AC 2009-2180: ASSEMBLING A "BEST COPY" ARCHIVAL JOURNAL COLLECTION: CASE STUDY OF THE UNIVERSITY OF CALIFORNIA IEEE PROJECT

Robert Heyer-Gray, University of California, Davis

Bob Heyer-Gray is a Reference and Collections Librarian at the Physical Sciences & Engineering Library at the University of California, Davis.

Jean McKenzie, University of California-Berkeley

Jean McKenzie is the Head of the Kresge Engineering Library at the University of California, Berkeley.

Lisa Ngo, University of California-Berkeley

Lisa Ngo is the Instruction & Electronic Services Librarian at the Kresge Engineering Library at the University of California, Berkeley.

Karen Andrews, University of California-Davis

Karen Andrews is the Head of the Physical Sciences & Engineering Library at the University of California, Davis.

Emily Stambaugh, California Digital Library

Emily Stambaugh is the Manager of the University of California Libraries' Shared Print Collections at the California Digital Library.

Assembling a "best copy" archival journal collection: Case study of the University of California IEEE Project

Abstract

As the IEEE digitizes its journals, librarians at two University of California (UC) campuses undertook a project to identify a single best possible copy to preserve in a regional shared storage facility. They describe the process: identifying who owned the most complete runs of each title, the standards they developed to validate completeness and acceptable condition, coordination among all the parties involved, communication with additional campuses as issues arose, cataloging challenges, cost determination, staffing used, space saved, and lessons learned. The methods developed for this pilot project will serve as a model for future endeavors involving multiple institutions that want to collaborate efficiently.

Genesis

When IEEE announced in 2003 that it was digitizing its journals back to the first volume, both Camille Wanat, Head of the UC Berkeley Engineering Library, and Karen Andrews, Head of the UC Davis Physical Sciences & Engineering Library, had the same thought. With both libraries having storage quotas to fulfill, there was always competition to be first to send items to the Northern Regional Library Facility (NRLF), where the five northern UC campus libraries store older and low-use materials. Guidelines for storing a title were that no duplicate items were allowed, and there was no requirement for a quality or completeness check. The UC Berkeley and UC Davis engineering librarians thus proposed a joint cooperative effort to preserve the best, most complete set of each IEEE journal title. This would benefit other UC libraries because they could potentially withdraw their IEEE back volumes knowing that a validated “best” set had been stored on behalf of all.

Background and Literature Review

The University of California (UC) is comprised of ten universities located in Berkeley (N), Davis (N), Irvine (S), Los Angeles (S), Merced (N), Riverside (S), San Diego (S), San Francisco (N), Santa Barbara (S) and Santa Cruz (N), nicely divided into five Northern (N) and Southern (S) campuses. Two Regional Storage Facilities (RLFs), north and south, located near Berkeley and in Los Angeles, have provided storage space for older and lesser-used materials for about three decades. For a conceptual description of the roles of the RLFs, see Schottlaender[1]. Recent policy changes have led to them being managed as persistent shared collections. This change guarantees that deposited volumes will remain available for all campuses, no longer under the control of the original depositing library. Details of UC shared print goals and objectives are described by Anderson[2] and give the context for this pilot project.

The importance of sharing responsibilities for collaboratively storing, preserving and managing print collections is attested to by the number of places that have taken steps in that direction. Recent articles describe the experiences of five libraries in Colorado[3], the Five College consortium of Massachusetts[4], and the Australian Academic Libraries[5]. A list[6] from 2006

(updated in 2007) includes other projects in Ohio, North Carolina, Utah and Sweden. A thorough review of all North American library storage facilities and their functions is detailed in the report by Payne[7] and her presentations[8,9] illuminate the very recent trends in shared collections management and stewardship. Payne[8] and Malpas[10] both cite recent unpublished research that Candace Yano, a professor of operations research at UC Berkeley, has completed on behalf of Ithaka [http://www.ithaka.org/]. Her calculations indicate that about fifteen print copies of any given item need to be retained to ensure that one good copy will be preserved in one hundred years.

Initial guiding principles

Acknowledging the prevailing environment of declining budgets and staffing levels, a methodology was proposed that would minimize any additional workload and ensure success. The guidelines included:

1. The UC Berkeley and UC Davis campuses would lead the project and try to gather the bulk of the volumes from within the two fairly comprehensive collections. Volumes from other UC libraries were to be sought to fill gaps or when both copies of a volume proved insufficient. The goal was to compile the best possible set from within the UC system and not attempt to acquire missing items beyond that.
2. Titles would be validated for completeness at the issue level and not include a page-by-page check.
3. A complete list of current and preceding IEEE titles would be compiled and only titles that were completely digitized would be sent. Each campus would take complete responsibility for submitting a title to NRLF. If the responsible campus did not have a complete set for that title, it would request the missing volumes from the other campus, or from elsewhere within the UC system. Then the entire title would be submitted as if it came from one campus. The other campus would give up ownership of the replacement volume in order to make the union catalog record cleaner in appearance.
4. The volumes would be sent to storage as part of the regular, ongoing workflow so that the staffs of the two campus libraries and NRLF would not be overwhelmed. Working the project into existing shipping schedules, budgets, and storage quotas would take longer but offered the best chance for sustainability in the uncertain environment.

Early Stages

The increasing prevalence of electronic journals and books led the California Digital Library (CDL) to recognize the importance of cooperatively stored print collections. The first UC Shared Print Manager, Nancy Kushigian, was appointed in mid-2004 and was apprised of the interest UC Berkeley and UC Davis librarians had in pursuing a prototype project based in the north in early 2005.

At that time, the only cooperatively owned print collections were all based at the Southern Regional Library Facility (SRLF)[11]. These consisted primarily of one print copy of each electronic journal title acquired on behalf of the entire UC system [see: http://www.cdlib.org/inside/resources/sharedprint/agreements_combined.html]. The unique

feature was that these materials were not purchased and deposited by one campus, and thus did not have campus "ownership." Procedures, staffing, and a special cataloging notation, "UC Libraries Collection," were established for these Shared Print materials acquired on behalf of all campuses. Similar activities did not exist for the northern UC campuses at the NRLF. Before UC Davis and Berkeley could proceed, programming changes were needed in the local Berkeley technical processing system to allow the "UC Libraries Collection" note to appear for NRLF holdings in the UC union catalog, and this delayed the IEEE Project. Bob Heyer-Gray at UC Davis and Jean McKenzie at UC Berkeley were chosen to lead project activities at the campus level. In the intervening time, a master list of IEEE titles documenting holdings information was compiled, and decisions about which older titles to include were made.

In late 2007, the catalog programming was completed. Kushigian had retired, and a new UC Shared Print Collections Manager, Emily Stambaugh, was hired. Lisa Ngo was added to the UC Berkeley Engineering Library team, and the first IEEE title shipments began early in 2008.

Selection of titles to send

The core list of IEEE titles was generated by Brian Quigley, former Electrical and Computer Engineering Librarian and current Head of the Mathematics and Statistics Library at UC Berkeley, by scanning and compiling the periodical titles from the *Index to IEEE Publications* and the *Index to IEEE Periodicals*. Working from this master list of about 225 IEEE and IRE (Institute of Radio Engineers, a predecessor to IEEE) periodicals, librarians chose titles that met three criteria to be included in the Shared Print Collection (SPC).

- Titles had to be completely digitized by IEEE and this was confirmed by checking IEEE *Xplore* Digital Library or the IEEE Periodical Legacy Inventory Report [see: http://www.ieee.org/web/publications/subscriptions/info/historic-content.html]. As of this writing, IEEE has yet to complete its digitization of legacy content despite its earlier estimates projecting the digitization to have been completed already.
- The holdings of UC Berkeley and UC Davis had to be on the same record in the union catalog (Melvyl), as this provided the basis for a separate SPC bibliographic record to be created. If the holdings were on separate records, librarians reported this to their e-serials cataloger to be corrected so the records would merge, and the title was sent later.
- The third criterion sets out the action required if some or all of the volumes had previously been sent to the storage facility by another UC library. In those cases, the librarian contacted the relevant person at the other library asking for approval for their volumes to become part of the SPC program. If agreed, NRLF staff retrieved the materials from the storage stacks, validated them at the issue level, and processed them for the SPC program. Once validated, the materials were returned to their existing location in the NRLF stacks. If approval was not given, the UC Davis or UC Berkeley set was used, and the stored volumes were returned to the owning library for disposition.

Validation of Titles at the Issue Level

UC's Shared Printer Manager led us through a process to define issue-level validation standards. The page-level validation standards developed for the JSTOR [see:

http://www.jstor.org/page/info/about/archives/index.jsp] project provided an excellent starting place, but were too detailed for our purpose. During several conference calls and one all-day meeting* where we could observe volumes of various ages and states, the criteria were established for issue-level validation, including identifying conditions that would cause volumes to be rejected outright, conditions that would require action before items could be accepted, and conditions that would be acceptable but noted on the bibliographic record.

Technical and scholarly content were the primary concern. Issues were rejected if content was missing or illegible through damage or unfortunate binding practices. Photocopies and photostats of issues were also rejected. Volumes with these conditions were requested from the other participating campus. Issues with missing covers, images or advertising pages were acceptable, with a note added to the bibliographic SPC record indicating such. Because of binding practices over the decades, and the fact that UC Berkeley and UC Davis used the same bindery, if covers or advertisements were missing from one collection, frequently they also were missing from the other.

Complete . . . Or Done?

For UC colleagues to use information about this Shared Print Collection in their collection management and storage decisions, they need assurance that the entire run is complete. Gaps left after both library collections had contributed all acceptable volumes are tracked and reported by Jane Kelley, who processes the materials at NRLF. Procedures are in development for acquiring missing materials from sister UC libraries, and for designating an appropriate communication channel for the completed list.

Generally, staff at the library originally sending a title will be responsible for asking other UC campuses to provide missing volumes. The library providing the copy will de-accession (withdraw) its material, send it to UC Berkeley or UC Davis to take ownership of it (accession) and submit it to NRLF to be added to the SPC. Though this may seem convoluted, it is much easier than having other contributing campuses have to learn the validation process and technical procedures required to make special notes in their local catalog and in the union catalog records.

All UC campuses that appear to own missing copy will be asked in turn and, if no copy is available or acceptable, the title will be declared "Done" even if it is not complete.

Since 2004, only UC Berkeley among all the northern UC campuses has continued to subscribe to the IEEE journals in print and they will contribute all titles from 2004 to current to the SPC. Whether a print copy will continue to be purchased in the future for the SPC, who will contribute, and how the materials will be paid for will be the subject of a broader conversation in the future.

---

* The all day meeting held at the UC Berkeley Engineering Library and at NRLF, attended by Emily Stambaugh (CDL), John Kiplinger (Director of Production, JSTOR), Ginny Moon (Deposit Services Manager, NRLF), Scott Miller (Operations Manager, NRLF), Jane Kelley (Deposit Services, NRLF), Jean McKenzie (UCB), Lisa Ngo (UCB), Barcley Ogden (Head, Preservation Department, UCB), and Bob Heyer-Gray (UCD), was essential to the creation of the procedures and documentation used in the project.

Sending Materials

Once titles are selected and issue-level validation performed, the volumes are prepped to be sent to the storage facility. Prepping involves several steps.

- Updating a "Sent or Pending" list on the project wiki (Figure 1) with the titles to be sent so the other campus doesn't start work on the same title

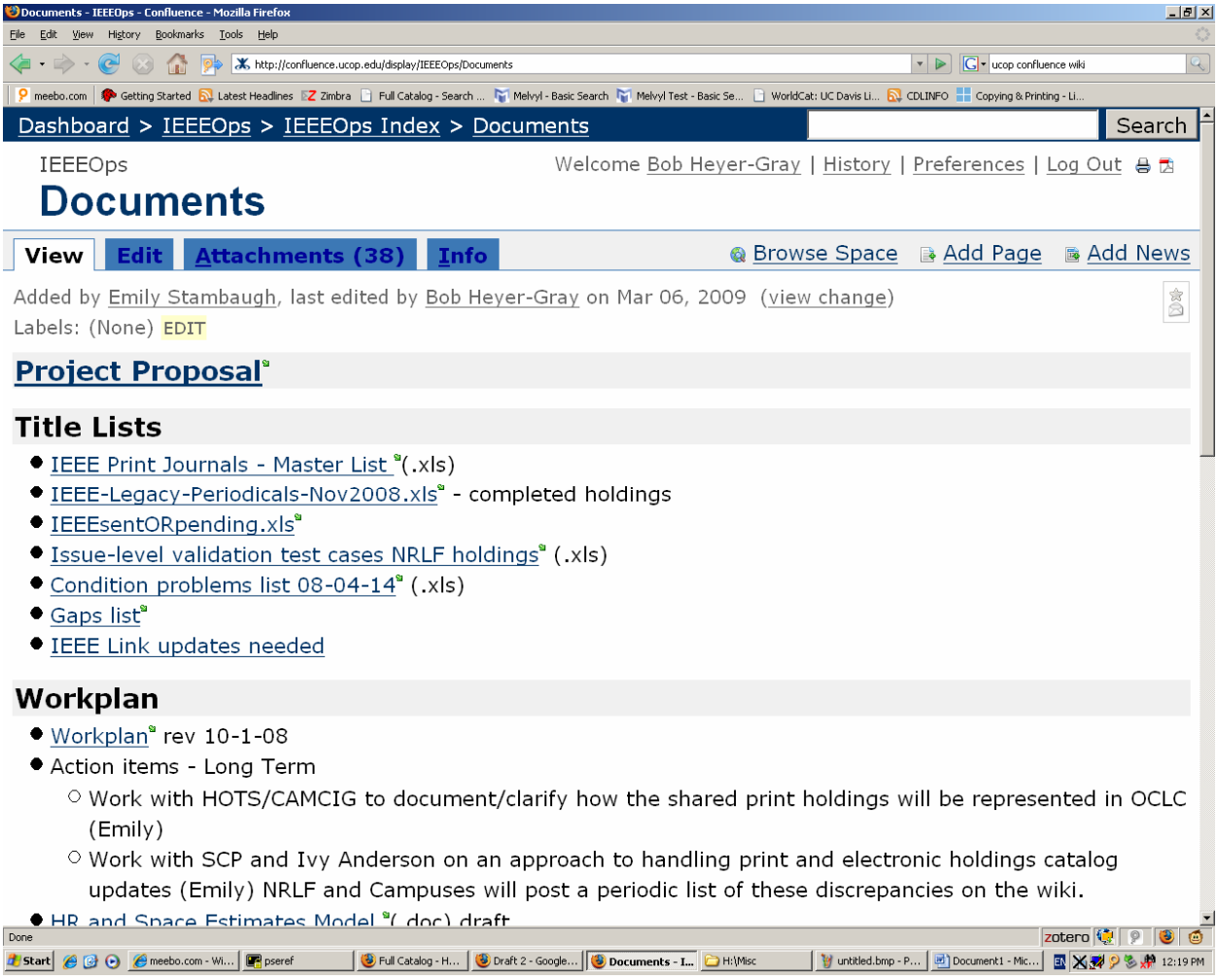


Figure 1, Documents page with links to shared files, including the "Sent or Pending" file from the IEEEOps wiki administered by Shared Print Manager Emily Stambaugh.

- Flagging with a "Public Holdings Note" (Figure 2) volumes that met the criteria but were less than perfect (missing covers, images, or advertisements) to indicate to the storage facility staff that the appropriate note must be added to the bibliographic SPC record.

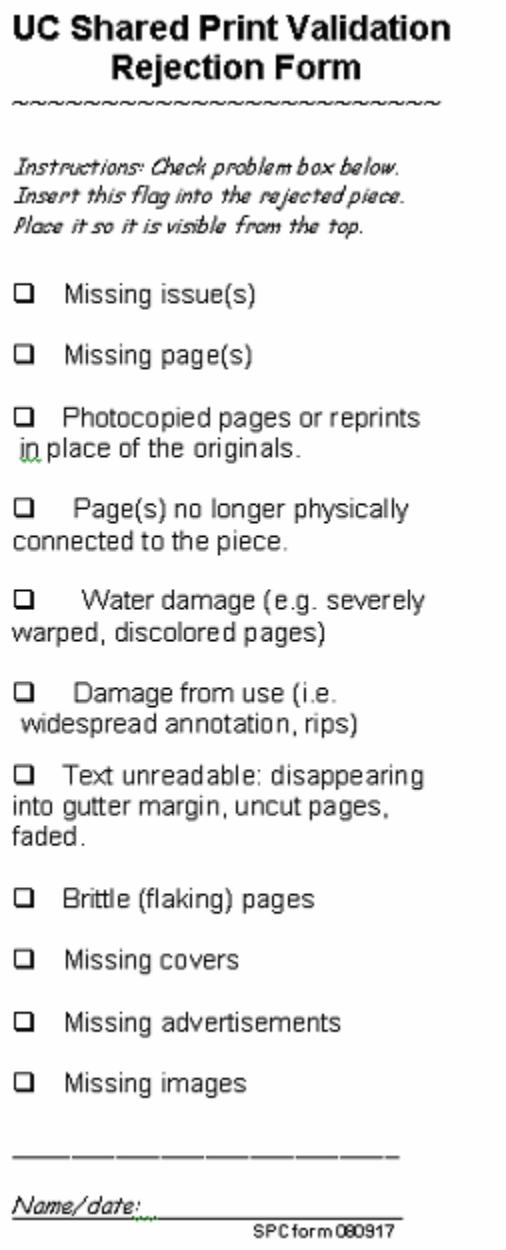
Figure 2, Flags notifying NRLF of acceptable but less than perfect volumes.

- Printing a copy of the union catalog record showing that holdings from both campuses are on the same record to include with the shipment
- Flagging with the "Start" and "Stop" flags (Figure 3) for each title. As the campuses send IEEE titles along with their regular storage shipments, these flags let the storage facility processing team know that the title is a part of the IEEE Shared Print project and needs to be sent to a different workflow.

Figure 3, Stop/Start flag identifying material as SPC.

- Recording statistics using a simple spreadsheet to measure the amount of time spent on the project by managers, librarians, library assistants and students for selecting, retrieving, and processing materials, number of volumes and items shipped, and total shelf space saved.

At NRLF, a new UC-Libraries bibliographic record with holdings notes is created for each title. If NRLF staff find volumes that do not meet the validation criteria, they return them to the sending campus with a Rejection Flag (Figure 4) indicating the reason for rejection. The UC-Libraries record is uploaded into the union catalog and a Gaps List indicating missing issues compiled for further action.



UC Shared Print Validation Rejection Form

Instructions: Check problem box below. Insert this flag into the rejected piece. Place it so it is visible from the top.

- ☐ Missing issue(s)
- ☐ Missing page(s)
- ☐ Photocopied pages or reprints in place of the originals.
- ☐ Page(s) no longer physically connected to the piece.
- ☐ Water damage (e.g. severely warped, discolored pages)
- ☐ Damage from use (i.e. widespread annotation, rips)
- ☐ Text unreadable: disappearing into gutter margin, uncut pages, faded.
- ☐ Brittle (flaking) pages
- ☐ Missing covers
- ☐ Missing advertisements
- ☐ Missing images

Name/date: _____

SPC form 080917

Figure 4, Validation/Rejection Form.

IEEE titles will continue to be sent to storage in this manner along with regular storage shipments for both participating campuses until everything in the master spreadsheet that has been digitized by IEEE has been stored. Through February 2009, a little over 100 titles have been sent with the entire project expected to last five years.

Behavior of the Shared Print Collection

The Shared Print Collection is a dim archive in that the materials are expected to be used only for extraordinary reasons. If patrons request one of the printed volumes (usually done online via the

union catalog), NRLF staff will refer the patron to the digital copy and provide the URL for the title. If the patron then provides a reference or has knowledge of something unique in the print version, NRLF staff will provide a copy or scan of the relevant pages. If the patron makes a strong case for seeing the printed volume, it will be sent to a UC Library for "Building Use Only", or the patron can visit NRLF and use the volume there. The reason given must be more than the user having a preference for print over electronic.

Communications

One of the more important aspects for the ongoing success of the IEEE shared print project is communication. As there are generally two or three separate units working on a title at any given time, it is vitally important that all three groups are kept in-sync and up to date with the details and status of titles being worked on. Modes and tools of communication include in-person meetings, telephone calls, conference calls, e-mails, a listerv and a wiki.

When the procedures and policies were in development, the conference calls and e-mails were more frequent. The conference calls, besides offering an opportunity to update the group on progress made, offered an excellent opportunity to surface additional issues that had not been considered in the initial phases of the project. Coordination between UC Berkeley, UC Davis, and NRLF as the procedures and policies were still in their formative stages clearly necessitated more frequent communication in order to avoid problems. Now that the campuses are in production mode, conference calls are less frequent. However, the use of the wiki and local communications (primarily e-mail between the campuses, but via the listserv as well) are playing a greater role. For instance, early in the process titles were designated for either UC Berkeley or UC Davis depending on such factors as pre-existing holdings at NRLF or more complete holdings for titles. As holdings for most titles actually tend to be identical between the two campuses, communicating in advance with the other campus about which titles are being prepared for storage has become key to the process. The wiki serves as the storehouse of the project proposal, the work plan, group agendas, meeting notes, and working documents and is currently restricted to the IEEE Operations Shared Print working group. Some of the working documents include IEEE titles sent or currently being worked on, shared print titles with gaps, and processing procedures/forms to name but a few. The more up to date the lists are kept, the less likely work is going to be duplicated at the campuses, and it cannot be stressed enough that communication and coordination among the participants are key.

Processing Procedures and Tools

While local processing techniques (e.g. how summary holdings notes are used for materials sent as part of the IEEE shared print project) differed between the campuses, the need to ensure that we all were working from the same principles and standards led to the creation of the working document entitled *Processing IEEE Shared Print Materials by Content Partners and NRLF*. This document combines the issue level validation standards the group adheres to with the practical steps for processing materials into a single usable document. Using consistent criteria for validation, rejection, sending materials, etc., ensures a minimum of misunderstanding regarding what can be sent, how it should be labeled, and with whom one must coordinate.

It is important to point out that both UC Berkeley and UC Davis use the GLADIS (UC Berkeley's online integrated library system) catalog for their NRLF holdings. A new location code needed to be created to accommodate the Shared Print Collection and a new export program for moving GLADIS records to the Melvyl union catalog was also needed. The creation of the export program allowed both UC Berkeley and UC Davis storage records/holdings to be transferred to Melvyl. While local notes and statements can vary among the campuses, the records for the titles in the union catalog need to be consistent and display the shared print location. Cooperation and input among NRLF staff, the UC Berkeley systems department, and the California Digital Library was vital for allowing this to happen with a minimum of delay. Below are two examples of records from the union catalog that reflect combined UC Berkeley and UC Davis contributions.

IEEE Micro holdings and MARC 852 field display in the Melvyl union catalog:

NRLF UC Shared Print-1 Circ status University of California Libraries - Building Use Only
v.1(Feb 1981)-24(2004), 26(2006)-27(2007)

852 |a GLAD |b UL01 |j XXX Shared Print-1 |3 v.1(Feb 1981)-24(2004), 26(2006)-27(2007)

IEEE transactions on antennas and propagation holdings and MARC 852 field display in the Melvyl union catalog:

NRLF UC Shared Print-1 Circ status University of California Libraries - Building Use Only
AP 11(1963)-AP 13(1965), AP 15(1967)-AP 17(1969), AP 20(1972)-AP 21(1973), AP 23(1975)-AP 35(1987); v.36(1988)-44:6(June 1996), v.45(1997)-54(2006)CUMULATIVE INDEXES: PGAP 1(1952)-v.39(1990)
Imperfect copies: issues in vols. 53:7-12, 54:7-12 have tight binding; issues in AP-30 are missing back covers

852 |a GLAD |b UL01 |j XXX Shared Print-1 |z ***Imperfect copies: issues in vols. 53:7-12, 54:7-12 have tight binding; issues in AP-30 are missing back covers*** |3 AP 11(1963)-AP 13(1965), AP 15(1967)-AP 17(1969), AP 20(1972)-AP 21(1973), AP 23(1975)-AP 35(1987); v.36(1988)-44:6(June 1996), v.45(1997)-54(2006)CUMULATIVE INDEXES: PGAP 1(1952)-v.39(1990)

The second record reflects not just the Shared Print location, but also an example of the consistent language in the notes for acceptable, but less than perfect copies.

More recently, UC Berkeley's move to the Millennium (III) catalog and the University of California pilot project with OCLC (Next Generation Melvyl Pilot) have added additional technical wrinkles to the project, some of which remain to be worked out as the project moves forward.

Staffing

As mentioned earlier, the project as proposed was to use only existing staffing and workflows at the campuses and NRLF. Up to this point, this has remained true. Based on early estimates for the initial phase of the project (February – October 2008), staffing at the campuses has come to a

total of about 1.27 FTE, with the bulk of the FTE being library assistant and student time for retrieving, validation, and processing of materials. FTE estimates from the storage facility are .12 FTE, again with the bulk of the work performed by library assistant and student staff.

Lessons Learned

Simple projects frequently are not simple. Despite our best efforts, problems arose that slowed the progress of what was originally thought to be a relatively straightforward storage project. Some problems were out of our control, such as IEEE taking longer than anticipated to scan their legacy content, therefore delaying the selection and storage of materials. Others were the result of mistaken assumptions on our part. Two hurdles of this kind that stood out were our overestimation of the condition of our collections; the other was the discovery that both content partners' holdings as reflected in the catalogs turned out to be less accurate than portrayed. We found that through good communication between content partners, the Shared Print Manager, and the storage facility staff, we were able to overcome problems that arose.

There were also some positive lessons that came from the project. We found that other UC campuses outside of the two original content partners were more than willing to participate and share their existing storage holdings for the project. IEEE, when made aware of a gap in the digitization of a title and knowing that UC Berkeley and UC Davis are in the process of creating a print archive, filled that gap from one of our collections, making IEEE*xplore* more complete for everyone. We also found that adding the Shared Print Manager to the team led to more structure in the process, as title selection and storage would likely have been a far more informal process if solely left up to the content partners.

What started out as a local project may have importance for similar collaborative efforts nationally to create retrospective shared print repositories. The final documentation and materials collaboratively developed for this project for communication and reporting will be made available for use by other institutions and/or consortia.

Bibliography

1. Schottlaender, B. E. C. 2004, '"You say you want an evolution...": The emerging UC libraries shared collection,' *Library Collections, Acquisitions, and Technical Services,* vol. 28, pp. 13-24.

2. Anderson, I. 2007, *Shared Print Goals and Objectives 2007-2009*, viewed March 6, 2009, <http://www.cdlib.org/inside/resources/sharedprint/documents/SHARED_PRINT_Goals_and_Objectives_2007-2009_final.pdf>.

3. Agee, J and S. Naper 2007, 'Off-site storage: an analysis,' *Collection Building,* vol. 26, pp. 20-25.

4. Bridegam, W. E. 2004, 'Print preservation at the local level--The Five College experience,' *Library Collections, Acquisitions, and Technical Services,* vol. 28, pp. 29-38.

5. Genoni, P.2008, 'Current and Future Print Storage for Australian Academic Libraries: Results of a Survey,' *Library Collections, Acquisitions, and Technical Services,* vol. 32, pp. 31-41.

6. Kemp, R. 2007, *List of cooperative print housing endeavors, with brief summaries*, viewed March 6, 2009, <http://library.uncw.edu/web/faculty/kempr/kemp-jstor-cooperative-print-endeavors-11-16-06.doc>.

7. Payne, L. 2007, *Library storage facilities and the future of print collections in North America*, OCLC, Dublin, OH, viewed March 6, 2009, <http://www.oclc.org/programs/publications/reports/2007-01.pdf>.

8. Payne, L. 2008, *The future of library collections: Access and stewardship in a networked world*, Presented at VALA-CAVAL 30th Anniversary Series, viewed March 6, 2009, <http://www.caval.edu.au/assets/files/members/paynecaval2008.ppt>.

9. Payne, L. 2009, *Trends in shared library storage and shared collection management*, Presented at Art Libraries Society of North America (ARLIS/NA), viewed March 6, 2009, <http://www.arlisna.org/news/conferences/2008/proceedings/ses_09-payne.pdf>.

10. Malpas, C. 2008, *Safety in numbers? Calculating optimal overlap for print preservation*, viewed March 6, 2009, <http://hangingtogether.org/?p=448>.

11. Anonymous, 2008, *University of California Libraries Shared Print Program JSTOR project*, viewed March 6, 2009, < http://www.srlf.ucla.edu/Jstor/Default.aspx>.