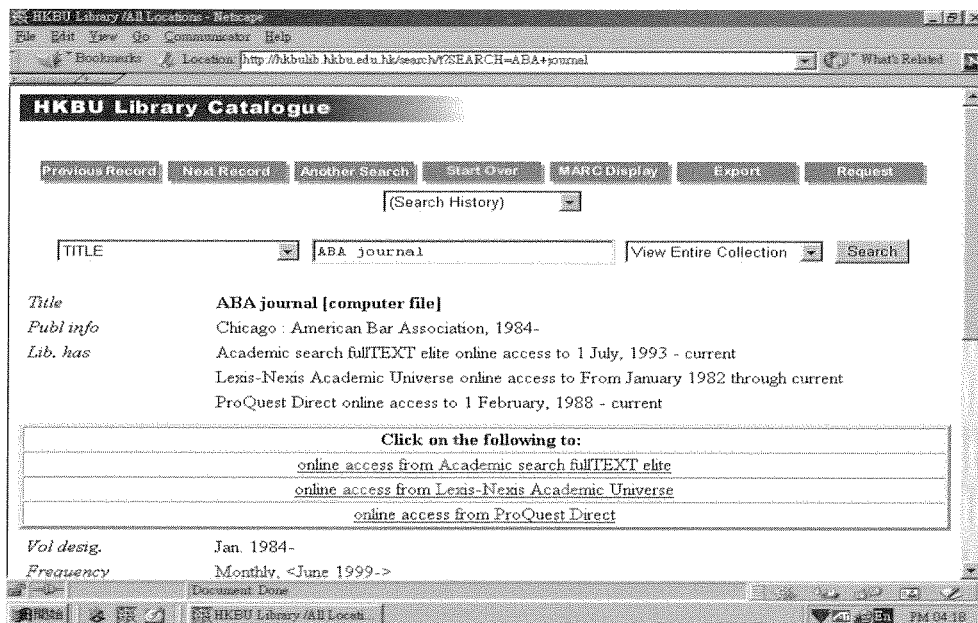


Figure 6. An Electronic Journal Title with Merged Full MARC Records in Public Display Format

```

001  hkbu_ejour01/13/01
006  m      d  |||||
007  cr anu
008  840106c19849999ilumxlp  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



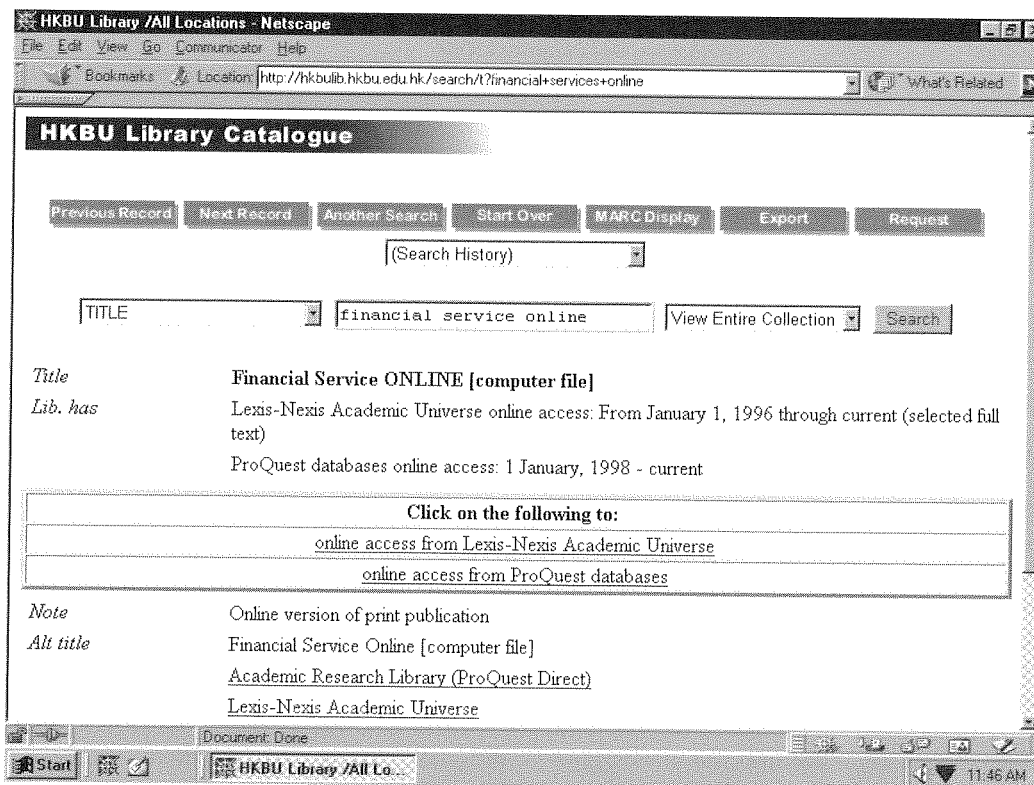


Figure 8. An Electronic Journal Title with Merged Brief MARC Records in Public Display Format

```

001  hkbu_ejour01/13/01
006  m      d  |||||
007  cr anu
008
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/pgd.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=ca  
d47e44-14ef-11d4-b2b3-8a0c585aaa77.1.  
3133490923.73209.+0.0&_state=&wchp=  
dGLkbb-lS1Al&_md5=a98aa4fef57  
ddd1310d948dad2a1eaf5&product=universe&  
unix=http%3A%2F%2Fweb.lexis-nexis.com  
%2F universe&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/ejourns/index.html](http://www.hkbu.edu.hk/lib/ejourns/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.

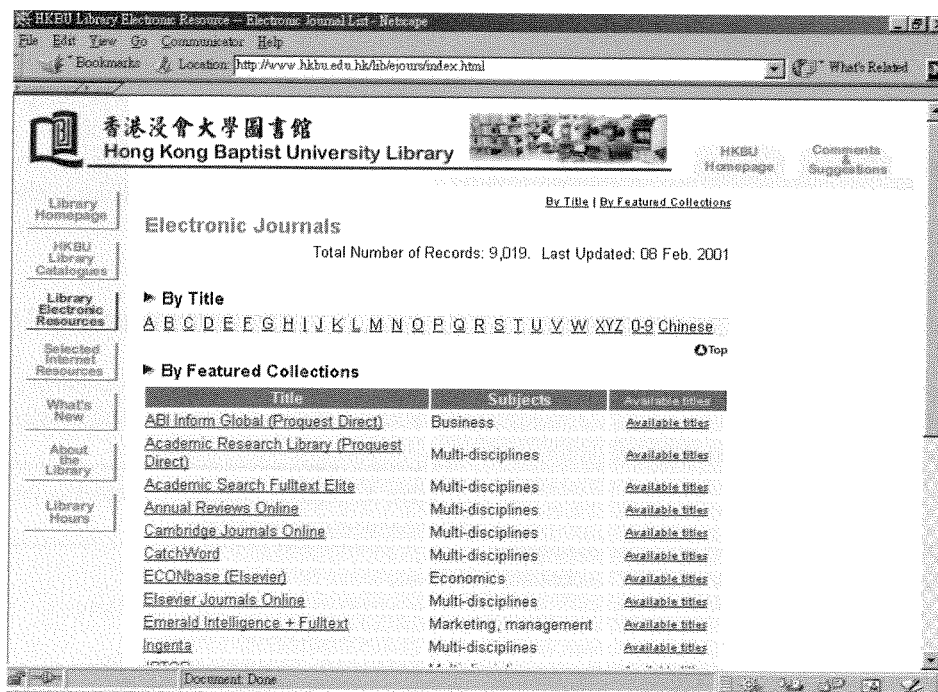
### Third Function of the EJCOP

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	<a href="http://www.hkbu.edu.hk/lib/cgi/au.pl?t=33+Metal+Producing">http://www.hkbu.edu.hk/lib/cgi/au.pl?t=33+Metal+Producing</a>	Journal title, space between each word is replaced by a "+"
Academic Search FullText Elite	<a href="http://www.hkbu.edu.hk/lib/cgi/asfe.pl?t=07470088">http://www.hkbu.edu.hk/lib/cgi/asfe.pl?t=07470088</a>	ISSN
ProQuest Direct	<a href="http://www.hkbu.edu.hk/lib/cgi/pqd.pl?t=19548">http://www.hkbu.edu.hk/lib/cgi/pqd.pl?t=19548</a>	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 charac-  
 ters 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 "ejbuin19diofmacaprssc-  
 ceer64". The EJCOP will automatically add "#1" to the sec-  
 ond string value to make it the distinct string  
 "ejbuin19diofmacaprssc-  
 ceer64#1."

Through this algorithmic method of generating match-  
 ing 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 min-  
 utes = 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

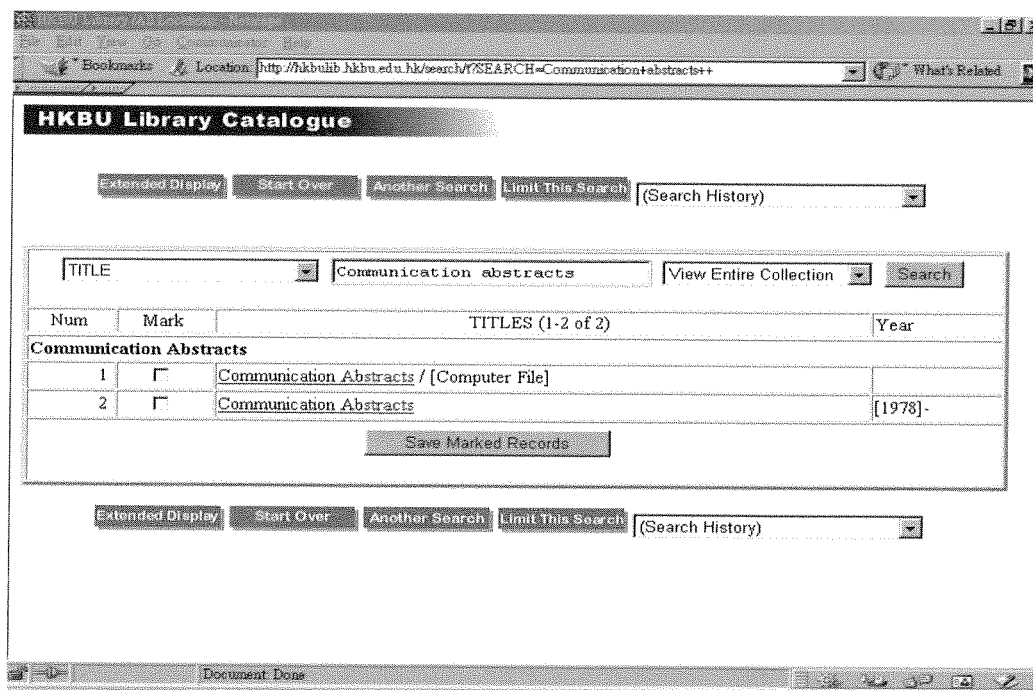


Figure 11. Duplicate Electronic Journal Title on HKBU Web OPAC

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.<sup>4</sup> 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](http://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](http://www.epnet.com/maglists/maglist.htm). Academic Universe title

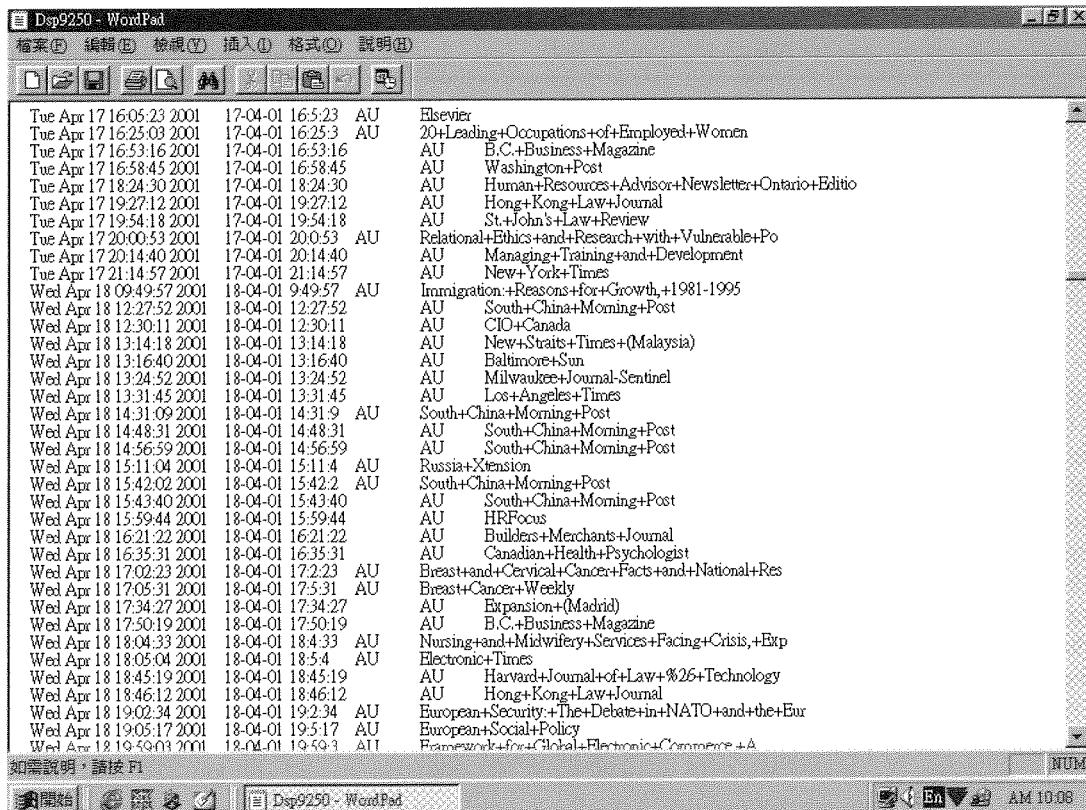


Figure 12. Usage Statistics for Academic Universe

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](http://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](http://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](http://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.