

Shared Cataloging Program - Reports to HOTS

Report to HOTS for the November 14 meeting by the Shared Cataloging Program (SCP)
November 10, 2003

SCP Web site: This site is scheduled to be released on "Inside CDL" on November 17; see <http://www.cdlib.org/inside/projects/scp/> for a sneak preview.

Accomplishments to date since July 1, 2002

Electronic Journals

This year, SCP added over 2,500 new serial titles records. Nearly half (1,176) of the electronic journals were from CDL's three licensed A&I databases: Expanded Academic ASAP, ABI/Inform, and Computer Database. Journals accessible through PCI full text represented another significant number of electronic journals (264). Otherwise, new titles were added steadily across-the-board to the 29 existing and 9 new journal packages. Though initiated in January 2003, cataloging of journals in the China Academic Journal package (1,700+ titles) was suspended in May due to instability of content and other issues. Though now suspended, UCSD supported the training of two Chinese catalogers in serials and electronic resources cataloging to tackle this assignment. (Should this package be reactivated, UCSD would need additional resources to support one or more Chinese catalogers to take on this workload.)

Integrating Resources

Forty-three new integrating resources were cataloged. Generally speaking, the integrating resources themselves represent an insignificant workload relative to journals and monographs. However, in many cases, the integrating resource is a database-such as PCI full text and the 22 Literature Online (LION) databases-and discrete, individual title access is possible and the SCP workload mushrooms accordingly.

Electronic Monographs

Over 15,000 electronic monograph records have been distributed for materials in ten packages. The largest numbers represent materials from the following packages: Making of America (7,373), IEEE Xplore conference proceedings (4,088), ACM Digital Library (1,233), University of California Press eScholarship editions (1,216), Lecture notes in computer science (900), and CRCnetBASE handbooks (629). To facilitate local processing by recipient libraries, some records were distributed in distinct package-by-package sets. Distribution of these began on March 10 with the Oxford reference online package and additional sets of records following approximately every other week. Other sets of records have been cataloged and are in the queue for distribution.

California Documents

In addition to the above, and during this period (July 2002-Sept. 2003), the SCP has cataloged and distributed 134 California document electronic journals and 734 electronic documents. California documents are not "pre-packaged" as a resource; all cataloging is done at UCSD on a one-by-one basis, working in close collaboration with the UC Government Documents Librarians group for priority-setting.

PIDs

At the end of FY 2002/03, over 32,000 PIDs were being maintained. Converting records that linked to the CDL-hosted ABI Inform and Computer databases so they now link to individual titles in the ProQuest ABI/Inform Global and the Gale Computer databases, and converting the old MAGS links to Expanded Academic ASAP were specific PID maintenance projects. Additionally, a couple of publishers changed the structure of their URLs requiring immediate action, the most notable being a URL change by ProQuest. SCP catalogers were able to update over 800 ProQuest URLs within a 48 hour period, a response time that earned us great praise from the vendor. In early November 2003, 450+ Wiley Interscience online journal URLs were updated in the same manner.

Production

FY 2002-03

Total number of records cataloged and distributed during FY 2002/2003 (includes regular and CalDocs combined):

Serials	2,840
Monographs	15,787
Integrated Resources	43
Total	18,670

Total production for SCP since inception

Serials	10,762
Monographs	15,614
Integrated Resources	162
CalDoc Serials	530
CalDoc Monographs	1,464
Cal Doc Integrating Resources	20

TOTAL NEW RECORDS DISTRIBUTED: 28,552

TOTAL number of PIDs maintained: 32,473

If there is interest, we can calculate the count of updated/modified records (e.g., when coverage of a journal changes). (URL changes are part of the PID maintenance workload.) Updates to serial records are a standard and accepted part of the serials cataloging workload, and are not counted above.

Other Accomplishments FY 2002-03

As of March 3, 2003, UCSD took over the SCP record distribution process from CDL. As a result, our sister UC campuses now receive their SCP records on a weekly basis, four to six weeks sooner than before.

Electronic monographs and how they should be handled by SCP was debated and resolved system-wide. HOTS and SOPAG endorsed the concept of "separate" records for electronic monographs; further input was sought from HOPS, RSC, and CDC with the University Librarians giving final approval.

While processes developed for the cataloging and cloning of electronic journals were employed in the cataloging of electronic monographs, the sheer volume of titles in these sets challenged the SCP staff to develop **new processing techniques**. Fortunately, the SCP was able to tap the skills of Karen Peters and Ryan Finnerty who developed automated techniques to create "separate" monographic bibliographic records based on the record for the print version. As an example of how this was done, a detailed procedure may be found at <http://www.cdlib.org/inside/projects/scp/SCPCreate.html>.

The SCP Steering Committee was discussed by HOTS and emerged with a new charge and a new name: the SCP Advisory Committee. The relationships and communication channels between various groups were clarified in a new organizational chart.

A "tracking database" at <http://www.cdlib.org/inside/projects/scp/SCPDatabase.html> was created so campuses could know the status of what packages are being worked on. This includes both monographic and journal packages.

Staffing on the cataloging side remains stable. The attached organization chart shows the internal structure for SCP work at UCSD. An SCP Management Group was formed, consisting of the Department Head, Division Head, and SCP Manager, to provide a formal mechanism for management and communication. All questions about the SCP should be directed to this group; we intend soon to set up a mailto link on the SCP Web site. In addition, a SCP Planning/Operations Group has been formed to plan specific workflows for specific packages. The SCP catalogers, who all work in the Serials Cataloging Unit in the Catalog Department, meet regularly as part of a larger SCP Information Group, which include CDL Acquisitions staff, Database Management staff, and others who are interested in the activities of the SCP.

Current Activities

Ongoing activities for serials continue to be cataloging new serial titles, updating journal coverage data, deleting withdrawn serial titles, and PID maintenance. There is, however, a definite shift in emphasis for SCP from serials to monographs. Additionally, SCP is processing and planning for the distribution of distinctive records sets for a few more monographic packages such as the IEEE standards and Black Drama. As appropriate, other monograph record sets acquired during the coming year will be processed in the same manner. Following up with the previous distributed monograph record sets, we will begin sending additions to those packages. Cataloging for materials in these packages is now mainstreamed with our other cataloging. Records for these items will be distributed as normal within the SCP weekly files. The SCP managers have set up a planning and operations group to develop processing procedures for vendor supplied record sets. We anticipate development of these processes to be challenging, because of the variability and inconsistency found with vendor records. As identified, issues that require policy setting will be referred to the SCP Advisory Committee for discussion and appropriate action. SCP managers will continue to develop necessary documentation and will work on improving the SCP web page

as a source of relevant information on the SCP. The Web site itself will be migrating to the Inside CDL server, with public release scheduled for November 17, 2003.

Goals for 2003-04

- Work with CDL to upgrade PID server hardware and software.
- Process the LION, EEBO, and Digital Evans packages.
- Improve communication from the SCP to HOTS and campuses.
- Upgrade and update the SCP Web site.
- Work collaboratively to establish and effectively utilize the SCP Advisory Committee.
- Revise the "Proposals for new cataloging projects" document and re-examine the process.

Trends and Issues for Consideration

The manipulation of large records sets, especially for monographs, is one of the most significant emerging trends. Our current experience with vendor records indicates that the quality and integrity of data is varied. As such, SCP catalogers will continue to rely on Karen and Ryan's expertise in manipulating records at various stages of the cataloging process. Additionally, much of their ability to manipulate records is dependent on functionality specific to UCSD's Innovative Interfaces system. One recent bright spot was Ryan's development of a macro that will create a separate electronic record from the record for the print version in two keystrokes. Concurrently, SCP managers, with support from the SCP Advisory Committee, will need to spend increasing amounts of time reviewing and setting up the parameters for manipulating vendor records for quality and technical aspects.

Current budgetary considerations are forcing us to consider, for the first time, **large scale withdrawal of CDL titles** from the SCP program. The impact of a major package cancellation is uncertain. For example, cancellation of the Elsevier contract by CDL but continuation of the contract by individual campuses will necessitate a discussion on the use of PIDs and withdrawal processing procedures. The reality of dealing with multiple campuses with multiple ILSs causes us to believe that withdrawals may not be as simple and straightforward as we might hope.

Related to the above, is the possibility of **expanding use of the PID server** to other campuses for use for local resources. Currently, UCLA is the only other campus that is using the PID server. This experiment has proved very successful, and use of the PID server by other campuses would seem to be the type of service CDL could provide. The SCP Advisory Committee has had initial discussions on this issue. Before the service can be expanded however, various technical and administrative details need to be worked out. Of these, the most significant are the technical issues as CDL staff is currently working on migrating the data and upgrading the hardware and software.

We at UCSD are very pleased with the success of this program. The number of titles funneled through the program has steadily increased, but **the potential for continuing workload increases is enormous** and must be carefully managed. CDL's Collection Development Committee (CDC) sent out a general call for UC bibliographers to identify freely accessed materials for "collecting." One journal collection falls into this category, the Digital Open Access Journals (DOAJ), but we anticipate that a significant number of single, free titles will be identified by bibliographers. The Making of America and National Academy Press, both free access monographic materials, have also already been identified for SCP cataloging. The China Academic Journals (CAJ) package could also be reactivated. Obviously, these types of materials could become a significant workload for the SCP and appropriate financial support will need to be

found. In addition, any cutbacks to Tier 1 subscriptions would not necessarily mean a corresponding reduction in titles to be cataloged and maintained. As suggested by the Elsevier example, the SCP could still be asked to maintain the Elsevier titles for the individual campuses (in essence, Elsevier turns into a Tier 2 package).

Various policy issues will no doubt emerge. One that is to be presented to HOTS soon is the **classification of monographs**. This has recently been discussed by the SCP Advisory Committee. Currently, all serials and integrating resources are classified specifically. For a variety of reasons, this workload is minimal. However, because of sheer numbers, for electronic monographs the workload potentially is quite large. For example, EEBO (125,000?) and LION (734) monograph records lack classification numbers. Even if we were able to find copy in our own catalogs or in OCLC with classification numbers, these two sets by themselves would require SCP to perform too much work at too high a level of granularity to be sustainable. Another avenue that should be explored is the possibility of broader classification rather than piece-by-piece classification assignment.

SCP managers wonder about **potential new demands made by the UC eLinks program**. It seems that there might be some relationship between these two programs that should be explored to increase overall efficiency and benefiting both staff and users.

CDL and UCSD Acquisitions staff have been exploring the creation and use of a **shared database/electronic resources management system** that could be shared for collection development and processing functions. The impact of such a development on SCP is uncertain, but it is something that we are actively tracking.

Finally, there is concern about the potential impact of the new **UC Shared Print Collection** and what a broadened program might mean for SCP. For the existing pilot projects at UCLA and UCSD, information about Shared Print monographs and integrating resources is not being distributed at all to the campuses for their local catalogs (although it is accessible through Melvyl). Information about the Shared Print journals will be distributed via an Excel file, and not through redistributed SCP records. Given the potential growth for shared print collections, SCP is tracking this project closely. What options are possible for the University of California system to improve and sustain collaborative cataloging?