Libri, 2003, vol. 53, pp. 237–241
Printed in Germany · All rights reserved

Copyright © Saur 2003
Libri
ISSN 0024-2667

A Comparison of Print and Electronic Journal Holdings in Academic and Public Libraries

PETER MCCRACKEN
Serials Solutions, Inc., Seattle, WA, USA

The past few years have seen an enormous growth in the number of journals that libraries can access electronically. This is a well-researched and well-documented phenomenon, but no studies have compared the actual number of electronic journals a library can access, particularly in comparison with the total number of journals they receive through all sources, including print. The growth of serials management companies suggests the possibility of this research. Through analysis of resources tracked by Serials Solutions, Inc., one can get a firm grasp on how many electronic journals a library is tracking, in comparison with the number of print or microfilm titles they have. This article presents the first comprehensive analysis of how many journals public and various types of academic libraries can access electronically, rather than in print. While highlighting the incredible increase in access to information electronically, it also identifies the limitations of the current research, particularly regarding libraries' print holdings. As data collection improves, future research will be more accurate, and will provide a basis for studying both the growth of electronic access to journals and the expected reduction in print access to journals.

Methodology

Serials Solutions, Inc., was the first company to provide a comprehensive solution to determining which journals a library could access through its electronic databases. The company has nearly 1200 separate clients, with products in use in over 1600 libraries around the world. Because Serials Solutions also offers complete MARC records for journals and an OpenURL resolver, accurate and authoritative data is a critical piece of the company's growth. While competitors have joined (and left) this market, none has anywhere close to as many clients as Serials Solutions, and therefore none has the ability to study this data as accurately as Serials Solutions.

The services provided by Serials Solutions require that we receive or collect lists of titles and ISSNs from database aggregators and publishers, and compile them into a single, large database. This database reflects holdings for all journals in all of the aggregated databases tracked. The manner in which the database aggregators describe their content, however, provides a challenge. Aggregators may or may not include an ISSN, and when they do, the ISSN may be incorrect. Aggregators may or may not include the publication location, abbreviations, subtitles, older titles, or other pieces of additional information. Often, the titles contain misinterpretations of the title or even basic misspellings. While the human eye can quickly see that variant title/ISSN combinations all intend to refer to the same serial, the computer cannot. Table 1 provides several examples of the irregular title and ISSN combinations that come from different providers.

The result of all of this variation is that Serials Solutions must spend a great deal of time "normalizing" these data to a single, agreed-upon title for each journal. To do that, and in order to deliver

This paper is part of the *Libri Special Thread on Internet Publishing*; started in no. 3 of the current volume

Peter McCracken, MLS. Co-founder, Serials Solutions, Inc., 444 NE Ravenna Blvd., Suite 211, Seattle, WA 98115 USA. Tel: +1 (206) 545-9056 ext. 17, +1 (866) SERIALS ext. 17. Fax: +1 (206) 525-9066. E-mail: peter@serialssolutions.com

Table 1. Three examples of variable title/ISSN combinations provided by database aggregators. In the second and third examples, invalid ISSNs are in *italics.*

Journal : Technological Horizons in Education	0192-592X
T H E Journal	0192-592X
T H E Journal (Technological Horizons In Education)	
T H E Journal (Technological Horizons In Education)	0192-592X
T.H.E. journal	0192-592X
T.H.E. JOURNAL (TECH HORIZONS IN EDUCATION JOURNAL	0192-592X
T.H.E. Journal (Technical Horizons in Education)	0192-592X
T.H.E. Journal: Technological Horizons in Education	0192-592X
T.H.E. Journal; technological horizons in education	0192-592X
Technological horizons in education	0192-592X
Technological horizons in education journal	0192-592X
THE journal	0192-592X
THE journal : technological horizons in education	0192-592X
Remix	
Remix	*1532-1327*
Remix	1532-1347
Remix	*1533-1327*
Remix (Emeryv. Calif.)	1532-1347
Remix (Emeryville, Calif.)	1532-1347
Hypatia	
Hypatia	*0886-5367*
Hypatia	0887-5367
Hypatia (Edwardsville, Ill.)	0887-5367
Hypatia- A Journal of Feminist Philosophy	0887-5367
Hypatia: A Journal of Feminist Philosophy	0887-5367
Hypatia: A Journal of Feminist Philosophy	*0887-5637*
Hypatia: a journal of feminist philosophy (Univ. of Oregon, Eugene)	0887-5367
Hypatia-A Journal of Feminist Philosophy	0887-5367

our MARC record product, we associate title/ISSN combinations with MARC records obtained from the Library of Congress. These records are all created to the highest existing cataloging standards, as set by CONSER, the Cooperative Online Serials Program, a component of the Program for Cooperative Cataloging at the Library of Congress. By associating these title/ISSN combinations to existing, quality MARC records, we are able to accurately determine the number of true serials that are available to a library in an electronic format.

To develop this study, the author created a cohort of libraries that are Serials Solutions clients from each of the major Carnegie classifications [1], plus a separate group of public libraries. The libraries in this study had to be Serials Solutions clients, as it is necessary to review each library's profile of subscribed databases to determine how many electronic journals they can access. Every fifth library in each Carnegie classification was selected, and checked to see if they were a Serials Solutions client. From the resulting set of Serials Solutions clients, sixty libraries were randomly selected from the Associates, Baccalaureates, and Masters classifications, and thirty from the Doctorals classification. Twenty-five public libraries, of all sizes, were also studied. No individual information about each library is expressed or described here; all information is presented in the aggregate.

The author compiled information about the number of journals the library can access electronically, and compared that with data about the number of journals the library can access in print. Two sources, the IPEDS data of 1998, [2] and the 2002 *American Library Directory* ("*ALD*"), [3] provided the most current information available about the number of journals each library can access. Discussion about the variation in those numbers will follow.

When Serials Solutions delivers A-to-Z title lists, these lists include multiple holdings for each journal. If a single journal is available through five separate databases from three different aggregators, the journal will be identified five separate times. In delivering MARC records, however, those five separate databases are listed separately, but on just one single MARC record. By counting the number of MARC records Serials Solutions would create for a client library if they purchased our MARC record service, we can determine how many true journals they can access. This number can then be compared with the number of titles listed in various other sources.

Many aggregators include non-serial items, such as transcripts, brochures, reports, supplements, corporate annual reports, etc., and these increase the number of titles they can claim to offer. These are not serial in nature, however, and should not be counted toward the number of journals their clients can access. Since no CONSER MARC records will be found for these non-serial items, they will not be included in the number of MARC records Serials Solutions would create for these libraries. On the other hand, many local or

regional papers, for which CONSER records do not yet exist, are included in these resources. In some cases, CONSER records exist, but have not yet been located. The result is that the numbers given as the number of serial titles available at an institution are somewhat undercounted.

After collecting data on the number of electronic journals a library can access, the author compared that with the number of journals a library has reported it can access. Because data from the two sources for print holdings (2002 *American Library Directory* and 1998 *IPEDS* data) often varied dramatically, the author always selected the higher of the two numbers. It appears that most institutions reported only print holdings, while some reported print and electronic holdings. See the *Additional Comments* section below for further discussion of this issue. The number of journals reported through IPEDS or *ALD* was added to the number of electronic journals, and the number of electronic journals was divided by the number of all journals. The results were computed and averaged with each Carnegie classification and in a separate section for public libraries. Those results are presented in Table 2.

Table 2. Comparison of e-journal and reported journal holdings by library type.

Institution type	Average number of reported journals [b]	Average number of electronic journals [c]	E-journals as a percent of all journals
Associate [a] (n=60)	874	5,617	86.5%
Baccalaureate [a] (n=60)	1,372	6,841	83.3%
Masters [a] (n=60)	3,178	7,910	71.3%
Doctoral [a] (n=30)	15,046	9,836	39.5%
Public, large (n=25)	3,529	6,362	64.3%

a. per Carnegie Classification, 2000.
b. greater of 1998 IPEDS or *2002 American Library Directory* (no IPEDS data available for public libraries)
c. Serials Solutions data, June & September 2003

Results and analysis

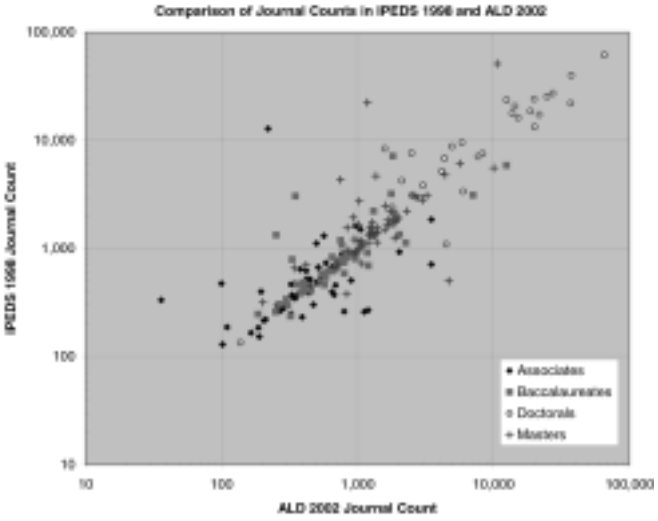
This study compared electronic journal holdings versus all journal holdings in sixty institutions in each of the Carnegie classifications of Associates, Baccalaureate, and Masters; thirty institutions in the Carnegie classification of Doctorate; and twenty-five public libraries. Twelve of those libraries are among the 75 largest public libraries in the country; twenty are in the top 10% by book volume; and twenty-three of the twenty-five are in the top 20% by book volume. As a result, the number of reported journals for public libraries is significantly larger than the national average for all public libraries. For Associates institutions, on average, 86.5% of their periodical titles are available electronically. For Baccalaureate institutions, on average 83.3% of their periodical titles are available electronically. For Masters institutions, on average 71.3% of their titles are available electronically. For Doctoral institutions, on average 39.5% of their titles are available electronically. For public libraries, on average 64.3% of their titles are available electronically. Since almost none of these institutions are tracking their electronic journals through their online catalogs, most academic libraries in the United States are unable to access between 80% and 90% of their entire periodical collection. They may have an A-to-Z list of their electronic journals, but this requires that patrons know to search both the OPAC and the A-to-Z list when seeking a specific title. Without effective quality control, some A-to-Z lists contain the numerous errors that were initially introduced by the aggregators themselves. Libraries without A-to-Z lists have almost no knowledge of what they can access through their myriad databases.

Students clearly prefer electronic access to journals, as any individual who has worked at a reference desk can attest. Libraries, in their attempts to fill these wants, have moved increasing amounts of their resources toward obtaining journals electronically, but have not kept up with tracking these journals for their patrons, or making these journals fully accessible to their patrons. When purchasing online resources, librarians need to ensure that their patrons can determine what is contained in these resources, and then access the resources in a manner with which they are familiar.

Additional comments

One great surprise from this study was the extreme variation in data reported by libraries to the various sources used for collecting informa-

Figure 1. Comparison of journal counts in IPEDS 1998 and *ALD* 2002 data.
Scales on both axes are logarithmic.



tion on the number of journals available in a library. The most recent IPEDS data available was from 1998; it was released in July 2002. While journal counts can certainly vary from year to year, and four years' separation between data collection could lead to some variation in the number of titles available at a specific institution, it cannot explain the massive variation between numbers provided by libraries for the two sources used in this study.

Knowing what is or is not a journal can sometimes be tricky. Knowing which ones to report, and when, can be downright hopeless. Libraries can be excused for not being certain which types of journals to report where, and some fault does lie with those creating data collection forms. But many libraries report very different numbers in the two different sources used in this study. The general instructions for 1998 IPEDS data collection sheet for the relevant fields reads,

> *Current serial subscriptions (lines 30 and 31)* – Report the total number of current serials received including those that are paid and those received without payment. Include government documents issued serially. Each available title counts as one when titles are received as part of an electronic subscription. [4]

Information Today, current publishers of the *American Library Directory*, does not provide a definition for the field in which they request information about periodicals: they simply request a number, and leave it to the library to decide how to define that number. [5]

Figure 1 shows a comparison of journal numbers reported to IPEDS and to the *American Library Directory*. One Associates institution reported 220 journal subscriptions to *ALD*, and 12,860 to IPEDS. Another reported 36 journal subscriptions to *ALD*, and 335 to IPEDS. Yet another reported 4700 to *ALD*, and 500 to IPEDS. Wildly varied results appear throughout the graph in Figure 1. (Note that scales on both axes are logarithmic, to better present the data.)

While it can be a challenge for an institution to determine data, and a challenge for others to collect data, little can be learned without accurate metadata. If researchers are to determine what works, and what does not, they must be able to access reliable data that accurately reflects comparable information across members of a cohort.

Conclusion

Electronic journals are a fact of life in today's library. Patrons and librarians recognize their value, and access through database aggregators, while occasionally problematic, is also incredibly cost-effective. Unfortunately, however, the vast majority of most libraries' journal holdings are now compiled within these databases, and very few libraries are able to track the journals in these databases, particularly through the OPAC. If libraries are to obtain the greatest possible value from the journals in these databases, they need to present information about these journals to their patrons.

Notes

1. The Carnegie Classification System is a systematic classification of institutions of higher education in the United States according to such variables as degrees offered, size, and commitment to research. The categories include: Doctorate-granting Institutions, Master's Colleges and Universities, Baccalaureate Colleges, Associate's Colleges, Specialized Institutions and Tribal Colleges and Universities. The Carnegie Foundation for the Advancement of Teaching, Menlo Park, California, offers a free online version of A Classification of Institutions of Higher Education (2000) at the URL: http://www.carnegiefoundation.org/Classification/index.htm [viewed 14 November 2003].

2. 1998 IPEDS Academic Library Survey File. Washington, DC: National Center for Education Statistics, 2002. Available online at the URL: http://nces.ed.gov/pubs2002/2002320.PDF [viewed 14 November 2003].
3. American Library Directory On Disc, 2002–2003. Medford, NJ: Information Today, Inc., 2002. CD-ROM.
4. US Department of Education, National Center for Educational Statistics. Integrated Postsecondary Education Data System: Academic Libraries Survey, 1998. Part L, General Instructions, page 3. (Form IPEDS-L.) Washington, DC: US Department of Commerce (as agent for the National Center for Educational Statistics), 1998. Available online at the URL: http://nces.ed.gov/pubs2002/2002320.PDF [viewed 14 November 2003].
5. Personal Communication with Beverly McDonough, managing editor, Information Today. 8 August 2003.

Editorial history:
paper received 3 November 2003;
accepted 11 November 2003.