How Much is Enough?

An Analysis of Electronic Serials Holdings Duplication at Indiana University Bloomington

Kris Stenson

S505

Spring 2011

Abstract

Budget pressures are forcing many academic libraries to reevaluate their collections and how they provide access to their patrons. Serials in particular are seen as a place where format duplication is common and budget cuts can be made. Full-text aggregated databases contain much of the same content that is found in thousands of journals, and may present an alternative to individual journal subscriptions. This paper examines current e-journal and full-text aggregated database subscriptions at Indiana University Bloomington in order to determine the current rate of content duplication. The goal is to determine how much cutting can yet be done, and which areas may be most suitable for such cuts. The paper does not make specific recommendations, but should provide useful inferences for acquisitions librarians at IUB Libraries. It also points to academic disciplines where more comprehensive future studies should be focused.

Introduction

We live in a world of virtual library resources. Journals, databases and books are all purchased or leased in electronic format in quantities unheard of even a few years ago. This is particularly true in the world of serials, where e-journals take up an ever-increasing proportion of library budgets (Zappen, 2010). Partially this is due to format changes enforced by the publishers themselves, but for the most part librarians have learned to embrace the necessity of moving journal subscriptions online. Online journals are free of physical constraints, allowing patrons to access resources remotely. Multiple users can access the same journals concurrently, multiplying the value of the purchases made. From the early days of e-journals, however, librarians and others have expressed concerns over the reliability of content in the electronic environment (Sprague & Chambers, 2000). The same lack of physical being that enhances the flexibility of electronic resources also gives them a sense of impermanence. The rapid expansion of e-journals in the last decade has started to alleviate this worry, but it has not disappeared entirely.

In conjunction with the rise of the e-journal is that of the full-text aggregated database. These databases have the advantage of concentrating information from myriad journals in one search interface, aiding the researcher greatly. One might ask whether such databases can serve as a substitute for individual subscriptions, saving money in the acquisitions budget while still providing the same content to patrons. Culpepper (2000) and Sprague and Chambers (2000) attempted to answer this question a decade ago, but found that it was not as simple as comparing lists and eliminating the duplicates. Their studies found that the definition of "full-text" varied from database to database, with less than full coverage of the journals contained therein. Concerns were also raised about the somewhat capricious addition and removal of journals from database holdings. Ultimately the studies cautioned against cancelling journals in favor of database holdings, while acknowledging that future budget crunches would likely make such actions necessary regardless.

The realm of full-text aggregated databases has changed a great deal since the late 1990s, however. Online formats have advanced considerably, and the content found in an e-journal can be said to adequately match that which is found in print in most cases. Many more databases contain full-text listings, and their reliability has been shown through time (Nixon, 2010). In short, what was once seen as an unpredictable and potentially risky area of expansion in 2000 has stabilized and become an integral part of library collections. Budget shortfalls continue, although the debate now is not so much about print versus electronic but rather individual subscription versus database. Nixon (2010) recently undertook an examination of journal holdings at Purdue University in light of continued budget cuts. The goal was to eliminate some of those titles which were held in full-text aggregated databases. While the concerns and risks associated with relying upon databases have not totally disappeared, Nixon found that in many

cases those risks were tolerable and the budget freedom to be gained from taking such steps was worthwhile.

With the recent success at Purdue it seems clear that librarians are now armed with a tool not previously available to them. Indiana University Bloomington, like all public institutions, faces continued budget cuts in this year and in those to come. It will soon be necessary, therefore, for the university to continue to save funds through the elimination of journal subscriptions. With this in mind, this study will seek to answer the following question: To what extent do e-journal holdings at IUB duplicate those found in subscribed full-text aggregated databases? I have chosen to focus on e-journals for this study as print journals are already being cut at a rapid pace. In most cases, however, they are only being replaced by their electronic counterpart, not eliminated entirely. The savings are thus limited from this practice. The elimination of all types of individual journal subscriptions represents a much more controversial step, and is an area where more discussion is needed. Only active subscriptions will be considered in this study, as cancelled serials represent no further drain upon the budget. Duplication in this case is defined as those titles which are present in full-text aggregated databases with the same content and currency as individual subscriptions. Full-text aggregated databases in this case means those databases composed of numerous journal titles, fully indexed and searchable, with all articles present in their entirety.

This study will not seek to make budgetary recommendations based upon the data collected and analyzed, but it should benefit those who make such decisions. Being as the focus is specifically on IUB's holdings it should be of interest to library administrators and acquisitions librarians at that institution. Acquisitions librarians in particular are called upon to make budgetary choices relating to journal and database subscriptions daily and as such are in need of

in-depth data to aid their decision-making. This study should provide them with just such data

so that they might find more practical and beneficial solutions to the current budgetary

challenges they face. While the specific results of this study will not be applicable beyond IUB,

the general trends in the findings should nonetheless prove informative to other similar academic

libraries as well.

Literature Review

From the late 1990's, when electronic journals first became widely available to libraries,

the debate over their relative value has ranged far and wide. They have been analyzed in terms

of cost, access and ownership (Culpepper, 2000; Smith & Cochenour, 2007; Zappen, 2010).

They have also been played against print in a variety of studies comparing these different

aspects, with widely varying results (Kalyan, 2002; Maple, Wright & Seeds 2003; Smith, 2009).

Discussion of format duplication has typically focused on cost versus reliability, but this debate

has evolved over time. Early discussions of e-journals focused largely on reliability and

completeness of content in the online environment (Tenopir, 1997). These concerns were

magnified when discussing aggregated databases, as they lacked even the assurance that came

with individual subscriptions (Hawbaker & Wagner, 1996). Studies by Culpepper (2000) and

Sprague and Chambers (2000) seemed to reinforce these misgivings, as they found that full-text

content in aggregated databases was not always what it seemed to be. Databases, while seen as a

good resource in tandem with journal subscriptions, were simply not trustworthy enough to rely

upon as the library's sole source for journal content.

Even e-journal subscriptions had some initial difficulty gaining acceptance among the

library community. The lack of physical possession caused a real sense of misgiving among

librarians, and certainly these worries were somewhat justified (McKay, Mouw, Kusma, Hurd & Sonberg, 1998). The acquisition and management of print journals is a straight-forward process consisting of ordering, receiving, binding and shelving materials. E-journals come with a host of new considerations, such as licensing, accessibility to users, digital archiving, provider stability and completeness of holdings, not to mention technological compatibility and cost (Rupp-Serrano, Robbins & Cain, 2002). Evolution of the library environment and ever-increasing costs would not allow librarians to stay with the comfort of print forever, however. These pressures prompted many libraries to evaluate their resources in terms of format duplication. A study conducted at Seton Hall University in 2001 found that print serials consumed nearly seventy percent of the library's materials budget (Kalyan, 2002). Another such project, undertaken at Penn State University in 2001, found that nearly $97,000 could be saved by eliminating all print subscriptions beyond the second copy (Maple, Wright & Seeds, 2003). Clearly there was money to be saved by reducing print subscriptions in favor of electronic.

It has been a number of years now since the initial push for e-journals began, but libraries continue to face pressure to evolve their holdings. Library budgets keep shrinking, and the cost savings associated with switching to the electronic format are not as great as once hoped. Like print, e-journal prices are subject to inflation, although not exactly in the same way (Smith & Cochenour, 2007). Individual license agreements, consortia and tiered pricing all make it difficult to predict the cost of e-journals from year to year. E-journal "Big Deal" packages also have faced scrutiny in recent years as prices continue to grow and librarians become frustrated with their lack of flexibility (Carlson & Pope, 2009). Possible solutions to this problem that have been proposed include interlibrary loan services and increased usage of open-access journals. While these options may be financially appealing they sometimes fail to fully address patrons'

access needs. Full-text aggregated databases have been proposed as an alternative, as they contain content from many major journals while costing less than individual subscriptions (Carlson & Pope, 2009). While not always perfect in their coverage, such databases have been used as a cost-cutting tool by a number of different libraries.

While the work of Culpepper (2000) and Sprague and Chambers (2000) found that full-text aggregated databases were not an optimal alternative to individual journal subscriptions, not all have agreed with their assessment. Kalyan (2002) found that, despite shortcomings, such databases could be used to provide access in lieu of journal subscriptions. While not ignoring the problems with databases (embargoed content, lack of 100% coverage of full-text listings), she found that such limitations could be minimized. In the process of eliminating duplicate holdings focus was given to those titles located in more than one database, thus increasing their chances of total coverage. She notes that while some small gaps are likely they can be handled through inter-library loan services on a case-by-case basis. A more recent project at Purdue University employed similar methods to trim the serials budget in 2009. That project focused on journals with low circulation, as well as those that were present in multiple databases. In addition, some of the money saved through cancellations was set aside in a new fund. This fund was specifically designed to re-acquire journal subscriptions should those titles disappear from databases in the future (Nixon, 2010). In this way some additional safety net was built into the process, guarding against the sudden loss of resources.

Analysis of coverage of titles within databases has been done through various methods, but typically involves a comparison of title lists. Typically, as in the two studies noted above, lists of subscriptions at a specific library or library system are compiled and then compared to holdings lists for selected databases. Since each library's holdings are unique in composition to

that specific location it is difficult to utilize any sort of master-list compiled by another source, at least in terms of cost-cutting projects. Others have taken overlap analysis in a different direction, however. Schroeder (2008) utilized subject-specific citation databases, faculty recommendations and citation analysis to compile a list of journals in the field of occupational therapy. The goal was to compare that journal list with holdings in different databases, identifying the optimal combination of database and individual e-journal subscriptions for the occupational therapy program at Grand Valley State University. While this project was budget-driven to some extent, it also sought to optimize the coverage of journals for a particular field of study within that budget. This demonstrates the benefit of overlap analysis not just as a way to trim collections, but also to improve them for the patron.

With the growing acceptance of e-journals and full-text aggregated databases as substitutes for print subscriptions, it seems only natural that librarians should begin to examine their print backfile as well. Even as e-journal subscriptions have gradually taken the place of print, libraries have traditionally maintained their own print backfiles of titles in their collection. There is evidence that this is beginning to shift, however. Christina Torbert (2009) of the University of Mississippi notes that libraries can benefit from both space savings and increased user access to materials through the conversion to electronic backfiles. Conducting a survey of library professionals, Torbert (2009) found that 81% had purchased electronic backfiles in the previous three years. Of those respondents, 37% were either discarding the duplicate print volumes or were moving them to storage. The same percentage of respondents were opting to keep the print backfiles in place. These results indicate that the discussion of format duplication is still very much in flux, and individual situations dictate different policies from institution to institution.

If full-text aggregated databases are being used in place of journal subscriptions by some libraries, can the same also be said of journal backfiles? While not yet prevalent, some institutions are doing just that. Adelphi University Libraries have chosen to eliminate a great deal of their print backfiles in favor of content available online in full-text aggregated databases (Smith 2009). Microfilm subscriptions, in the past a space-saving alternative to print backfiles, have also been eliminated if titles meet certain online availability requirements. Exceptions have been made for titles that contain superior graphical content in print. Much like with similar projects involving active subscriptions, Adelphi University Libraries have focused on titles that are present in more than one database as a way of ensuring content security. The results of their efforts were evident in an annual budgetary savings of $27,664 (Smith 2009), coming from both cancellation of microfilm subscriptions and reductions in binding costs.

Overlap analysis can also be used as a tool for the streamlining of database subscriptions themselves. Banks (2005) conducted overlap analysis of 5 different bibliographic databases at the University of Illinois Urbana-Champaign to determine whether a particular database, Wilson Social Sciences Abstracts (WSSA), could be eliminated from the budget. That study found that only 23 of the 615 indexed journal titles were unique to that database, and over 85% of titles were indexed in at least two other databases. This data was then combined with a cost-use analysis showing WSSA to be far more expensive per use than other reviewed databases. After some debate among faculty and librarians the database was ultimately eliminated from the library's budget. This demonstrates that the potential for overlap analysis is not limited to different formats, but also to different resources of the same type. It is thus a highly valuable collection evaluation tool for acquisitions librarians to utilize.

Librarians continue to face the challenge of providing optimal resources with a limited budget, and must use all evaluative tools at their disposal. This study is a natural progression of the above noted work. It operates with the assumption that overlap analysis is a proven, effective means through which to evaluate a library's holdings in a variety of formats. Furthermore, it addresses the growing need to evaluate resources not just in terms of print vs. electronic, but also e-journal subscription vs. database. The former debate has already run its course, while the latter will have primacy in the library environment for years to come.

Methods

This section will first explain the reasoning and methods by which e-journal titles were selected for the study. It will then discuss the same for the sampled databases. Special cases and exceptions will be discussed in greater depth. Methods for analysis of the results will be explained as well. Finally, the section will examine ethical ramifications and limitations of the study.

Indiana University Bloomington currently has on order around 3400 individual journals in electronic format. This number can be deceiving, however. It does not include certain society memberships, publisher bundles or subscription plans, each of which can contain dozens or more titles and which can frequently change in composition. It does not include thousands of open-access journals that the IUB Libraries link to through their website. The exact number of e-journal subscriptions also varies from month to month, as subscriptions are added and dropped regularly. Given all of this variability it is nearly impossible to pin down an exact answer to the question "how many e-journals do IUB Libraries subscribe to?" It was decided, therefore, to frame the current study as an exploratory case study with fairly narrow definitions and aims.

While every attempt was made to generate statistically relevant results, exhaustive and comprehensive coverage was not attempted.

Two main sets of data were required to perform this study: a list of all current e-journal subscriptions and a list of currently subscribed full-text aggregated databases. E-journal data was gathered from the Sirsidynix Symphony integrated library system currently used to manage IUB Libraries collections. A report was created using the IU Information Environment (IUIE) that pulled title, publisher and format data for all current IUB subscriptions. This data was in turn imported into an Excel spreadsheet for further manipulation. In this case "current" refers to those subscriptions that have been active in the 2011 fiscal year (July 1 2010 - June 31 2011). The report was run in late February 2011, and reflects subscriptions active up to that time. Subscriptions that contained the format designation for either e-journal or print & e-journal were collected in this report (4178 orders in total), but those for print alone were not included. There are two main reasons for the exclusion of print. For one, it was felt that print subscriptions are already heavily analyzed for possible conversion to electronic format and thus an additional study on such materials was unnecessary. Secondly, IUB has been actively seeking to convert resources to electronic format for a number of years, and thus many of the subscriptions remaining in print do so because they are not available in any sort of electronic format. To include such materials in the research sample would thus have the potential to skew the results without greatly adding to their informational value.

Once the spreadsheet was created the data was sorted to further narrow the scope. First, duplicate title entries were removed so that only one instance of each title remained. These duplicate orders typically fall in one of two categories: subscriptions maintained by IUB for exchange partners and titles for which there has been a format change during FY 2011. An

example of this latter category would be a subscription that was maintained as print and electronic through the end of calendar year 2010 and then was switched to electronic only with the beginning of 2011. Both orders would thus exist in FY 2011. Orders for subscription plans, memberships, and bundles were removed as well, as each one contains a number of different titles. This decision was made due to practical considerations. Each of these orders would need to be manually examined, with lists of included titles checked for electronic access, and it was deemed impractical to include them. Since the study seeks to be descriptive rather than comprehensive the author does not believe this omission undermines the results in any way. Finally, all titles included in one of the "Big Deal" packages from Springer or Wiley-Blackwell as well as those from Elsevier were excluded. These titles typically exist solely within the publisher's proprietary databases, and as such would not be present in the databases being used for this study. Furthermore, since Big Deal packages are subject to cancellation restrictions, among others, it seemed unlikely that any results gleaned from this study could be put to practical use with those titles.

After removing the above mentioned titles from the spreadsheet 1853 of the original 4178 titles remained. For practical reasons it was decided to sample from these titles rather than check each one. Simple random sampling was used to avoid bias from the author and most accurately reflect the collection as a whole. First, the necessary sample size was calculated using the online Creative Research Systems sample calculation tool.[1] A 95% confidence level and +/-5% confidence interval were used, resulting in a sample size of 318. Next a random number table

---

[1] http://www.surveysystem.com/sscalc.htm.

was generated using the Research Randomizer online tool.[2] That table was then used to create a study sample of 318 unique titles from the main title list.[3]

Regarding the databases to be used in the study, an initial list was created using the Serials Solutions e-resources management service. Serials Solutions acts as the portal through which most electronic journals, books and databases are managed at IUB Libraries, and maintains detailed holdings information and usage statistics related to those resources. Using assessment tools within the client, a report was run in late February 2011 detailing all IUB database usage for a time period from September through December 2010. This time period was chosen as it roughly equals the Fall semester at IUB, as well as being contained entirely within FY2011. The resulting list needed a great deal of filtering before being of use, however. Serials Solutions includes in its definition of "databases" all e-journal hosts, not just aggregators. All those resources listed that were not true aggregated databases thus had to be removed from the list. The electronic resources librarian at IUB, Lori Duggan, was consulted for much of the database filtering process, as her expertise in the area helped resolve several "grey area" type problems. Publisher-specific databases such as Elsevier ScienceDirect and Springerlink were excluded for the same reason Big Deal titles were not included in the e-journal list. Since they comprise a closed system and are essentially just content hosts for their own titles they are not the focus of this study. Databases comprised of different parts within one main system, such as ABI/INFORM, were compared to master listings in the A-Z resource list on the IUB Libraries website.[4] In most cases these were treated as one database, and the usage data for the various parts combined in the final tally. Databases that are primarily abstract or bibliographic rather than full-text in nature were excluded as well, as were those that focus on backfile rather than

---

[2] http://www.randomizer.org/form.htm.
[3] See appendix A.
[4] http://www.libraries.iub.edu/index.php?pageId=1044&mode=alpha&letter=ALL.

current content. Some, such as LexisNexis Academic, contain both full-text and abstract

components, but were included since the full-text was considered significant. Individual

decisions had to be made to include or exclude several databases that did not perfectly fit the

study parameters. Journals@Ovid contains both individual subscription and database collection

components, but was included due to it being primarily a database in nature. Factiva was

excluded due to inaccurate reporting of contents and inconsistent access to full-text as noted by

the Serials and Electronic Resources Acquisitions unit at IUB Libraries. JSTOR represented a

different conundrum. While it is consistently one of the most heavily used resources at IUB its

contents are traditionally embargoed for 3-5 years. Closer examination revealed a significant

number of titles listed as current, and so it was included in the study.

   With the list trimmed of inappropriate resources the decision had to be made as to how

many databases to include in the final study. While wanting to most accurately reflect IUB

Libraries' holdings it was acknowledged that too many selections would make the study

unwieldy to perform. It was decided that the databases should be selected first based upon

usage, rather than more arbitrary criteria such as subject area or "significance." The question

then was then where to cut off the list. Due to the hazy categorization of many of the listed

"databases" a clear count was deemed unfeasible, and thus no attempt was made to ensure that

the selected databases consisted of a statistically valid sample of all databases. Ultimately all

databases with more than 200 uses during the selected time period were chosen (minus omissions

noted above). This resulted in 28 total databases, with usage ranging from 209 to 38,265 uses

and including both subject-specific and generalized content. This represents approximately 94%

of database usage for the time period examined.[5]

---

[5] See appendix B. The 94% figure is based upon filtering the entire list through the same criteria used for the sample.

With the samples selected, the next step was to identify which of the selected e-journals were present in those databases. Several methods were considered for this process. Consulting database provider-supplied titles lists was the first consideration. Ultimately this was eliminated as not all companies provide or clearly identify such lists. Additionally, copies of such lists on file with the library are often out of date, as database holdings frequently change. Manually checking each database for the titles in question would no doubt yield more accurate results, but given the large number of titles and databases to examine it was deemed too inefficient and time-consuming.[6] A far more efficient solution was identified through the use of the aforementioned Serials Solutions. That service maintains extensive listings of all holdings for any given title. Those listings are updated regularly through a combination of publisher updates, vendor notifications and information from libraries using the system.

All 318 titles were searched individually within the Serials Solutions client interface, and the selected databases were reviewed as well to double-check the results. The next question that arose was that of what would be counted as current access. Both individual e-journal hosts and aggregated databases update holdings at different rates. Furthermore, different journals publish on vastly different schedules. No one hard and fast rule could thus be applied as to what "current" really meant. Ultimately a multi-level approach was taken. If a title was identified as being present in one or more of the selected databases the first step was to identify the most recent issue currently online on the primary e-journal host. With that as a benchmark the database holdings were then examined for comparison. If the holdings were up to the same issue as the e-journal host they were naturally considered current and counted. If those holdings were not current but the coverage was within the last six months the amount of lag was recorded (one issue behind, etc.), and the title was counted. It was decided to use "issues" as the unit of

---

[6] Checking for every selected title in each of the 28 databases would result in 8904 unique searches.

measurement rather than months due to the different publication schedules. For example, if a title is an annual and the holdings are one issue behind that equals 12 months. If that same title is a monthly and it is one issue behind it equals only one month. The author felt that expressing gaps in terms of time could thus bias the results.

Once all title holdings had been searched and duplication recorded using the above method the results were categorized in terms of types of access (no database coverage, current, etc.).[7] Totals and percentages were calculated for number of titles with any duplication, as well as the specific nature of that duplication. Titles with duplication were subsequently categorized by discipline to identify any trends that might be found in that way. Major publishers present in the sample were identified and their titles present/not present in databases were compared. Finally, the results were broken up by database in order to provide clearer focus for future research.

Ethical Considerations

As a matter of disclosure it should be noted that the author is currently employed in the Serials Acquisitions department of IUB Libraries, which has served as the impetus for this study. That position has also afforded the author access to a number of closed systems that would not otherwise be available to him. When beginning the planning process it was necessary to first consult with the head of Serials Acquisitions, Pam Owens, regarding the appropriateness of using certain internal data for the study. While there were no ethical concerns about study participants or the like it was nonetheless important that no internal data was used in a way deemed inappropriate by library administrators. Ms. Owens, in conjunction with the head of

---

[7] One special case relates to titles which are hosted by JSTOR as part of their Current Scholarship Program (CSP). While still part of the JSTOR framework, the CSP represents individual subscriptions rather than aggregated content.

Acquisitions, Lynda Clendenning, granted approval for the use of library data from both the

SirsiDynix and Serials Solutions systems for this study.

Implications

While this study focuses on the unique collections of Indiana University Bloomington it

is hoped that the results will be applicable at more than just that institution. The study does not

seek to be comprehensive, but rather to provide a picture as to the current level of content

duplication between individual subscriptions and database holdings. The study is exploratory in

nature, and thus will not be able to, on its own, provide direction for action either at IUB

Libraries or other institutions. The findings should, however, contribute to the ongoing

discussion of the management of e-resources within the academic library environment.

Results

This section will examine the findings of the study from several different perspectives.

In addition to the basic numbers, trends will be identified in terms of disciplines, publishers, and

databases related to the sampled titles. Possible meanings and implications of these findings will

be subsequently discussed, including limitations to the data.

Currency

A total of 47 of the 318 sampled titles, or 14.8%, were found to have current coverage in

one or more of the selected databases. Some titles were present in multiple databases, leading to

a total of 71 distinct "instances" of coverage (1 title in 1 database = 1 instance). As noted before,

this includes some cases where the title was not completely up to date with the e-journal host, but

was within a reasonable time frame (within 6 months and 2 issues). Of those 71 instances, 14

were within one issue of current and one was within two issues of current. The remaining 56

instances represent coverage that was up to date with the primary e-journal host. Not included in

these numbers are cases where a title was present in the database but was not complete. Six

cases were noted where the database contained the majority of, but not totality of, current issues

of a title. Cases where a database contained only a small percentage of a title's content were not

recorded.

Titles Found

Of the 47 titles found in at least one database, 10, or 21.3%, were found in two or more.

The most widely found title, *The Journal of Leisure Research*, was found in a total of six

databases. This was followed by *Environmental Hazards* (5), *The Journal of Public*

Administration Research and Theory (5), *Health Affairs* (4), *Language Testing* (3) and *Teacher*

Education and Special Education (3). Four more titles were present in two databases each (see

fig. 1). The mean number of databases for any given title present in at least one database was

1.51, the median and mode both 1 and the standard deviation 1.2.

Figure 1:

Journals Found in Databases
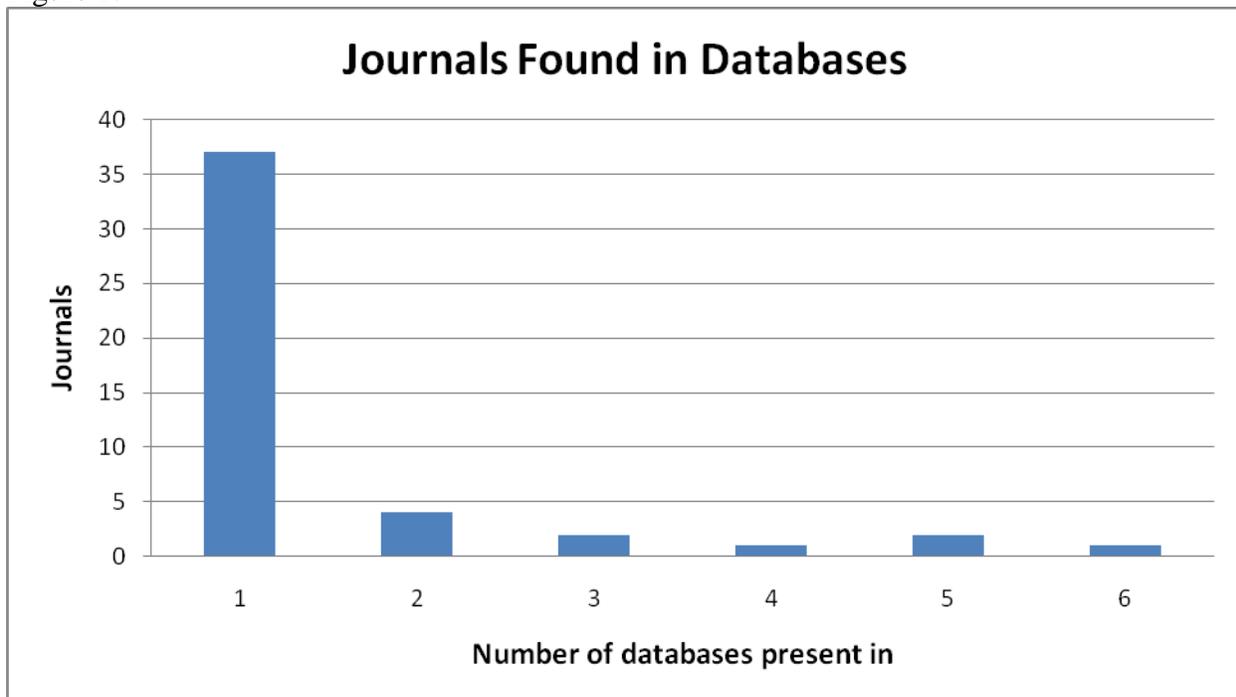
Disciplines

In order to determine whether the frequency of titles in databases differs from subject to subject the title list was sorted according to academic discipline. This was done initially through the use of university fund codes, which identify which collection manager and subject area each title falls under. A total of 55 fund codes were represented in the sample. Fund codes do not necessarily equal disciplines, however, as some cover multiple areas (such as "reference" or "general") and others represent only a subset of one area (e.g. communications and journalism). Furthermore, 55 categories are far too many to glean useful data out of a sample of 318. As definitions of academic disciplines are frequently changing and often debated it was decided to use the highest level classifications possible. Five classifications were identified: humanities,

social sciences, natural sciences, formal sciences and professional and applied sciences.[8] Fund

codes were identified as belonging under one of these categories and the total number of titles for

each was calculated. In the case of fund codes that did not correlate to a discipline, such as

"general," the individual titles under that code were classified into other appropriate categories

individually. As fig. 2 shows, disciplines with more sampled titles predictably had more titles

present in databases as well. It should be noted, however, that the actual percentage of titles did

vary. Close to 20% of titles in the humanities and professional and applied sciences were found

in databases, compared to less than 10% of titles from the social, natural and formal sciences (see

table 1). Of the 10 titles found in more than one database, 7 were within the professional and

applied sciences, while 3 were from the humanities.

Figure 2:

---
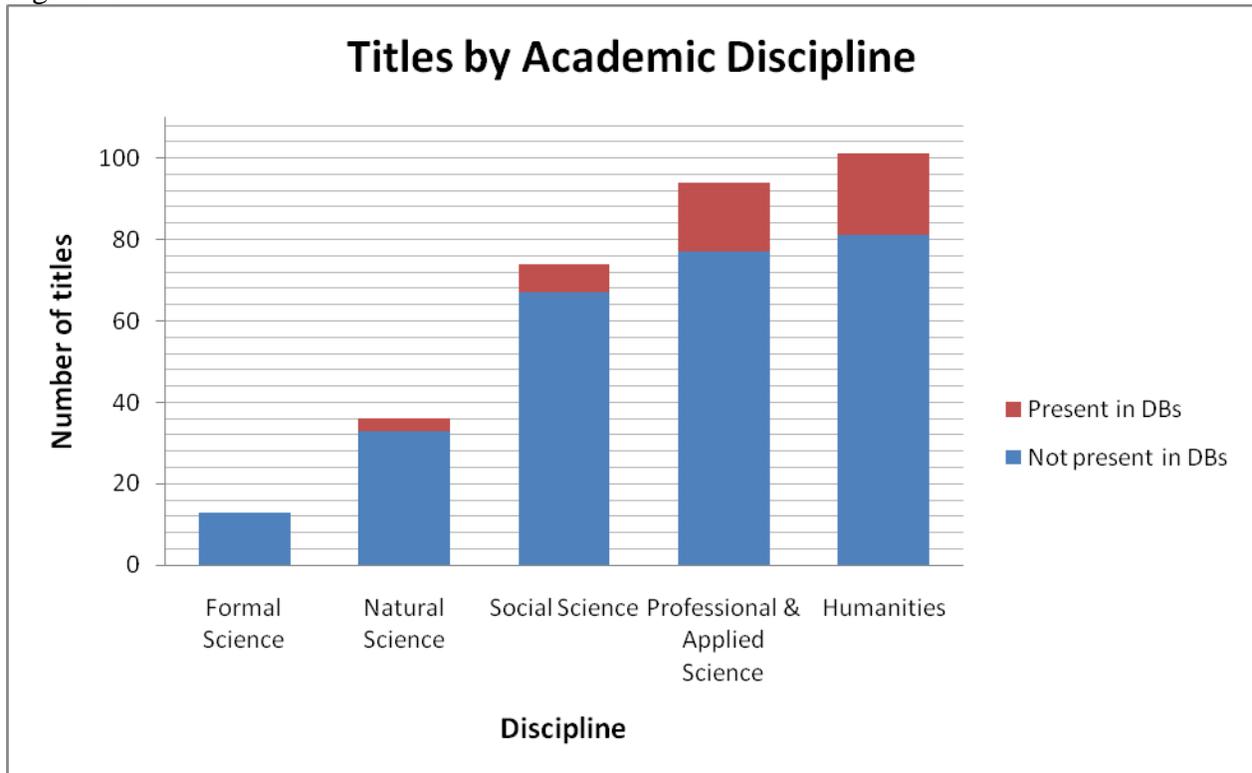
[8] See Appendix C for a complete list of how subject areas were grouped into academic discipline. Specific fund codes are not listed due to protect confidentiality, but the areas they represent are.

Table 1:

Discipline	Titles in Databases	Titles in Sample	Percentage
Formal Science	0	13	0%
Natural Science	3	36	8.3%
Social Science	7	74	9.5%
Prof. & App. Science	17	94	18.1%
Humanities	20	101	19.8%
Total	47	318	14.8%

Of those titles found in databases, 42.5% were in the humanities, 36.2% in professional and applied sciences, 14.9% in social sciences and 6.4% in natural sciences. No titles found in databases were from the formal sciences.

Publishers

In an attempt to identify the likelihood of any given title being present in one or more databases the sample list was sorted by publisher. Within the sample 80 distinct publishers were identified. All publishers with five or more titles within the sample were then separated for further analysis. This resulted in a list of seven major publishers: Brill (9 titles), Cambridge University Press (13), Maney (7), Oxford University Press (20), Sage (33), Taylor & Francis (69) and Walter de Gruyter (7), representing 49.7% of all sampled titles. Past these seven publishers there was a steep drop-off in the number of titles per publisher, hence the cutoff of five titles. There were also 42 titles for which no publisher was identified in SirsiDynix[9]. The seven selected publishers were then compared with the list of titles found in databases. None of the titles from Brill, Cambridge or Maney were present in any of the selected databases. Taylor & Francis had two titles present in at least one database, which represents a mere 2.9% of its sampled titles. Oxford had but a single title present, representing 5% of the titles from that

---

[9] Small publishers that may only produce one or two titles are not typically recorded or tracked in the IU system.

publisher. There were six Sage journals present in databases, or 18.2% of sampled Sage titles. Walter de Gruyter had the largest representation, with four of its seven titles, 57%, present in at least one database (See fig. 3).

Figure 3:
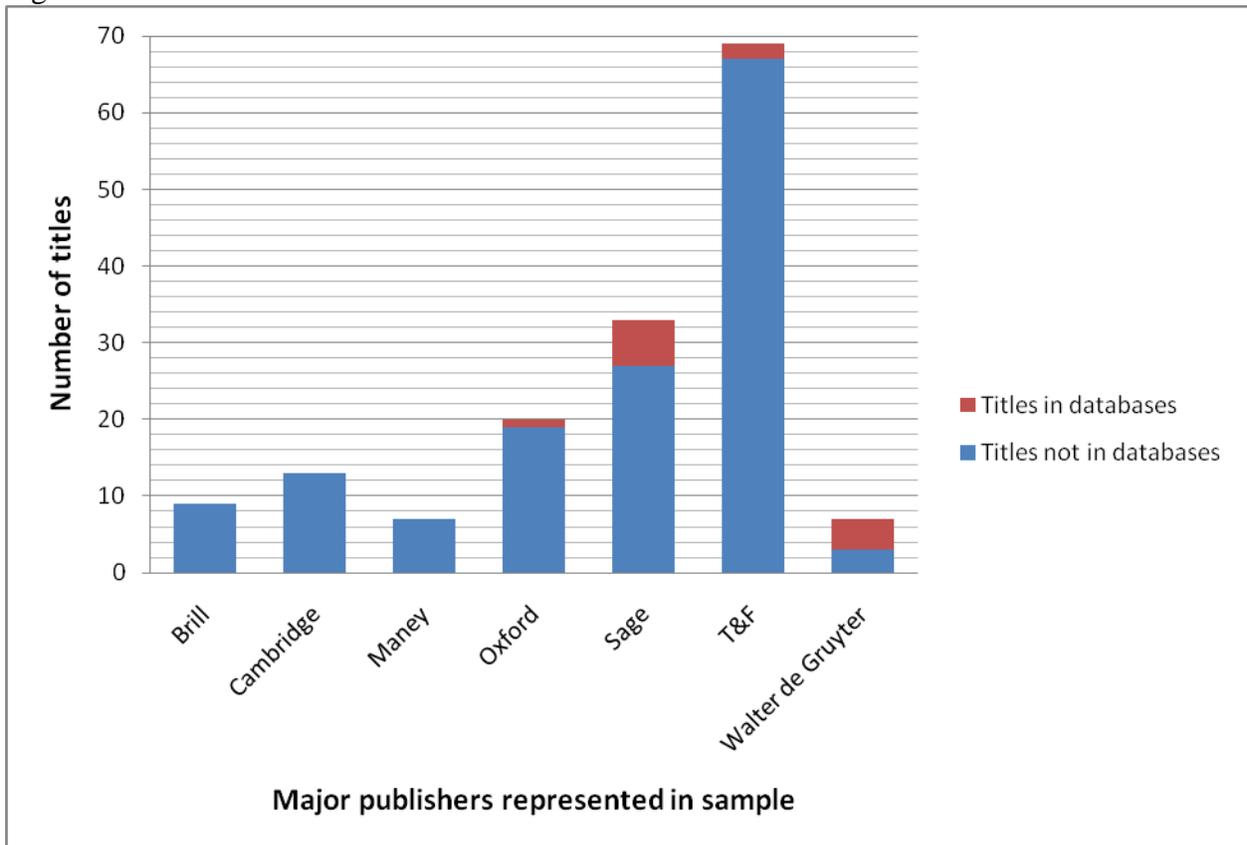
Databases

While 28 databases were included in this study, only 17, or 60.7%, contained current holdings for any of the sampled titles. Table 2 shows the differences between all selected databases and those with sampled titles present. All but one database had less than 10 titles, with the one outlier of 21, Academic Search Premier, working to skew the numbers upward (see fig. 4). The 21 titles which were found in that database represent 6.6% of the sampled titles, and

29.6% of total instances of coverage. Communication & Mass Media Complete was second with

nine titles (2.8% and 12.7%, respectively) and MasterFILE Premier came in third with six titles

(1.9%, 8.4%). Four databases contained only a single instance (0.3%, 1.4%): CINAHL Plus,

GREENR, Health & Wellness Resource Center and Journals@Ovid.

Table 2:

	All DB	DB with titles
Mean number of titles:	2.63	4.18
Median number of titles:	1	3
Standard Deviation:	4.3	4.8
Mode:	0	3

Figure 4:
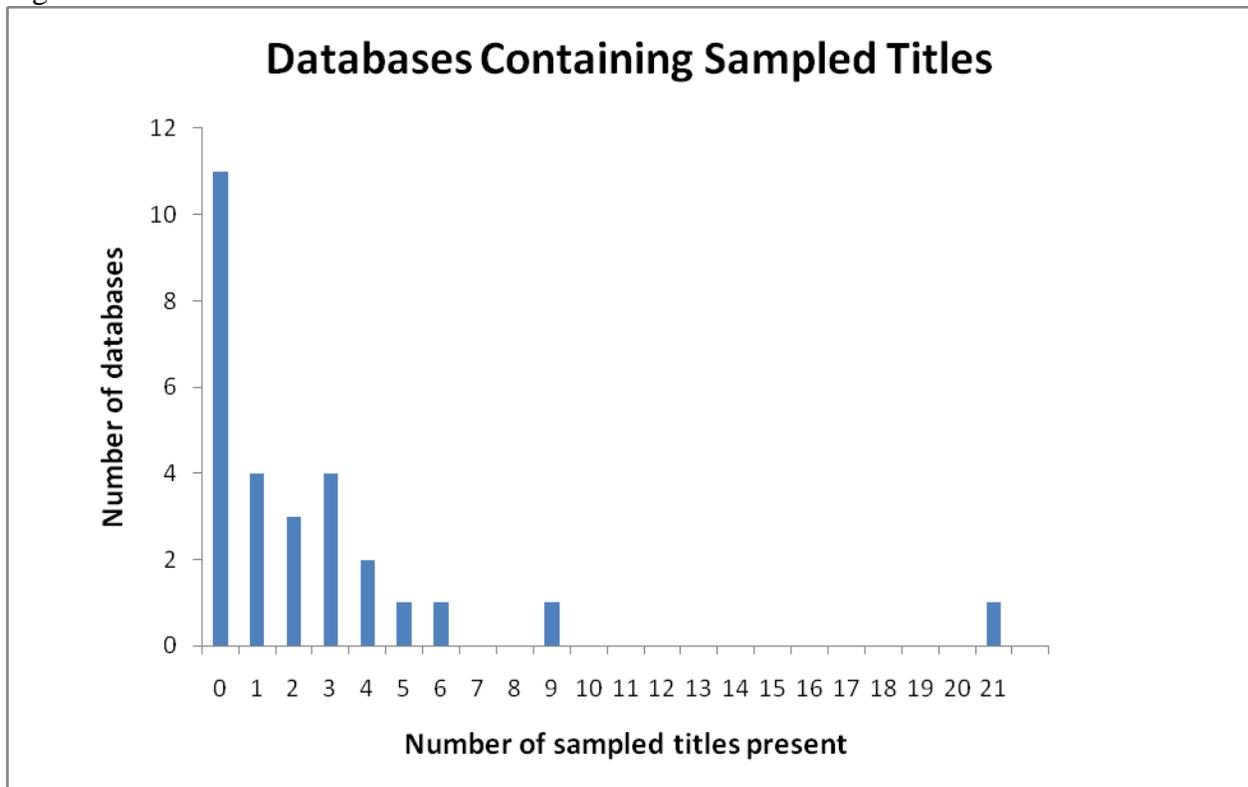
Discussion

The discovered rate of journal duplication was lower than initially expected. While no particular ratio of titles was assumed, 14.8% seems low given the wide range of titles covered by full-text aggregated databases. It is important to note, however, that this percentage represents a snapshot of collections specific to IUB, and does not reflect upon the entire e-journal universe. It is entirely possible that if one was to examine all currently available e-journals and compare them to all currently available aggregated databases one would find a higher percentage of convergence. Such a study would be highly difficult, if not impossible, to implement, however. The constantly changing serials environment means that any study is forced to be a snapshot of a particular place and time.

Examining the title duplication by academic discipline may provide greater potential for more comprehensive projects. The data shows that journals in the humanities and professional and applied sciences have close to double the rate of duplication as those in other fields. As such, these areas would be excellent candidates for more focused examination with an eye on reducing resource duplication. The same trend appears to hold true in terms of how many databases each duplicated title is found in, although it is possible that those numbers are unduly influenced by the small number of titles. While a comprehensive examination of all titles across all disciplines may be impractical, more limited projects focusing on where there is the most likely benefit are certainly within reach. Further work is needed to determine with more specificity just which subject areas within those disciplines are most prone to duplication. The sample used for this study is too small to accurately determine trends at that level.

There is naturally a high concentration of titles from large, established publishers such as Taylor & Francis, Sage, Oxford, etc. at IUB Libraries. The amount of intellectual control each of these publishers exerts over their respective titles has a direct effect upon what is found in full-text databases. That factor weighs heavily on the results of a study such as this. As it is, the results clearly show that publishers such as Sage and especially Walter de Gruyter grant the dissemination of their titles much more freely than others. This is of course a limited selection of publishers to analyze, but it serves to illustrate the wide variety of situations encountered when dealing with journal publishers. It bears restating that some of the largest publishers, Wiley-Blackwell, Springer and Elsevier, were not included at all in this study since current content in their titles is distributed exclusively through their own databases.

Another factor influencing the results of this study is the collection development efforts previously undertaken at IUB Libraries. The institution has been slowly but surely converting a great deal of its journal subscriptions from print to electronic over the last decade. In the last several years such efforts have accelerated as the libraries attempt to provide patrons with the best available resources. This drive, combined with increasing pressure to trim budgets and become more efficient may have already eliminated a great deal of duplicate holdings that might have once existed. Had a similar study been conducted three or four years ago one likely would have found very different results.

The choice of how to define "current" coverage is another that is open to interpretation. While being within one issue may be reasonably current for most disciplines, for others absolute currency might be critical. Had the author chosen to define currency in a different way the results would have changed accordingly. In this we see another of the inherent limitations of

studies such as this- the variables selected represent one particular framework, and thus the results gathered here will not be directly applicable to other situations.

Examining the results from the database side, it should come as no surprise that the greatest amount of titles were found in the largest and most wide-ranging database: Academic Search Premier. While the selection of databases was based solely upon usage, it is apparent that some of the selections have a very narrow focus, and as such perhaps were not the most applicable for a non-field-specific study such as this. The fact that LGBT Life contained none of the selected titles is not surprising, nor is this knowledge particularly informative to the researcher. While the aims of the study were exploratory rather than descriptive the inclusion of irrelevant data does limit the utility of the results. That said, the results show that studies such as this can be used to evaluate databases as well as journals. Collection managers looking to purchase specialized databases might be able to use studies such as this to determine how unique the contents of a given database are, and whether such a purchase is warranted. With larger databases it may be more desirable to have a large amount of duplication, as that might allow for greater savings in subscriptions. Studies such as this can clearly provide the data; how it is utilized is up to the librarian.

Conclusion

The basic objective of this study was simply to determine the degree of duplication present between e-journals and databases, which it has been able to do on a limited basis. This was not purely an intellectual exercise, however. The initial impetus for it revolved largely around budget pressures, and how libraries are forced to cope with them. Few librarians would say they want to cancel journal subscriptions, but the reality is that with price inflation and

budget reductions such actions are inevitable. What librarians need, therefore, are tools to help make such difficult decisions a little more manageable. Can overlap analysis such as this study provide such a tool? The results point to yes. Even within a limited sample one can clearly identify certain titles that have a high degree of duplication, and would thus be excellent candidates for subscription cancellation. Conducting such an analysis on the entirety of a journal collection would be extremely time-consuming and labor-intensive, but may be a worthwhile endeavor for libraries facing the budget axe. This study has identified several areas at IUB Libraries where further research is needed, notably in terms of academic disciplines. The generality of this study allowed for "big picture" results, but narrower focus is needed in future work. Subject area journal collections should be compared with subject-specific databases to gain more complete understanding of content duplication. Ultimately, such analysis may not directly lead to wholesale journal cancellations: relying upon databases in lieu of journal subscriptions is still a controversial measure. Such actions are increasingly necessary, however, and studies such as this will help librarians to negotiate the delicate balance between budget restrictions and patron needs.

References

Banks, R. A. (2006). Decision-making factors related to bibliographic database cancellation. Behavioral & Social Sciences Librarian, 25(1), 93-110.

Carlson, A., & Pope, B. M. (2009). The "big deal": a survey of how libraries are responding and what the alternatives are. *The Serials Librarian*, 57(4), 380-398.

Culpepper, J. C. (2000). Electronic resources: Just what is online?. *The Acquisitions Librarian*, 24, 21-28.

Hawbaker, A. C., & Wagner, C. K. (1996). Periodical ownership versus fulltext online access: a cost-benefit analysis. *The Journal of Academic Librarianship* 22(2), 105-109.

Kalyan, S. (2002). Non-renewal of print journal subscriptions that duplicate titles in selected electronic databases: a case study. *Library Collections, Acquisitions, & Technical Services*, 26, 409-421.

Maple, A., Wright, C. & Seeds, R. (2003) Analysis of format duplication in an academic library collection. *Library Collections, Acquisitions, & Technical Services,* 27, 425-442.

McKay, S. C., Mouw, J. P. & Kusma, T. (Workshop Leaders), Hurd, S. H. (Moderator), Sonberg, P. (Recorder) (1998). Partnering in a changing medium: the challenges of managing and delivering e-journals. *The Serials Librarian* 34(3), 295-299.

Nixon, J. M. (2010). A reprise, or round three: Using a database management program as a decision-support system for the cancellation of serials. *The Serials Librarian*, 59, 302-312.

Rupp-Serrano, K., Robbins, S., & Cain, D. (2002). Canceling print in favor of electronic: criteria for decision making. *Library Collections, Acquisitions, & Technical Services*, 26, 369-378.

Schroeder, J. (2008). An overlap analysis of occupation therapy electronic journals available in full-text databases and subscription services. *Journal of Electronic Resources in Medical Libraries*, 5(4), 346-361.

Smith, D. A. (2009). A journal back file overlap analysis: looking back to move forward. *Library Collections, Acquisitions, & Technical Services*, 33, 25-30.

Smith, P., & Cochenour, D. (2007). Predicting serials inflation rates in the era of electronic resources. *Against the Grain*, 19(3), 72-74.

Sprague, N., & Chambers, M. B. (2000). Full text databases and the journal cancellation process: A case study. *Serials Review* 26(3), 19-31.

Tenopir, C. (1997). The complexities of electronic journals. *Library Journal* 122(2), 37-38.

Torbert, C. (2009). Purchasing of electronic backfiles: results of a survey. *The Serials Librarian* 57(4), 410-422.

Zappen, S. H. (2010). Managing resources to maximize serials access: the case of the small liberal arts college library. *The Serials Librarian* 59(3), 346-359.

Appendix A: Selected e-journal title list (with corresponding sample number)

9	ACTA POLITICA
10	ACTA SOCIOLOGICA
11	ACTION LEARNING RESEARCH AND PRACTICE
13	ACTIVE LEARNING IN HIGHER EDUCATION
28	AFRICAN STUDIES
36	ALTERNATIVES
38	AMBIX
45	AMERICAN ECONOMIC REVIEW
50	AMERICAN INDIAN CULTURE AND RESEARCH JOURNAL
52	AMERICAN JOURNAL OF BOTANY
53	AMERICAN JOURNAL OF DISTANCE EDUCATION
54	AMERICAN JOURNAL OF EDUCATION
58	AMERICAN JOURNAL OF HEALTH EDUCATION
71	AMERICAN JOURNAL OF THEOLOGY PHILOSOPHY
73	AMERICAN LITERARY REALISM
78	AMERICAN SOCIOLOGICAL REVIEW
83	ANALES CERVANTINOS
98	ANNUAL REVIEW OF EARTH AND PLANETARY SCIENCES
99	ANQ
105	ANUARIO DE ESTUDIOS MEDIEVALES
108	APPLIED LINGUISTICS
109	APPLIED MEASUREMENT IN EDUCATION
111	APPLIED SPECTROSCOPY
117	ARCHIVES OF CLINICAL NEUROPSYCHOLOGY
126	ARIZONA QUARTERLY
129	ARTS & HUMANITIES IN HIGHER EDUCATION
140	ASPASIA
143	ASTROBIOLOGY
148	AUSTRALASIAN JOURNAL OF PHILOSOPHY
153	AUSTRALIAN JOURNAL OF POLITICAL SCIENCE
159	BEHAVIOR MODIFICATION
165	BEHAVIOUR
168	BELFAGOR
173	BINOCULAR VISION AND STRABISMUS QUARTERLY
185	BOOKBIRD
187	BRAIN :
192	BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE
195	BRITISH JOURNAL OF CRIMINOLOGY
213	BULLETIN OF HISPANIC STUDIES
216	BULLETIN OF THE CHEMICAL SOCIETY OF JAPAN

221	BUSINESS AND SOCIETY
230	CAMBRIDGE OPERA JOURNAL
241	CANADIAN JOURNAL OF PROGRAM EVALUATION
242	CANADIAN JOURNAL OF ZOOLOGY
251	CELL BIOLOGY INTERNATIONAL
256	CHALLENGE
257	CHANGE
267	CHLOE
273	CLASSICAL ANTIQUITY
284	CLINICAL SUPERVISOR
303	COMMUNICATION MONOGRAPHS
313	COMMUNITY DEVELOPMENT JOURNAL
314	COMPARATIST
326	COMPUTER MUSIC JOURNAL
338	CONTEMPORARY PHYSICS
343	CONTINUITY AND CHANGE
349	CRIME, MEDIA, CULTURE
350	CRITICAL ARTS
351	CRITICAL ASIAN STUDIES
353	CRITICAL INTERVENTIONS
367	CULTURAL AND SOCIAL HISTORY.
374	CURRENT EYE RESEARCH
380	DEFENCE STUDIES
387	DEVELOPMENT
407	EARLY AMERICAN LITERATURE
436	EDUCATION FOR INFORMATION
442	EDUCATIONAL GERONTOLOGY
450	EIGHTEENTH CENTURY
454	EIRE IRELAND
465	ENGLISH LITERATURE IN TRANSITION 1880-1920
470	ENGLISH WORLD WIDE
473	ENVIRONMENT AND DEVELOPMENT ECONOMICS
476	ENVIRONMENT AND PLANNING D SOCIETY SPACE
477	ENVIRONMENT AND URBANIZATION
481	ENVIRONMENTAL HAZARDS
491	ESQ
496	ETHNOS
504	EUROPEAN JOURNAL OF OPHTHALMOLOGY
505	EUROPEAN JOURNAL OF PUBLIC HEALTH
515	EUROPEAN SOCIOLOGICAL REVIEW
518	EVALUATION
522	EVOLUTIONARY ECOLOGY RESEARCH
526	EXPLICATOR

737 INTERNATIONAL PLANNING STUDIES

754 IRISH STUDIES REVIEW

761 ITALIAN CULTURE

762 ITALIAN STUDIES

764 ITINERARIO

766 IZVESTIYA MATHEMATICS

768 JAHRBUCH FUR WIRTSCHAFTSGESCHICHTE

773 JEWISH QUARTERLY REVIEW

780 JOURNAL FOR SPECIALISTS IN GROUP WORK

782 JOURNAL FOR THE STUDY OF JUDAISM IN THE PERSIAN, HELLENISTIC, AND ROMAN PERIOD

783 JOURNAL FOR THE STUDY OF THE NEW TESTAMENT

788 JOURNAL OF ADVERTISING

795 JOURNAL OF AGING AND SOCIAL POLICY

805 JOURNAL OF APPLIED COMMUNICATION RESEARCH

814 JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY

816 JOURNAL OF BALTIC STUDIES

819 JOURNAL OF BIOLOGICAL CHEMISTRY

845 JOURNAL OF COLLEGE STUDENT RETENTION

853 JOURNAL OF CONTEMPORARY AFRICAN STUDIES

867 JOURNAL OF DRUG EDUCATION

873 JOURNAL OF EARLY MODERN HISTORY

880 JOURNAL OF ECONOMIC METHODOLOGY

885 JOURNAL OF EDUCATIONAL AND PSYCHOLOGICAL CONSULTATION

886 JOURNAL OF EDUCATIONAL COMPUTING RESEARCH

893 JOURNAL OF ENDOCRINOLOGY

908 JOURNAL OF FEMINIST STUDIES IN RELIGION

916 JOURNAL OF GENDER STUDIES

918 JOURNAL OF GENETIC PSYCHOLOGY

921 JOURNAL OF GLACIOLOGY

930 JOURNAL OF HISPANIC HIGHER EDUCATION

949 JOURNAL OF ISLAMIC LAW CULTURE

955 JOURNAL OF KNOT THEORY AND ITS RAMIFICATIONS

964 JOURNAL OF LEGAL STUDIES

966 JOURNAL OF LEISURE RESEARCH

968 JOURNAL OF LIBRARY ADMINISTRATION

971 JOURNAL OF LITERACY RESEARCH

981 JOURNAL OF MATERIAL CULTURE

982 JOURNAL OF MATERIALS RESEARCH

998 JOURNAL OF MOLECULAR MICROBIOLOGY AND BIOTECHNOLOGY

1000 JOURNAL OF MOTOR BEHAVIOR

1008 JOURNAL OF NEUROSCIENCE

1010 JOURNAL OF NIETZSCHE STUDIES

1018	JOURNAL OF PACIFIC HISTORY
1024	JOURNAL OF PEDIATRIC PSYCHOLOGY
1028	JOURNAL OF PETROLOGY
1037	JOURNAL OF PLANNING EDUCATION AND RESEARCH
1046	JOURNAL OF PUBLIC ADMINISTRATION RESEARCH AND THEORY
1047	JOURNAL OF PUBLIC CHILD WELFARE
1055	JOURNAL OF REVENUE AND PRICING MANAGEMENT
1057	JOURNAL OF SEMANTICS
1060	JOURNAL OF SOCIAL AND CLINICAL PSYCHOLOGY
1067	JOURNAL OF SPANISH CULTURAL STUDIES
1081	JOURNAL OF SUSTAINABLE FORESTRY
1082	JOURNAL OF SYMBOLIC LOGIC
1099	JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR
1102	JOURNAL OF THE HISTORY OF IDEAS
1103	JOURNAL OF THE HISTORY OF MEDICINE AND ALLIED SCIENCES
1107	JOURNAL OF THE INTERNATIONAL PHONETIC ASSOCIATION
1111	JOURNAL OF THE SOCIETY FOR AMERICAN MUSIC
1113	JOURNAL OF THE VIRTUAL EXPLORER
1115	JOURNAL OF THEOLOGICAL STUDIES
1118	JOURNAL OF TROPICAL ECOLOGY
1131	KANT STUDIEN
1135	KONSTHISTORISK TIDSKRIFT
1141	LAKE AND RESERVOIR MANAGEMENT
1146	LANGUAGE AND SPEECH
1149	LANGUAGE TESTING
1151	LARES
1156	LEARNING & BEHAVIOR
1166	LIBRARY & INFORMATION HISTORY.
1171	LIBRI
1172	LIGHTING, RESEARCH, & TECHNOLOGY
1177	LINGUISTIC TYPOLOGY
1182	LITERATURE & THEOLOGY
1183	LITERATURE AND MEDICINE
1188	LOW TEMPERATURE PHYSICS
1200	MAS?Q
1204	MATHEMATICAL RESEARCH LETTERS
1207	MATHEMATIKA
1213	MEDIEVAL ARCHAEOLOGY
1225	MICROBIOLOGY
1236	MLN
1237	MNEMOSYNE
1259	MOTOR CONTROL
1262	MUSEON

1280	NARRATIVE INQUIRY
1286	NATURE
1289	NEURO OPHTHALMOLOGY
1304	NEW ZEALAND JOURNAL OF BOTANY
1308	NINETEENTH CENTURY THEATRE AND FILM
1313	NORDIC JOURNAL OF LINGUISTICS
1320	NOVUM TESTAMENTUM
1324	NURSING SCIENCE QUARTERLY
1329	OEUVRES CRITIQUES
1339	OPTICAL ENGINEERING
1341	OPTOMETRY TODAY
1351	OXFORD ECONOMIC PAPERS
1353	OXFORD LITERARY REVIEW
1357	PACIFIC REVIEW
1366	PARAGRAPH
1371	PEACE REVIEW
1373	PEDAGOGY CULTURE AND SOCIETY
1375	PERCEPTION
1380	PERSPECTIVES ON POLITICAL SCIENCE
1391	PHILOSOPHY EAST WEST
1392	PHILOSOPHY OF SCIENCE
1396	PHRONESIS
1399	PHYSICAL REVIEW.
1411	PLANTA MEDICA
1417	POLICY AND PRACTICE IN HEALTH AND SAFETY
1435	PRIMERA REVISTA LATINOAMERICANA DE LIBROS
1436	PROBATION JOURNAL
1441	PROCEEDINGS OF THE EDINBURGH MATHEMATICAL SOCIETY
1442	PROCEEDINGS OF THE LONDON MATHEMATICAL SOCIETY
1444	PROFESSIONAL SCHOOL COUNSELING
1448	PSYCHOANALYSIS AND HISTORY
1457	PUBLIC LIBRARY QUARTERLY
1462	PUBLIC WORKS MANAGEMENT AND POLICY
1465	PUBLICATIONS OF THE MODERN LANGUAGE ASSOCIATION OF AMERICA
1475	QUANTITATIVE FINANCE
1484	QUEST
1488	RATIONALITY AND SOCIETY
1489	READING PSYCHOLOGY
1494	REINARDUS
1495	RELIGION
1510	RESTAURATOR
1514	RETINA
1518	REVIEW OF ECONOMICS AND STATISTICS

1520	REVIEW OF EDUCATIONAL RESEARCH
1525	REVIEW OF INTERNATIONAL POLITICAL ECONOMY
1535	REVISTA DE HISTORIA ECONOMICA
1538	REVISTA DE MUSICA LATINOAMERICANA
1544	REVUE ROMANE
1556	RUSI JOURNAL
1567	SCHOLARLY COMMUNICATIONS REPORT
1574	SCIENCE FICTION FILM AND TELEVISION
1576	SCIENCE MAGAZINE
1585	SEEING AND PERCEIVING
1586	SEFARAD
1594	SEVENTEENTH CENTURY
1597	SHAW
1601	SIMULATION & GAMING
1605	SLAVONICA
1611	SOCIAL COGNITIVE AND AFFECTIVE NEUROSCIENCE
1621	SOCIAL RESEARCH
1622	SOCIAL SCIENCE COMPUTER REVIEW
1632	SOCIETY FOR AMERICAN MUSIC.
1633	SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS.
1639	SOCIOLOGY OF SPORT JOURNAL
1644	SOUTHERN CULTURES
1650	SPORTS MEDICINE
1659	STUDI MUSICALI
1660	STUDIA IRANICA
1661	STUDIA MUSICOLOGICA
1662	STUDIA NEOPHILOLOGICA
1663	STUDIA SLAVICA
1668	STUDIES IN CONFLICT AND TERRORISM
1669	STUDIES IN EIGHTEENTH CENTURY CULTURE
1671	STUDIES IN LANGUAGE
1673	STUDIES IN SCIENCE EDUCATION
1676	SUICIDE LIFE THREATENING BEHAVIOR
1679	SYMBOLAE OSLOENSES
1689	TATE ETC
1691	TEACHER EDUCATION AND SPECIAL EDUCATION
1692	TEACHER EDUCATOR
1696	TECHNICAL SERVICES QUARTERLY
1698	TECHNOLOGY ANALYSIS STRATEGIC MANAGEMENT
1699	TECHNOMETRICS
1700	TECTONICS
1704	TESOL JOURNAL
1707	TEXT

Appendix B: Selected databases with Sept.-Dec. 2010 usage data

Database	Usage
Academic Search Premier	38,265
JSTOR	19,647
ABI/INFORM Global	9,898
PsycARTICLES	8,695
Journals@Ovid	8736
Business Source Premier	5,314
Education Full Text	4,821
LexisNexis Academic	4,058
PubMed Central	3,371
Communication & Mass Media Complete	2,961
SPORTDiscus with Full Text	1,326
Emerald Management Plus	815
CINAHL Plus with Full Text	783
Health Reference Center Academic	731
Art Full Text	561
ulrichsweb.com	520
Health & Wellness Resource Center	504
BioOne.1	498
Library Literature & Information Science Full Text	485
Business & Company Resource Center	458
LGBT Life with Full Text	405
Corporate ResourceNet	350
MasterFILE Premier	338
Ethnic NewsWatch	293
Military and Intelligence	287
MIT CogNet	263
GenderWatch	252
GREENR (Global Reference on the Environment, Energy, and Natural Resources)	209

Appendix C: Fields of study present in sample grouped by academic discipline

Formal Sciences:
Computer Science
Information Systems
Mathematics

Humanities:
Classics
Comparative Literature
English
Fine Arts
Folklore
German
History
History and Philosophy of Science
Linguistics
Music
Philosophy
Religious Studies
Spanish
Theater

Natural Sciences:
Astronomy
Bioscience
Chemistry
Geology
Physics

Professional and Applied Sciences:
Business
Communications
Criminal Justice
Education
Health
Journalism
Library & Information Science
Medicine
Nursing
Optometry
Public & Environmental Affairs
Public Health
Social Work
Sports

Social Sciences:
African Studies
Anthropology
East Asian Studies
Economics
Gender Studies
Geography
Jewish Studies
Latin American Studies
Middle Eastern Studies
Political Science
Psychology
Slavic Studies
Sociology
West European Studies