

Shared Cataloging Program - Reports to HOTS

November 8, 2005
Accomplishments FY 2004/2005

Electronic Journals

This year, SCP increased access to serials by 3,213 links¹. Looking at specific packages, the most substantive growth resulted from the acquisition of a single new paid package subscription, EBSCO Business Premier (1,607), which was purchased to replace the ABI/INFORM package, and from open access journal additions (1,458), many of which were titles from the Directory of Open Access Journals (DOAJ). Substantive growth was also seen in Blackwell-Synergy (289), JSTOR (119), ScienceDirect (108), and Expanded Academic ASAP (105). New titles were added steadily across-the-board to various other packages. Counting EBSCO Business Premier, 16 new packages were added to the SCP workload. Except for EBSCO Business Premier, SAGE (352), SciELO (246), and SourceOECD (205), the number of titles in each was under 100, and most were under 50.

Integrating Resources

Not surprisingly, the number of integrating resources added was pretty small. We saw only a net increase of 55 titles (including CalDoc IRs). Reiterating our comment from our report last year, the integrating resources themselves represent an insignificant workload relative to journals and monographs. It is the potential for discrete, individual title access within a database that we tend to be concerned about since the SCP workload mushrooms accordingly.

Electronic Monographs

Four new packages were processed with one, Early American Imprints, accounting for about 68% of the total growth in e-monograph access (34,049). Our continued work on titles within the LION database resulted in our next most substantive output, 11,627 titles. Standing packages experiencing notable growth were: SPIE (1050), LNCS (613), IEEE Xplore conference proceedings (581), UC Press e-Scholarship (467), Methods in Enzymology (391, new package), CRCnetBASE (315), Black Drama (291), ACM Digital Library (194), IEEE Xplore standards (185), and Digitale Bibliothek (146, new package).

California Documents

SCP provided access to 151 serials, 488 monographs, and 6 integrating resources during this report period. We continue to note that the publishing pattern of print and online to online only continues in force.

PIDs and BibPURLs

Over 96,000 PIDs and nearly 3,500 BibPURLs were being maintained at the end of the fiscal year. BibPURLs are created using the same technology as PIDs and their use is guided by the same principles. We adopted the use of BibPURLs in keeping with national level cataloging practices. BibPURLs are used for *free* access materials and can be added to the OCLC master record for national, and UC, distribution. In contrast, PIDs are created for UC *licensed* resources, and being UC specific, should not be entered into the OCLC master record.

Productivity/Statistics

Net production statistics since inception through:

	June 30, 2004	June 30, 2005	FY 2004/05 Net Increase
Serials	12,126	15,339	3,213
Monographs	23,961	74,114	50,153
Integrating Resources	314	363	49
CalDoc serials	768	919	151
CalDoc monographs	1,723	2,211	488
Cal Doc IRs	20	26	6
Total	38,912	92,972	54,060

PIDs	40,168	96,371	56,203
BibPURLs	1,849	3,466	1,617
Total	42,017	99,837	57,820

FY 2004/2005 Production Details

	New Access ¹	Modified Records ²	Withdrawn Access ³	Total Transactions
Serials	3,149	4,265	197	7,611
Monographs	52,983	421	70	53,474
Integrating Resources	48	38	7	93
CalDoc Serials	161	269	31	461
CalDoc monographs	448	203	35	686
CalDoc IRs	8	---	---	8
Total Transactions	56,797	5,196	340	62,333

¹ New Access: all instances of adding an 856 link, either because a new title was cataloged, or a new link was added to previously cataloged title.

² Modified Records: all instances of bibliographic record maintenance such as updating holdings data, processing title changes, correcting cataloging errors, etc

³ Withdrawn Access: all instances of the removal of an 856 link

Other Accomplishments FY 2003/2004

Continued growth in the program has necessitated a greater tracking and accounting of the SCP. To facilitate appropriate allocation of resource, workloads, and cataloging priorities, and to aid in the analysis and projection of future program needs, the **Shared Cataloging Program Unit** was established as an independent administrative unit within the Metadata Services Department of the UCSD Libraries. Organizationally, this allowed us to streamline administrative oversight and eliminate the SCP Management Group. The Planning/Operations group continues to meet and has continued to devise and refine local monograph processing techniques for large record sets. The group also developed, and members have begun implementing, an action plan to update all existing serial records to current cataloging conventions.

Staffing on the cataloging side increased by .70 FTE to accommodate an increased workload, specifically, that brought upon by the cataloging of monographic packages.

CDL and SCP staff continued to meet and to plan a transition from the OCLC PURL resolution software to mixed environment using **SFX OpenURLs** and a proposed NOID resolution service. CDL needs to write the programming before the migration can be attempted. A pilot test using the JSTOR titles was being developed to allow and reveal potential problems so as to resolve outstanding technical questions and to create a prototype interface for review by public services and other stakeholders.

Throughout the last several months of the fiscal year, SCP staff worked with UCSD Acquisitions staff and the SCP Advisory Committee on the **793 \$g proposal** . Via this subfield, the campuses would have a mechanism by which to track the nature and currency of access for each vendor for each serial title.

SCP staff and Nancy Kushigian, Director of Shared Print, held a series of meetings to explore and define the relationship between the SCP and the **Shared Print Collection** . Of immediate concern was the degree to which SCP would have aided in the processing of the shared JSTOR print collection. SCP lent advice on issues and processes as well as provided data on potential workflows and staffing requirements. SCP completed a time study to collect data on the record distribution process and workload to help forecast costs should SCP be asked to distribute shared print records. Ultimately, it was decided that the SCP would not involve itself in the cataloging of shared print titles or in the distribution of bibliographic records for those titles to the campuses. However, Shared Print Collection staff would make use of whatever cataloging work SCP had already done, as appropriate.

An ongoing problem for CDL staff had been the tracking of variant names used for various packages, databases, and vendors. At CDL's request, SCP staff created an **authorized headings list** . Largely the work of Becky Culbertson, this list < <http://www.cdlib.org/inside/projects/scp/authorizedheadings.html> > is in an Excel format that allows various types of sorting such as by name, vendor, publisher, tier level, etc. In the future, this authority file can be used to standardize terminology used in an ERM.

A **special file distribution request** for nearly all the SCP records was processed for UC Merced so as to seed their new catalog. This involved coordinating work with CDL Acquisitions to properly identify their subscriptions. Their file also included all open access resources currently available. A similar request from UCLA was also honored as a follow-up to their ILS migration.

Goals for 2005-06

- Process the ECCO, EEBO, and Early American Imprints packages
- Transition from ABI/Inform to EBSCO Business Source Premier
- Work with CDL on URL resolution service
- Continue to improve communication from the SCP to HOTS and campuses
- Complete "SCP Road Show" report and follow up on findings
- Revise the "Proposals for new cataloging projects" document and re-examine the process
- Initiate overview analysis of the SCP's role and benefits within UC

Horizon Issues

UCSD Database Management staff continues to devise various creative strategies for **the manipulation of large records sets**. Our experience with vendor records indicates that the quality and integrity of data remains varied. Much of our ability to manipulate record sets remains dependent on functionality specific to UCSD's Innovative Interfaces system. We relied on that functionality to batch process two large record sets, LION and Early American Imprints, which combined, totaled just over 47,000 records. Other large record sets are in the pipeline.

The **large scale withdrawal of CDL titles** caused by CDL's cancellation of the Computer Database package was successful for two reasons: first, at ~300 titles, it was a moderately sized project, and secondly, only one package of titles was involved. The switch from ABI/Inform to EBSCO Business Source Premier will be more challenging as there will be many more titles involved and because, since subscription access to both packages will overlap through the end of the year, the switch needs to be done in stages. One stage, which is first priority and currently underway, consists of adding the EBSCO titles not found in ABI. The second stage consists of adding overlapping titles. The final stage will be deletion of ABI access. Because of this staged process of first adding and then deleting, several hundreds of these records will be distributed at least twice. For those campuses that do automatic overlay of records, this will likely prove inconsequential, however, those campuses that don't will find this process a significant workload.

Expanding use of the PID server to other campuses for use for local resources is being deferred until the URL resolution service issue has been settled. UCLA has successfully used the PID server for over two years now. It has proved invaluable to them for all the same reasons it has been invaluable for SCP—it is an effective, efficient, and economical means for performing URL maintenance. As such, the option for campuses to use SCP's URL resolution service is being considered for inclusion as part of any service tendered by CDL for migration from the OCLC PID service

The potential for continuing workload increase remains enormous and must be carefully managed. If completed, the addition of the EEBO, Early American Imprints, and the Congressional Serial Set would add nearly 500,000 titles to our catalogs. Cataloging of individual titles in Asian language packages remains suspended. Originally only a single

package, (China Academic Journals (CAJ), ~1,800 titles), CDL has now acquired Sibiu Congkan (~500 monographs). Additionally, we are aware that some bibliographers are interested in SCP cataloging for Siku Quanshu (~3,700 titles), SuperStar (~50,000 titles), National Diet Library (~30,000 titles), and JapanKnowledge (~600 titles). Shirley Leung has been hired as a consultant by CDL to specifically review and recommend a plan of action for the continued acquisition and cataloging of CJK resources system-wide. SCP staff will be keenly interested in her findings.

Request for the cataloging of **new material types** , other than serials and textual monographs have surfaced. These include cartographic materials from the Rumsey Collection (~45,000 titles) and the Library of Congress American Memory Project (7,000+ titles) and musical sound files from Alexander Street Press, Naxos Digital Services, and New World Records (the number of titles is unknown but is presumed to be in the thousands of titles). The request for the cataloging of cartographical materials has been sent back to its originators for refinement, and a final acquisitions decision for the various sound files has yet to be made. Still, this points to greater workloads for the SCP, and for processing formats not currently mastered by SCP staff. Additionally, it points to a significant decision making process for the development of cataloging policies and procedures for these new material types.

Fortunately, hopefully, some of these packages SCP will be able to handle via one time **batch processing projects** , but even for these, some will remain ongoing workloads as additional titles are added to the package. As just these enumerated materials total over half-a-million titles, obviously, the SCP will continue to need appropriate financial support and staff with the appropriate cataloging and language expertise to successfully complete its mission.

As of this report date, a preliminary decision on a **shared database/electronic resources management system** was made favoring the Verde system from Ex Libris. The total impact of this decision on the SCP is uncertain, but already the SCP has been asked to enter into each serial record the unique SFX ID number for future ERM functionality.

As the conclusion to our report last year, we set ourselves the task of performing an **analysis of the role and benefits of the SCP** . To this end, we planned to investigate at least three aspects of our current environment. These were 1) a more detailed look at services offered by outside vendors, 2) a means to solicit greater feedback on needs and improvements to services rendered by the SCP to the campuses, and 3) to explore and identify additional value-added services the SCP could provide the campuses. While we did not make any headway on item one, for item three, we took a request for developing a means to track subscription status and put forth the 793 \$g proposal. On another front, we will soon be working with CDL staff on adding data to serial records to aid with future ERM functionality.

With regards to item two, we decided to set up a series of presentation by SCP staff to the individual campuses, the "SCP Road Shows." In these presentations we would explain in greater detail aspects of the SCP, learn first hand of the processes used by the campuses for processing SCP records, address issues raised, and bring back ideas and concerns for further action. On November 10, we will have completed the last of these presentations with our visit to UC Riverside. A fuller report of our experience will follow to HOTS and the SCP Advisory Committee, but some preliminary comments are as follows.

Over the course of the presentations, a common understanding of the purpose, scope, goals, and history of the SCP was created system-wide. An overarching best practice,

that campuses load SCP records as is, emerged from the campuses. For three of the campuses, UC Santa Barbara, UC Berkeley, and UC San Francisco, the afternoon focus sessions proved an extremely fertile environment for all parties involved to brainstorm and develop ideas addressing local processing issues so that each could successfully work towards the overarching best practice. And finally, for SCP, it was extremely educational to hear first hand of the variety of local processing work streams and campus needs.