Data Driven Collection Assessment using a Serial Decision Database

By:

Diane Carroll, Owen Library 135, Head of Collections, Washington State University, PO Box
643200, Pullman, WA 99164-3200; carroldi@wsu.edu, (509) 335-7151 (Phone), (509) 335-2534
(Fax)

Joel Cummings, Owen Library 139, Collection Manager for Agriculture, Engineering and the
Sciences, PO Box 643200, Pullman, WA 99164-3200, jcummings@wsu.edu, (509) 335-6493
(Phone), (509) 335-2534 (Fax)

Abstract

Washington State University (WSU) Librarians expanded the concept of a
comprehensive serials decision database (SDD) first proposed by Metz and Cosgriff[1] in 2000.
Information about subscribed and unsubscribed serials was merged from the integrated library
system and subscription agent, interlibrary and citation databases, as well as journal use sources
onto one spreadsheet. The SDD, used for selection and cancellation projects, evaluation of
electronic journal packages, and collection assessment, was recognized as a necessary serials
management tool for WSU. There is a need for the commercial development of products that will
replace this manual way of organizing information critical to collection development.

THE PROBLEM

With the continuing financial crises looming over university budgets and library funding in particular, collection budgets have become increasingly at risk for funding cuts. Every dollar needs to be accounted for and provide evidence that the library collections funds were used in a wise and prudent manner. In the past decades, an increasing amount of data relevant to serials collection management has become available to librarians that can assist in providing the evidence needed to defend collection budgets.[2] Metz and Cosgriff [3] explained how they merged information from faculty surveys, citation analysis, and other measurements of print usage into a Comprehensive Serials Decision Database which resulted in bringing "together everything a library knows or can learn about its resources." Gallagher et al (2005)[4] described their efforts to evaluate the value of journals at Yale University by combining data from disparate sources. They concluded that "It is crucial…for libraries to determine how best to manage these resources, since failure to do so will result in the continued inefficient use of resources, and an underserved user-base whose perception of the value of the library will diminish." This project illustrated the need for a systematic approach to combining relevant collection development data from many sources for use by collection development librarians.

PRODUCTS AVAILABLE FOR SERIALS ASSESSMENT

This need for more complete analysis of serial data elements was described in 2009 by Robin Paynter. Paynter analyzed four commercial decision support systems, which like the SDD would enable library staff to spend more energy analyzing data and making decisions rather than attempting to extract and collect data: Serials Solutions' 360 Counter®, Thomson Reuters Journal Use Reports (JUR), Swets' ScholarlyStats and Ulrich's Serial Analysis System (USAS) ™.[5] She described in detail the usage reports supported, the types of analyses, ability to upload custom data, reporting features and product support.

A successful commercial product would allow the merging of generic serials knowledge base information with customized information from the local library for titles both subscribed and unsubscribed regardless of whether they have been assigned an ISSN. Evaluating these commercial products on just these two criteria, ScholarlyStats is useful as a source of use data, but no other customized data can be uploaded. Use data can be uploaded into JUR and combined with publication activity reports but only for titles index in the Web of Science. USAS[TM]'s strengths is it contains serials information from the serials knowledge base Ulrich's Periodical Directory and that it allows uploading of three columns of data. Its weakness is only three columns are available to upload customized local information and titles without ISSN cannot be matched and data uploaded. Serials Solutions' 360 Counter ® is also a source of use data that can be uploaded by the library staff or this service can be purchased via the Data Retrieval Service. Serials Solutions' 360 Core® is a serials knowledge base which allows local holdings to be entered by provider. A combination of USAS[TM], acquired by Serials Solution in 2007, and Serials Solutions' 360 Counter® products with customized reporting has possibilities of creating a serials decision support system[6].

New generations of Electronic Resource Management (ERM) systems also show promise of creating improved serials collection assessment tools. Subscription services have extensive serial knowledge bases and the ability to add use data. SUSHI (**S**tandardized **U**sage **S**tatistics **H**arvesting **I**nitiative) protocol has been incorporated into the COUNTER Code of Practice (Project COUNTER)[7]. Implementation will allow the automated retrieval of the Project COUNTER compliant usage reports into many products such as ScholarlyStats, Serials Solution's 360 Counter®, subscription services and ILS systems, etc. For example, Swets who acquired ScholarlyStats in 2007 has integrated use data into its new ERM service. This system is reported to "Manage licenses and evaluate acquired content, calculate and analyze the cost per use of electronic journals quickly and efficiently and provide full text access to your end users."[8] EBSCO recently announced EBSCONET® ERM ESSENTIALS[TM] that it is designed to locate

subscription history, details for electronic journal packages, review publishers' standard license details and terms of use. EBSCONET®Usage Consolidation service is currently under development. It will load Project COUNTER compliant use data using SUSHI to the ERM thereby allowing for cost per use calculations.[9]

Products currently on the market lack the functionality to produce reports that include title, journal coverage years and download use data by provider. Without the coverage years, the librarian may not know the subscription status of the titles. Without the provider, a librarian may not know if the title is a paid subscription or part of an aggregator collection. Librarians managing a collection need a list of subscriptions that reveals all of the nuances of subscription status, post-cancellation rights, coverage, format and provider in order to make effective decisions about selections, cancellations, formats, and providers.

SIMULTANEOUS DEVELOPMENT OF SERIAL ASSESSMENT DATABASES

Oregon Health & Science University (OHSU) librarians built an SDD based on Metz and Cosgriff's work in 2000[10] to assess the collection and change from print to electronic format. An overview of the data included in this spreadsheet and how the use of this information evolved through five annual updates (2001 to 2005) was report at the Online Northwest Conference (2006)[11]. This database provided objective information that allowed the OHSU Library Collection Development Committee to make subjective decisions regarding journal matters.[12]

In 2003, Galbraith[13] discussed the development at WSU of the Collection Management Access Database that allowed librarians to make data driven journal retention and cancellation decisions for the Owen Science and Engineering Library and the Fischer Agricultural Sciences library. Procedures on how to download order record information from Innovative Interfaces, Inc and standardize the data were developed by WSU Library collections and systems staff. However, as data for journal assessment became available from more sources, it was discovered that information processed via this method had been superseded by the need for a model that

would be able to incorporate the increasing complexity of serials management. Therefore in 2007, the first SDD created for the WSU Libraries was a hybrid of the OHSU SDD and WSU Collection Management Database model. This article describe in detail the construction of WSU's 2009 version of the SDD and illustrates how this system has proven itself to be an invaluable asset in shaping and managing the serials collection.

Creating the SDD is a labor intensive process. It is the hope of the authors that those developing future ILS or serials management products will understand the need for a comprehensive serials decision support system. The ability to upload local data, match it to general information from a knowledge base, and create complex reports would make this process much easier and less time consuming for individual libraries.

CREATING THE DATABASE

The following section describes the creation of WSU's 2009 SDD. Detailed procedures, a sample SDD (18 titles) and copies of the Microsoft Excel (MS Excel) macros used to create the SDD can be found on the WSU Research Exchange[14].

Tools

MS Excel was chosen to create the SDD because of its functionality and familiarity. In 2006, Nick Peterson of the OHSU Library designed a merge macro[15] which allows elements to be merged from two worksheets onto one by matching any data element, such as an ISSN or title. In 2007, Cindy Ellis of the WSU Libraries Systems Department created a macro to remove initial articles, such as "A, An, and The," and to remove periods from the end of the titles exported from the Integrated Library System (ILS).[16]

Sources of data

The 2009 WSU SDD was created by merging journal information from a variety of sources:

1. Integrated Library System (ILS) contributed data from the bibliographic, order, serial, and the Electronic Resource Management (ERM) module including the coverage database and resource and license records (Innovative Interfaces, Inc.).

2. Journal use data was retrieved from locally collected print usage for some titles, from commercially collected COUNTER download usage (Swet's ScholarlyStats), or directly from electronic journal providers.

3. Subscription and general information about journals was collected from serial knowledge bases, such as serials the subscription agent's database (EBSCONET®) and USAS^{TM}.

4. The ILLiad^{TM} system contributed title, ISSN and number of requested non-returnable interlibrary loans (ILL).

5. Citation databases provided data on the number of times WSU authors had written articles in a particular journal or had cited a journal over a specific period of time (Web of Science®). Also, data from a locally compiled list of authored papers from a WSU annual reports database was included. The latter data was collected from WORQS (WSU's Online Review and Query System) – the annual review software used across the WSU campuses.

6. Calculated values included total use; cost per use; percentage of downloads from aggregators, and priority assignments based on usage.

Results of loading data sources

Data was merged into the 2009 WSU SDD in the following order: ILS records (bibliographic, order, serials, ERM); subscription agent data; citation databases data; ILL data; journal use data; and USAS^{TM}. The ILS was the core of the SDD and other information was added to that initial list of titles. Starting with 4,251 titles culled from the ILS, the SDD grew to 16,158 titles when all sources of data were added. Table 1 lists the data sources, the number of titles added as each source was merged and the table location on the sample SDD (Tables 2 and 3).

DATA ELEMENTS

Many of these elements came from more than one source. Some of the challenges involved with collecting and merging the data in the SDD are addressed in this section.

Journal titles

Current subscriptions were contributed to the SDD from ILS reports, print and online download usage reports and lists of current subscriptions from the serial agent knowledgebase. Sources of subscribed and unsubscribed titles included journals requested on ILL, journals with WSU authored papers or journals cited by WSU authors from the citation databases. Collecting from these sources resulted in the inclusion of journals of interest to the user community regardless of whether the library subscribes to a title or not (Table 2A).

ISSN

The unique identifier used most frequently to merge data in the SDD was the ISSN. The mixing of print (pISSN) and electronic ISSN (eISSN) in the same MARC field made use of the ISSN from the bibliographic record problematic. The SDD has a pISSN and eISSN column to help with the matching, but frequently titles were merged one by one because no ISSN was assigned or print and electronic ISSNs were undifferentiated in a single column from titles being added (Table 2A).

Acquisitions information

Order records included the subscription status, format of journal and location, and accounting information such as cost, vendor, invoice paid date and fund code. Using the Innovative Interface's report export function, titles were selected that had an invoice posted within the specified time period. However, because the results included invoices posted from other time periods, the exported spreadsheet had unwanted information and jumbled columns. The ability to limit payments to just those made within a particular time period was found only in the Innovative's legacy character-based system (Innopac). The exported spreadsheet had data for just the desired time period in correctly labeled columns (Table 2A and 2B).

Subscription status

The subscription agent database identified print titles that were part of memberships or print packages (Table 3A). Many electronic journals had a serial record in the ILS but did not have an order record because electronic journal packages had been posted on one order record which covered many titles. Therefore, coverage dates from the ERM were used to determine the subscription status of many electronic journals. After coverage dates were added to the SDD, titles without order records were assigned a subscription status. Access to titles that were parts of consortial agreement shared title agreements were assigned "Consortial" instead of "Current." Titles where current access (defined as less than one month embargo) was only from an aggregator were labeled "Current-Aggregator."

Journal usage

Title coverage dates from the ERM and downloads use data were merged by provider and loaded into the SDD on the same row allowing the total uses to be calculated (Table 2C). Merging title, coverage and download use by provider was the single most time-consuming activity in creating the SDD.

Only one collection unit in the WSU Libraries counts in-house print journal usage. To gain additional information on print usage, the cumulative number of times a bound journal were checked out was collected and divided by six years, the number of years circulation data had been collected in the ILS (Table 3B).

Interlibrary loans

A spreadsheet of title, ISSN and publication year of journal article interlibrary loans that were requested over five times in 2008 was separated by publication year of the requested article – 1997 to present and before 1997 (Table 3B). This data was loaded in the SDD in two columns

so the importance of acquiring a current subscription could be differentiated from the need to acquire archival access.

Authored and Cited Papers

The number of articles authored by WSU scholars was retrieved by searching the Web of Science® (including Science Citation Index Expanded, Social Sciences Citation Index and Arts & Humanities Citation Index) for Washington State Univ or 99164 in the address field. Both the journal and the cited references were included in the exported spreadsheet. Three years of authored and cited articles (2006-2008) were included in the 2009 WSU SDD to provide current data because a single year's was thought to be too short of a time frame for many titles. All journals that contained WSU authored articles were included in the SDD, while only journals with five or more references cited by WSU authors were included in the SDD (Table 3B).

WSU also collected references to authored papers during the institution's annual faculty review process (Table 3B). This data was added to the SDD for 2005-2008 without removing duplicates between local data and Web of Science® data or adjusting for duplicate reporting from multiple authors at WSU. This information was valuable as not all WSU authored articles were located in journals indexed in the Web of Science®.

Calculated values

In the SDD, all uses (ILL, Web of Science®, print use, electronic article downloads) were totaled in one column. The cost was divided by total uses (Table 3C). When a title had no uses, the cost per use was listed as the full cost of the title. The percentage of download use from aggregators was calculated by summing the uses from aggregator collections divided by total uses and the percentage of use from JSTOR was calculated by dividing JSTOR use by total uses.

Priority assignment

Titles were assigned to a high, medium or low use category or no use data (priority 1, 2, 3 or 4 respectively; see Table 4). To start, all titles were assigned a priority 4 and then any individual data element raised the title's priority to a 1, 2 or 3 (Table 3C). To illustrate, a journal may be a Priority 1, if it exceeded one of the following thresholds: 15 interlibrary loan requests, 250 downloads, 40 WSU cited papers or 10 WSU authored papers.

MAINTAINING THE SERIALS DECISION DATABASE

The first time the SDD was created at WSU has been described in the previous section. The development followed this time table. In March 2006, a meeting of collection managers to review data elements to be included was held. Data specific to WSU was reviewed and priority parameters were established. The Head of Collections worked with WSU staff to learn how to download order records from Innovative and to use the wizards developed there. Procedures were written for all new processes. The first version was completed by one person and was ready for review in July 2006.

Maintaining and updating the SDD occurs throughout the year. For example, in 2010 the following occurred.

- January – Replaced WSU authored and cited papers with the most current three year (2007 to 2009); replaced interlibrary loan, circulation and print use data from 2008 with 2009 data

- February and March – Matched electronic journal 2009 use data with coverage years by title and provider and merged into the SDD; calculated cost/use with 2009 costs; total uses and assign priority; added subject headings, list price and publisher from USAS^{TM}

- April, May and June-SDD was available for review by collection managers and selectors with 2010 price estimates

- July – Finished merging all order records with payment information from FY08-09 and close the SDD for 2009. Renamed the file and created a column for cost estimates for

2010. Estimated all costs based on knowledge of contracts and inflation estimates. Created a column for actual costs that will replace estimated costs as invoices are received throughout the year. Collection managers used the estimated costs when working to meet a cancellation target.

- August-Collection managers and selectors presented information to faculty on possible cancellations for 2010, format changes or additions; Serial vendor information was added to the SDD.

- September-Decisions on all changes in format or cancellations/additions were given to the Head of Collections on a copy of the SDD. Detailed information from the SDD with the changes was given to the acquisitions staff by September 15.

- October 1-Acquisition staff submitted changes to serials vendor or directly with publisher.

- November and December – The serials staff managed all changes resulting from print cancellations and additions; the Head of Collections used the SDD to report to the ERM team to close a holding for an electronic journal or remove the title because there were no post-cancellation rights.

Activities with the SDD that occurred throughout the year included loading order records with new payment information at least quarterly and more often towards the end of the fiscal year It was also used to keep notes in a comment column of collection recommendations such as title transfers, the ability to cancel print because post cancellation rights have been obtained, etc.

Technical services staff members have assisted with harvesting data from other sources and merging electronic journal use data with coverage data. This information was given to Diane Carroll who merged the data into the SDD. The SDD is backed up frequently but the master copy is held by one person. The macros and wizards have increased productivity and time needed to update the SDD; however, it is very labor intensive. All of March work time was spent on

meeting the deadline of April 1 to finish the SDD. Merging the electronic journal use and

holdings was in essence an audit of the collection. Data can be manipulated using MS Excel

filters that revealed information that needed to be investigated and reported. For example, prices

that were much higher than estimated, unintentional duplicate copies, title with stats that were not

loaded in ERM, new titles added as part of memberships, titles now available online that were not

listed in our ERM, etc. Updating the SDD often gets side tracked into investigating these

situations since it is unlikely they will come to light in other ways. This can significantly add to

time included in updating the SDD, however, acquisitions staff report that problems revealed in

the SDD were rarely duplicates of problem they discover using their own quality control

procedures.

SPREADSHEET VERSUS RELATIONAL DATABASE

The 2009 SDD was a single spreadsheet of serial titles listed in rows with consistent

metadata. It permitted users with a basic understanding of MS Excel to quickly obtain data about

subsets of the entire collection by an elementary combining of filters on available columns. This

capability enabled librarians to repurpose subsets of the data contained in the SDD for a variety of

uses. Additionally for the more advanced user, the simple structure of the database makes it

elementary to convert the SDD into a form that was usable by MS Excel's pivot table and

charting functions. Pivot tables and charts allow the SDD users to analyze data from different

columns against each other and perform a wide variety of calculations on that data and display

this data in summary charts and tables. This simple structure also allows for additions and

deletions of columns, or new data elements, as needed. For example, in 2009 iteration of the SDD

several columns of data from the serials record relating to binding were added to assist the

Preservation and Binding task-force to analyze and make binding decisions.

A relational database such as MS Access has advantages in its ease of merging data;

however, merging on ISSN or titles regardless of whether using a spreadsheet or relational

database is prone to errors due to confusion over print and eISSNs, absent data, ISSN changes over time, and other issues. Being able to view the data on one spreadsheet allows for quality control checks and helps to find errors when matching the newly added information. The SDD master copy was updated and backed up by one person while other used copies for analysis. While MS Excel and MS Access, enable users to manipulate detail in sophisticated ways, MS Excel was chosen since it would present a shorter learning curve for employees, be easier to use, and therefore be of greater value to the libraries[17].

Minor problems that require careful attention when analyzing data in the SDD emerged. For example, subscription information frequently was not wholly independent of other titles on the list. For example, many individual serial titles were included in packages (i.e. memberships, 'big deals', parent-child bundles, etc.) Therefore, cost, usage data, and other information about individual titles may not appear in the same row as the title, but rather with the parent title for the package.

HOW WAS THE DATABASE USED?

The SDD has proven itself to be an invaluable asset in shaping the collections, making individual selection and cancellation decisions, managing the budget, marketing and assisting with serials management.

Cancellations

The primary use of the SDD by collection development librarians has unfortunately focused on annual serial cancellation projects. Following are two examples of how the SDD was customized for presentation to library selectors and faculty from the science, engineering and agriculture and for the humanities and social sciences.

In the first example, the science collection manager developed a preliminary cancellation list with extensive analysis of the SDD in consultation with science selectors. The selectors then

disseminated this information to faculty in relevant departments. First the titles and costs were itemized as part of the website introducing the cancellation process (Appendix 1).

An abbreviated version of the SDD was linked to the message sent to WSU Faculty so they could download and review more detailed data (Table 5). This spreadsheet displayed the title, current format, cost per use, percentage of download use from aggregators or JSTOR, citation data, print and electronic journal download usage, and subject from USAS^{TM}. An additional column was added by the collection managers to give a predicted fate on online access to titles in the advent of cancellations. In the cases where Ulrich's Subject field was blank, the collection manager manually added a subject term.

Given this information, individual faculty members in the relevant departments and colleges replied with comments and concerns about potential cancellations to the collection manager or individual selectors. Once this input had been collected from the Science, Engineering and Agriculture faculty, the science collection manager led a final cancellation meeting with the science selectors. The science collection manager gathered Journal Citation Report data and data from the SDD and by use of mail merge, formatted the data into a layout that was then converted indirectly into a PowerPoint presentation (Figure 1). The science selectors then used this collective session to make decisions about which serials to cancel. This group method has worked well for several years, and as the science selectors have appreciated being able to discuss journal issues as a group together. By using the SDD information in combination with faculty input and JCR data, they have been able to make thoughtful and well-rounded decisions.

In the second example, the humanities and social sciences collection manager provided faculty with a spreadsheet with two tabs or worksheets. The first worksheet labeled Master List, listed all titles in humanities and social science titles to which WSU Libraries had a subscription. In addition, faculty in each subject area was given a second worksheet of titles associated with subscription costs in their academic area for their review. Many titles were purchased as part of

14

packages or memberships and did not have an individual price. Faculty was concerned when they did not see those titles in their subject areas. Providing the big picture of all subscriptions on the Master List provided more complete information for the faculty about the entire collection. Faculty members were asked to evaluate the paid subscriptions in their subject areas.

Data and labels from the SDD were reorganized or simplified. The library location and format was altered to include information about back file access via JSTOR (Table 6). The priority column was relabeled to "Summary of usage information." Priority 1 and 2 were assigned "High Use," priority 3 to "low use" and priority 4 to "no use data". Subjects from USAS^{TM} were changed to the name of the department and a new column labeled "Priority" was added for input by faculty. Evaluation of the humanities and social sciences collection was completed in the spring. Allocations and cancellation targets were announced later, and the lower priority titles were reviewed as a cancellation list was developed.

In conclusion, these approaches provided a means to build consensus on challenging and sometimes emotionally charged cancellation decisions. Sharing portions of the SDD with faculty helped them to set their priorities while demonstrating that the library was making evidence-based decisions. Providing objective information in addition to input from faculty gathered for that cancellation project gave the librarians the ability to make difficult choices between journals with greater confidence.

Journal Selection

Unsubscribed titles were added to the SDD from WSU authored papers, WSU cited journals or articles borrowed on interlibrary loan. Article download data was collected for unsubscribed titles that were accessible by library users when IP ranges were registered for the product. For example, open access titles (BioMed Central), titles available via aggregators with or without embargoes, titles included in JSTOR, previously cancelled titles for which some online access was maintained by post cancellation rights, or titles for which there was a free trial period

were recorded in the SDD. Unsubscribed titles in Priority 1 were evaluated for addition to the collection.

The separation of interlibrary loan data by date (articles prior to 1997 and articles thereafter) allowed the selector to decide if a current subscription and/or online archive would satisfy the faculty's need for a given title.

Aggregators and Package Agreements

The SDD has provided the WSU Libraries with a simple method to examine data for aggregators and packages and to evaluate of the importance of consortial and other serial agreements. For example, librarians used the SDD extensively during the selection of a general full-text aggregator as part of the process conducted by the State of Washington's Co-operative Libraries Project (CLP). In this instance an extensive free trial of one aggregator and paid access to a second aggregator provided the WSU Libraries with data that enabled the librarians to evaluate these two major full-text databases. The difference in numbers of titles, or subjective analysis of the content of these different titles, was done by analyzing spreadsheets made available by the respective companies. An analysis of local usage and citation data allowed the librarians to make a decision using information gleaned from the SDD. The differences between the products as they related to WSU use made the potentially challenging job of consensus building within the libraries relatively easy – even when considering dramatic impacts the decision would have across the WSU library system. In this process, the SDD was used to identify high use titles to which the institution would lose online access to due to the change in aggregator. Selectors added subscriptions as thought necessary based on the analysis of the SDD data to ensure continued access to these titles and a smooth transition period.

Collection Assessment by Ownership and Format

The WSU Libraries has canceled approximately $750,000 of its subscriptions since the first SDD was developed in 2007. The library must support the information needs central to the mission of WSU and one way to accomplish that was to rely less on purchasing titles permanently and pursuing access to materials without permanent ownership. Those included open access journals, shared titles available through consortial agreements, and subscriptions to aggregator collections. "Permanent" was defined as titles where online or print access would be retained up to point of cancellation. Table 7 compares the download usage of titles in the 2009 SDD based on their format and ownership status. It includes only titles that were current subscriptions or had download use data (14,657 of 16,090 titles in the SDD). Print subscriptions were broken down into detailed categories of current print only, current print and online, current print and JSTOR archive, current print online access via aggregator collection. The term "other" included non-current titles with downloads uses where online access was either permanent or access only.

This analysis helped to answer difficult questions. WSU Libraries subscribed to 2,575 titles that have no online access at a cost of $377,709 (Table 7). When the print only titles were evaluated by assigned priority, only 129 of the titles were priority 1 or 2. They were the most vulnerable part of the collection and will most likely continue to appear in high numbers on potential cancellation lists. Only 531 titles were print and online subscriptions. Print was canceled for most titles that have post cancellation rights over the last eight years. However, an additional 643 print subscriptions have archival online access (JSTOR) or aggregator access. When a print subscription was not available online except for the archival access in JSTOR, the current print subscription became vulnerable to cancellation because of low use of the print or no record of use. The same was also true of the current print that was not available online except through an aggregator. Even though online access through aggregators was unstable, with budgets strained, many cancellation decisions made were the lesser of two evils. The SDD with the

priority analysis helped to change some lower use titles to access only while protecting those titles that are critical to the university' mission.

The permanent collection included online only titles that were 32% of the number of titles within the collection and 79% of the serials collection budget; print was 18% of the titles and 9% of the budget; and print and online were 8% of the titles and 9% of the budget (Table 7). The total for permanent access was 69% of the collection and 98% of the budget. Access only titles represented 31% of the titles and only 2% of the budget (Table 8). Not included in the cost was an aggregator package purchase with state fund through the CLP. The cost per download was $3.89 for titles permanently held and $0.15 for access only titles, however 70% of the uses came from permanently held titles at 30% from access only (Table 8). The role once filled by ILL just decades ago, where ILL was the nearly exclusive method of obtaining materials not subscribed to by an institutions library, now appears to be shared with access to sources of journal articles through aggregators and consortia agreements.[18]

Collection Assessment by WSU Authored and Cited Papers

Cost per download use for packages can be easily obtained by using a commercial service. With the SDD, a less common analysis of publisher titles and packages can be calculated: cost per cited paper. Table 9 was created from the 2010 SDD and reflects costs for 2009, and titles referenced by WSU authors in their 2007 to 2009 publications.   The total number of cited papers in the SDD was 127,596. Table 9 presents information on 78% of those papers. The average cost per cited paper was $32.

About 43% of the cited papers were from Wiley, ScienceDirect and SpringerLink and these ranged from $38 to $52 per cited paper. About 30% of the cited papers cost $10 or less per cited paper and included Open Access journals, titles from Highwire Press, AAAS, Emerald, a variety of society publications, the American Chemical Society, and BioOne.   Science and Nature were $5 per paper cited by a WSU author. Nature publications without the title "Nature"

were $65 per cited paper. Analysis of cost per referenced paper will be added to the decision making process for the next cycle of cancellations at WSU in FY11.

Budget

The SDD was useful in forecasting inflation based on the contract terms for titles within the collection instead of using a generalized inflation forecast for all academic serials. Customized cancellation targets were projected in August for the overall collection budget before renewals were submitted to the subscription agent in September. During the fiscal year, payment information replaced estimated costs and enabled the library staff to track expenditures and to predict problems that could, if not noticed ahead of time, become crises. Conversely the data could reveal opportunities before it was too late to take advantage of them. By employing subject data obtained via Ulrich^{TM}, the SDD can show collection development librarians where their funds are having the most impact and that information can be used to support requests for re-allocations or attempts to secure additional funding for these areas.

Above all else, the SDD enabled the library to explain its value to the University, demonstrating what the value the University gets for the money that is allocated for the library's collection's budget. It has been especially valuable in an era where the popular and even sometimes library personnel perception is that the value of the library is declining, when the data would strongly suggest that the value of the library to its academic users is increasing.

For example, in 2007, low use titles were identified in several package deals. A list of both the subscribed titles and shared access titles with citation and usage data was presented to the Deans at WSU. Upon seeing the numbers they contributed funding from their own budgets to allow the libraries to keep the packages for several years, thus saving several hundred titles from cancellation that year. This list produced by the SDD during an extremely tight budget year was easy to create and had a very high impact. Faculty appreciated receiving this straightforward, objective data.

Marketing

The SDD can also be used to inform the user community of high-use journals and the efforts made by the library to match their research, teaching and outreach needs. In August 2009 an article from *WSU Today* [19] was well received as faculty and other stakeholders became more aware of this data-driven approach to supporting WSU research and ensuring that is used most effectively (Figure 2).

CONCLUSIONS

With increasingly complex data available for managing both print and online serials and declining budgets, there is a strong need for a method of organizing serials data beyond what was available from an integrated library system. In the future, all ILS will need to generate reports that integrate information about journals from all records and allow data generated from other sources to be uploaded. There is also a need for an ILS or other commercial product to compile data from various sources easily, so less time can be spent coordinating the data and more time can be devoted to analyzing it.

A method of creating priorities has been suggested here, but the adaptability and different methods of analyzing data available to users of the SDD has not been explored in-depth. However, the ability to analyze library collections using data driven measures enabled by the SDD further enables librarians to demonstrate the value of their collections at a time of tremendous change in libraries and increasing budget pressures felt by academic institutions. While these quantitative measures are not the exclusive measure of value to academic libraries, for libraries to effectively serve the educational and research needs of their institutions this data driven approach should form the foundation of collection management. The SDD is one solution that can assist in shaping collection formats and determining ownerships, as well as facilitating consensus building by a providing communication tool based on objective information to make difficult subjective decisions.

ACKNOWLEDGEMENTS

NOTES

1. Paul Metz and John Cosgriff, "Building a Comprehensive Serials Decision Database at Virginia Tech," *College and Research Libraries* 61, no. 4 (2000): 324-334.

2. Andree Rathemacher, "E-journal Usage Statistics in Collection Management Decisions: A Literature Review," Chap. 6 in *Library Data*, ed. Darby Orcutt, (Santa Barbara: ABC-CLIO, LLC, 2010), 72-89.

3. Metz and Cosgriff, "Building a Comprehensive Serials Decision," 324-334.

4. John Gallagher and others, "Evidence-Based Librarianship: Utilizing Data from all available Sources to make Judicious Print Cancellations Decisions," *Library Collections, Acquisitions & Technical Services* 29 (June 2005): 169-179.

5. Robin A Paynter, "Commercial Library Decision Support Systems: An Analysis Based on Collection Manager's Needs." *Collection Management* 34 (January 2009): 31-47.

6. Paynter, "Commercial Library Decision Support Systems". *Collection Management*, 31-47.

7. "Introduction to Release 3 of the COUNTER Code of Practice for Journals and Databases (Published August 2008). http://www.projectcounter.org/code_practice.html (Accessed August 22, 2010).

8. "ERM as a Service. Powered by SwetsWise."

  http://www.swets.com/web/show/id=3124136/langid=42 (Accessed August 22, 2010)

9. EBSCONET® ERM ESSENTIALS™ Fact Sheet. http://www2.ebsco.com/en-

  us/Documents/prodServices/EBSCONET_ERM_Essentials_Fact_Sheet.pdf (Accessed

  August 22, 2010)

10. Metz and Cosgriff, "Building a Comprehensive Serials Decision," 324-334.

11. Diane Carroll, "Evaluation and Management of Print and Electronic Journals Using a Serials

  Decision Database" (paper presented at the 23rd Annual Online Northwest Conference,

  Corvallis, Oregon, February 10th, 2006),

  http://www.ous.edu/onlinenw/2006/program.html (Executive Summary access August

  22, 2010).

12. Ibid., 15

13. Betty Galbraith, "Journal Retention Decisions: Incorporating Use-Statistics as a Measure of

  Value." *Collection Management* 27 (May 2003): 79-90.

14. Diane Carroll, "Procedures for creating a serials decision database."

  https://research.wsulibs.wsu.edu:8443/jspui/handle/2376/2277, (accessed August 22,

  2010)

15. Diane Carroll, "Procedures for creating a serials decision database. Use of Excel Merge

  Macro to merge holdings and use stats for ejournals."

  https://research.wsulibs.wsu.edu:8443/jspui/handle/2376/2277, (accessed August 22,

  2010)

16. Cindy Ellis and Diane Carroll, "WSU Wizard using Excel."

  http://research.wsulibs.wsu.edu:8080/dspace/handle/2376/2246, (accessed August 22,

  2010)

17. "Using Access or Excel to manage your data." http://office.microsoft.com/en-us/help/HA010429181033.aspx, (accessed August 22, 2010)

18. Joel Cummings, "Full-Text Aggregation: An Examination of Metadata Accuracy and Implications for Resource Sharing." *Serials Review* 29 (Spring 2003): 11-15.

19. "Assisting faculty, saving money. Database helps Libraries manage journals." *WSU Today Online*. Monday, Aug. 24, 2009**.** http://wsutoday.wsu.edu/pages/Publications.asp?Action=Detail&PublicationID=15305&PageID=21 (accessed August 22, 2010)