RESEARCH
SAVE
COLLECT
SHARE
PUBLISH
CREATE KNOWLEDGE

2012–2013 ANNUAL REPORT

University of California
CDL
California Digital Library

SUPPORTING THE RESEARCH LIFE CYCLE COMES FULL CIRCLE

While preparing a presentation during the past year to summarize CDL's work, I suddenly realized that we had achieved a significant overall goal. We have over 20 services addressing all stages of the research life cycle. The "upstream" stages that provide new opportunities for libraries to engage with the research process are well represented with established services such as eScholarship, but also with new tools such as DataUp and the DMPTool. And the more traditional library services all have characteristics of achieving greater impact by being part of larger initiatives such as HathiTrust and OAC, or by connecting researchers to resources beyond UC. Certainly, there are other opportunities for services within this life cycle, and our campus libraries provide many of the critical services that rely on their direct relationships with students and faculty. Nonetheless, CDL's portfolio represents a solid backbone on which the campuses can base their offerings, and they can make use of CDL's services in a variety of ways with local branding, embedding in their web site or other services, or integrating CDL services with workflows and local systems.

The 2012-2013 year saw the maturation of many of these services, whether by enriching features for local use, such as the addition of streaming media support in OAC/Calisphere and print on demand services for eScholarship, or by expanding and connecting to larger communities for increased impact, such as the Western Regional Storage Trust (WEST).

A major driving force during the past year that will continue at an even greater pace next year is the evolution of open access to research publications and data. From the passage of the UC systemwide open access policy, to interest in the state legislature in requiring open access to state funded research, to the Office of Science and Technology Policy mandate for federally funded research, open access is taking hold as a value for researchers, a return on taxpayers' and funders' investments, and a disruptive force in business models for publishers and libraries. CDL will continue to invest, investigate and invigorate its services in support of this significant worldwide movement.

It was especially gratifying that so many of our services received awards during the past year. These awards range from recognizing innovation and doing the right thing at the right moment in time, to recognition of community uptake and usage.
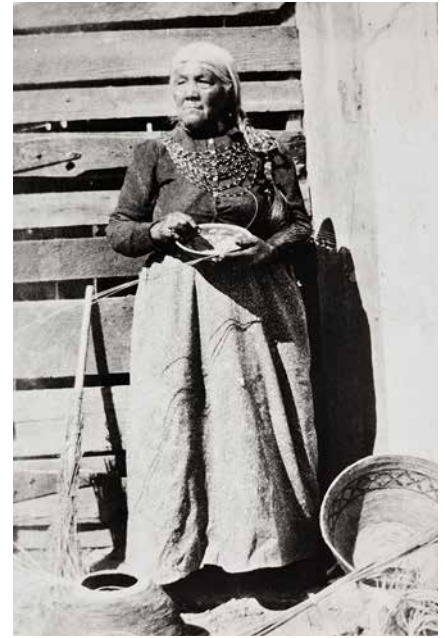
As we enter 2013-2014, we look forward to refocusing our lens on the following high level strategic themes: **Explore, Advocate, Build, and Connect**. Our goals are interconnected with those of the Council of University Librarians, to enrich the systemwide library collection; capitalize on technological opportunities to accelerate the transition to a primarily digital environment; maximize discovery of and access to information resources; optimize and repurpose physical library space; expand engagement in scholarly communication; build and leverage expertise.

Our work is also influenced by changes in scholarly practices that continue to shape what researchers, and by association students, need to accomplish their work. Consequently, we will still use the research life cycle as an organizing principle, but we will shift focus to researchers themselves, putting "scholars at the center" to better understand their individual environments and intersections with library services. Much of our work will focus on a research agenda whereby we seek to better understand the needs of researchers and our campus library partners.

Exciting and perhaps even revolutionary times are ahead as technology propels research practices forward and the desire for openness and sharing takes an even stronger hold. As we compare our work to "top trends" in the library and higher education realm, CDL is well positioned to benefit from and contribute to this environment.

Laine Farley
Executive Director

Petra Pico, famous basket maker and a leader of Ventura's Chumash community: circa 1890. Black Gold Cooperative Library System
http://content.cdlib.org/ark:/13030/c8r49pjk/?query=&brand=calisphere

INTRODUCTION

The California Digital Library's overarching mission is to support the University of California community's pursuit of scholarship and to extend the University's public service mission. CDL's 2012-2013 themes reflect current trends in higher education and librarianship—data curation; digital preservation; changing user behavior and expectations; growth of digital humanities; libraries as publishers; new publication formats, curation and dissemination; open access; collaboration in collection building and resource sharing; and collaboration in large regional print repositories.*

*ACRL Research Planning and Review Committee, ACRL Environmental Scan 2013, April 2013, http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/EnvironmentalScan13.pdf

CDL has aligned its initiatives around a research lifecycle in which new knowledge is created by researching, collecting, publishing, sharing and saving scholarship.

LIFECYCLE STAGE		SERVICE	FUNCTION
RESEARCH – Locate	Books and other formats	Melvyl® UC Berkeley Library University of California	Search for materials owned by UC libraries and libraries worldwide
	Licensed collections	CDL COLLECTIONS, WORKING WITH UC LIBRARIES, NEGOTIATING, ANALYSIS	Acquire 50k journals, 500k books, over 500 databases
	Primarily books	HATHI TRUST	Search 10M books, digitized from over 80 libraries and view over 3M in the public domain
RESEARCH – Use	Primarily articles	UC-eLinks	Connect from citations to publications
	Any format	Request Interlibrary Loan and Document Delivery	Request materials not owned by UC libraries
RESEARCH – Special	Primary sources – images	calisphere	View primary sources contributed by 120 CA libraries, archives, and museums
	Special collections	OAC Online Archive of California	Search detailed descriptions of primary resource collections maintained by more than 200 CA libraries, archives, historical societies, and museums

AND EMERGING SERVICES FARTHER UP THE LIFE CYCLE

LIFECYCLE STAGE	SERVICE		FUNCTIONS
PLAN	Grant application	**DMP**Tool	Create, edit, share, and save data management plans to meet funder requirements
COLLECT	Data collection	DATA**UP**	Open-source add-in and web application for Microsoft Excel as a data collection tool
		Web Archiving Service	Collect and manage ephemeral web-published content
MANAGE, SHARE, SAVE	Tracking, management, citation	EZID	Create and manage persistent identifiers
	Storage, management, sharing	Merritt	Curation repository: store, manage, and share research data
PUBLISH	Scholarly publication	eScholarship University of California	Open access scholarly publishing services: 50,000 papers; 50 journals; books, seminars, and more
	Data publication	**DATA PUBLICATION PLATFORM**	An infrastructure to publish and receive credit for sharing research data

THIS YEAR CDL HAD THREE MAJOR GOALS:

- To advance the transformation in scholarly communication related to openness, sharing and validation.
- To enable a holistic approach to managing the information lifecycle at the appropriate scale, to elevate certain components for more impact as well as to support local needs.
- To adopt a strategic financial plan focused on efficiency and value.

A sampling of our lifecycle activities this year demonstrates how we reflect these themes.

ADVANCING THE TRANSFORMATION IN SCHOLARLY COMMUNICATION: OPENNESS, SHARING AND VALIDATION

Supporting campus open access pilots and publishing funds: Open access (OA) literature is free, online, and available to anyone. An open access article has limited copyright and licensing terms which means that anyone with access to the Internet may read, download, copy, and distribute that article. The business models for some open access journals shift the cost of publication from subscriptions to authors by charging the author an article processing fee. In many fields, particularly the sciences, that fee is then charged to a research grant. Recognizing that not all scholars have grant monies to draw on, CDL collaborated with the UC campus libraries to develop campus-based open access publishing fund pilots to support UC faculty members who wish to make their research findings immediately and freely available to the public. Funding from the California Digital Library was matched by campus libraries or other campus sources to help pay the article processing charges for UC authors who want to make their articles available via open access publishing.

Implementing UC open access policy as adopted by the campuses: In May 2012, The UCSF Academic Senate voted to make electronic versions of current and future scientific articles freely available to the public, helping to reverse decades of practice on the part of medical and scientific journal publishers to restrict access to research results. In support of this policy, and as a first step toward implementation, UCSF and CDL developed a mechanism for faculty to request waivers of the license and place a temporary embargo on the availability of published articles deposited in eScholarship. The UCSF library and CDL notified over 300 publishers of the policy, responded to author questions, and prepared for the next phase in which authors start to deposit their publications. CDL hired Katie Fortney as CDL's Copyright Policy & Education Officer to support open access and other initiatives. Katie brings to CDL degrees in librarianship and the law, and experience in both fields.

MEET Justin Gonder

Access & Publishing Group

Justin Gonder, eScholarship Operations Coordinator, wondered if he'd ever use his master's degree in research psychology in his work at CDL. But his skills in gathering requirements, designing interfaces, and utilizing feedback to create intuitive user interfaces have proved essential to his interactions with faculty and librarians in creating a successful implementation of UC's Open Access policy.

After UCSF faculty passed its Open Access policy in 2012 and it became a real possibility that a wider University of California OA policy might pass, CDL staff leapt into action to create an article deposit process (and an 'opt out' option), if and when the time came. (A UC systemwide policy was passed on July 24, 2013.) Justin joined forces with UCSF faculty and librarians to create mockups in preparation for those depositing articles, building on his in-depth knowledge interacting with eScholarship users through the eScholarship Repository helpdesk. His experience in understanding users' pain points and new feature requests helped immensely in this task. Talking to UCSF faculty and understanding different deposit workflows and scenarios, coupled with the terms of the policy, formed the basis for deposit mockups. That work will directly affect what unfolds this upcoming year, engaging with librarians and faculty now developing a working deposit model to be reviewed by the broader UC community for input. CDL has attempted to prepare for the new policy from the moment it became a glimmer on our horizon, and Justin has been central to that effort.

Justin views a primary goal of his efforts to create support for the open access policy that is as effortless as possible for users and that provides guidance and information for faculty to feel well informed and confident in participating in the University's open access structure.

Justin's a natural for this project. He loves taking a challenging and complex problem like the OA policy deposit system and boiling it down to a workable process. He unbundles the pieces, describing how we accurately reflect our relationship with publishers; explaining embargo terms; depositing with PubMed or eScholarship—what can be daunting processes. His dream is to take this really complex problem and create a mechanism so that, "in the end, we have something that's so super simple—users aren't aware of the process that makes it intuitive and efficient. " If anyone can do it, Justin can.

Evaluating options for harvesting publications in support of the UC Open Access policy: In order to make it easier for UC faculty to deposit their articles in an open access repository, CDL staff have scanned the landscape and identified requirements for a harvesting system to gather publications by UC authors, and written a Request for Proposals (RFP) to select a vendor. Having a system in place for harvesting articles by UC authors will be a first step towards making sure an Open Access Policy operates automatically with a minimum of manual input needed for including publications in our repository. Now that the systemwide Open Access policy has been passed by the Academic Senate, this capability will be used in pilots at UCI, UCLA, and UCSF during the 2013-2014 academic year.

Exploring new approaches to data metrics and data citation practices to validate open access content: CDL's eScholarship provides an open access scholarly publishing platform and institutional repository for more than 60,000 papers; 59 journals; books, seminars, and more. Article metrics have been implemented in eScholarship, displaying total number of hits and downloads for each item, plus a historical display month-by-month so that authors and readers can track the interest in individual articles.

Modeling a data publication service with the UC community: In early 2012, UC Riverside's Bourns College of Engineering staff approached CDL with a request to help their affiliated researchers comply with federal requirements to make publicly available the publications and data resulting from federally funded research. To meet this need, CDL technologists developed integration between three CDL services—EZID, eScholarship and Merritt—allowing a researcher to upload a publication to eScholarship and, as part of that submission, upload one or more datasets that are automatically registered in EZID and submitted to Merritt for preservation.

Cultivating DataCite US Alliance for EZID: The EZID service provides Digital Object Identifiers from DataCite, an international organization supporting citation and archiving of data sets. DataCite US is a collaboration between the three U.S. providers, CDL, Purdue and the federal Office of Scientific & Technical Information (OSTI). Its goals are to improve search, retrieval, access, and re-use of scientific datasets and other scholarly digital objects. The U.S. DataCite partners achieved an active working relationship and are now more coordinated in responding to DataCite's initiatives, which can often be Euro-centric; in addition, a DataCite US website was created to clarify goals.

ENABLING A HOLISTIC APPROACH TO MANAGING THE INFORMATION LIFECYCLE: ELEVATE SELECTED COMPONENTS FOR IMPACT & SUPPORT LOCAL NEEDS

Researching a systemwide digital asset management system (DAMS): One of the long held goals of the University of California libraries has been to create a systemwide digital asset management system for the digital content housed in the campus libraries. A DAMS is a centralized repository for digital files that allows the content to be described and managed, and supports connections to archiving and access solutions. CDL staff were involved deeply in a series of assessment studies, reports, and recommendations relative to a systemwide DAMS under the auspices of the Next Generation Technical Services initiative of the UC libraries. The result of the research is being referred to as UCLDC, the UC Libraries Digital Collection which will include a DAMS, a discovery and display layer to unite collections from the systemwide DAMS with those housed in campus asset management systems, and preservation of digital content in Merritt. The project team is based at CDL and also includes a small group from UC Berkeley's Research IT staff. The end result will allow researchers and scholars to search the rich digital content of the UC libraries from a single portal.

Supporting digitization of campus library collections: Two large scale digitization projects of the UC libraries were completed with UCLA and UCSF, with work continuing at UC Davis and the Southern Regional Library Facility (SRLF). UCLA contributed works from the East Asian, Management, Chemistry, Biomedical, and Arts libraries, and children's books, Italian comedies and rare business and economics texts; UCSF contributed medical library collections. UC Davis is contributing selected California government documents and SRLF is contributing pre-1923 works in all languages. To date, the UC libraries have digitized 3.5 million books with Google and some 200,000 with the Internet Archive, all of which reside in HathiTrust along with works from over 80 member libraries. CDL coordinates activities among the libraries, the digitizing partners and HathiTrust.

228,216
IMAGES IN OAC AND CALISPHERE CONTRIBUTED BY MORE THAN

120
CALIFORNIA LIBRARIES, MUSEUMS AND ARCHIVES

MEET Adrian Turner

Access and Publishing Group

Adrian Turner is a natural born describer. While pursuing his MLIS at UCLA's School of Library and Information Science, he had a strong interest in bibliographic description and was employed as cataloger for Bennett Gilbert Rare Books and B & L Rootenberg Rare Books & Manuscripts, early printed book sellers. He subsequently worked as an archivist for the city of Sacramento and UC Irvine Libraries, Special Collections & Archives—taking a deep dive into the world of library descriptive standards. Thus a metadata and description specialist was hatched.

Adrian has been active in the professional archival community for many years. Since joining the CDL in 2002, he has supported OAC and Calisphere operations, and developed tools and services for contributors. He has also supported a range of libraries, museums, and archives throughout California in expanding access to their holdings via these sites.

Adrian is currently part of a CDL project implementation team focusing on implementing a UC Libraries Digital Collection (UCLDC). The UCLDC is a high priority, collaborative initiative of the ten campus libraries that will result in a shared platform for managing and displaying selected unique digital resources. It will reveal the strength and breadth of the UC Libraries' combined digital collections to end-users, while providing a streamlined underlying technical infrastructure for campus libraries to manage and surface their content.

Adrian will be involved in configuring the DAMS and its underlying metadata model (described above). He enthusiastically characterizes the solution to a longstanding need across campuses as "great and awesome—this problem goes back 10 years!"

Adrian has been a key player in envisioning and implementing the DAMS. He contributed to the UC's Next Generation Technical Services (NGTS) initiative, serving on planning groups in the initial NGTS phases that developed recommendations for a systemwide DAMS. He also participated on and co-chaired "lightning teams" in subsequent NGTS phases. One such team gathered requirements for the UCLDC.

Adrian's passion for supporting the description of and access to unique collections—digital and otherwise, for making invisible archival collections materialize—is infectious. After talking to Adrian for a few minutes you just can't imagine why we waited so long to tackle this project!

Exploring taking the Online Archive of California to the network level: Further supporting UC's research mission, CDL continues to provide online infrastructure for California's special collections housed in more than 200 libraries, archives, historical societies, and museums represented in OAC and images from 120 institutions in Calisphere. The CDL conducted a survey to determine what libraries, museums, and archives would need from online hosting services for archival materials; consulted with experts in the field; and gathered cost projections to run such a system. We decided to continue monitoring these services, especially in light of the University of California Libraries' Next Generation Technical Services initiative recommendation for a UC libraries digital asset management system. With the new system, we will be better aligned to support future developments in this area.

Meanwhile, over 2,000 items were added from ten California libraries with funding from the California State Library's Local History Digital Resources Project, ranging from photographs of minority communities along the California coast, to records documenting the history of community colleges in the state, to letters penned by students incarcerated in Japanese American internment camps. In addition, information about 41 important collections held by nine institutions throughout California documenting environmental history within and beyond the state was added to OAC. The availability of these previously "hidden" archival collections represents the successful conclusion of a cross-institutional project entitled *Uncovering California's Environmental Collections: A Collaborative Approach*. The guides (also known as finding aids) to these collections are available in the Online Archive of California.

Two other notable collections were added to OAC and Calisphere this year: the Milton "Hal" Halberstadt (1919-2000) papers at UC Davis Library Department of Special Collections documenting the life and work of a skilled photographer most famous for his depictions of food and another 530 photographs of Southern California in the first part of the 20thCentury taken by Edward W. Cochems, a self-taught photographer based in Santa Ana from 1915 until around 1945. They were contributed by the UC Irvine Libraries Special Collections and Archives.

751,077
VISITS TO OAC IN 2011-12

821,928
VISITS TO OAC IN 2012-13

986,559
VISITS TO CALISPHERE IN 2011-12

1,015,057
VISITS TO CALISPHERE IN 2012-2013

Conducting streaming media assessment and implementing for Calisphere and eScholarship: CDL Access & Publishing staff, led by Brian Tingle, successfully added streaming media to OAC and Calisphere. Materials were digitized as part of the California AudioVisual Preservation Project and samples include a 1963 interview with Malcolm X and a demo of "Lisa," an early personal computer developed by Apple, contributed by Stanford University Manuscripts Division.

Streaming media is part of the eScholarship article submission process whereby videos may be added as part of an item's supporting materials. See an example in *California Italian Studies*.

Implementing eScholarship Plus—print-on-demand ebooks for eScholarship: CDL established a new print-on-demand (POD) and ebook distribution/sales service for books and journals published in eScholarship, the University of California's open access (OA) publishing platform. This service enables publishing units within UC to develop hybrid OA business models for their publications. Books and journals enrolled in this service will be freely accessible via eScholarship while simultaneously offered for sale in POD and ebook formats via self-branded ecommerce storefronts and through retail affiliates such as Amazon.com. In order to support this new publishing model, eScholarship has contracted with Lulu to provide the platform for these services. As an example, publications in the University of California Perspectives in Medical Humanities series are available in print from Lulu, as well as online in eScholarship.

Shaping strategies for HathiTrust collection development, including print monographs archive: CDL's Ivy Anderson, Director, Collection Development & Management, chaired the HathiTrust Collections Committee, which produced a discussion document with recommendations for handling duplicates in HathiTrust.

Expanding ebook collections: On behalf of UC libraries, CDL licensed new collections of ebooks from Harvard University Press, Karger, Royal Society of Chemistry, CRC Press, Springer, American Geophysical Union and Wiley in 2013. These consortial collections represent significant user benefits, campus cost savings and wider reach within the UC system in the transition from locally-purchased print copies (and the limitations of ILL) to systemwide '24/7' online availability.

> "It's been a very easy process for us… once we identified the Web Archiving Service at the California Digital Library as the vehicle that we wanted to use, it was like falling off a log. We acquired it, we had great help from the organization and we began. We've been going very strongly since."
>
> —**Kevin Leonard,** *University Archivist & Asst Director of Special Collections, Northwestern University Library*

5,700,000+
CLICKS TO ELECTRONIC
RESOURCES VIA UC-ELINKS

A major new initiative began in 2013 with a 'demand driven acquisition' (DDA) task force comprised of campus and CDL members. The task force is currently managing three DDA and 'evidence-based' ebook pilots that include consortial purchasing of ebook collections in several engineering fields, Chinese-language monographs and social sciences/arts and humanities titles from international university presses. A fourth pilot is under investigation to develop an automatic purchasing plan through an ebook vendor based on a consortial profile that is designed collaboratively by all campuses.

In addition, the UC libraries issued an E-book Values Statement with the goal of working with publishers, aggregators and others in the academic community to institute best practices and standards for ebooks as business models continue to evolve.

Integrating DataUp development into researcher workflows:
In using DataUp, an open source add-in and web application for Microsoft Excel, researchers have a data collection tool to integrate into their research workflow. The first version of the DataUp tool was launched in October 2012. Since then, CDL staff members have worked to promote its use in the researcher community by attending conferences and meetings, presenting on DataUp. We held a meeting of stakeholders focused on the future of DataUp where we discussed how best to reach out to potential users of the tool. We also submitted a successfully funded grant proposal to the National Science Foundation. In the coming year we will work closely with the DataONE community and Microsoft Research to continue improving the tool's functionality.

DataUp was cited as one of the Library of Congress's Top Ten Digital Preservation Developments of 2012, as well as receiving a National Digital Stewardship Alliance (NDSA) Innovation Award.

Delivering PAPR registry and collection analysis to support WEST archive cycles: The Western Regional Storage Trust (WEST), a partnership to create a distributed retrospective print journal repository in the western United States, completed its first round of print journal archiving. Under a three-year program jointly funded by WEST members and the Andrew W. Mellon Foundation, WEST goals are to preserve the scholarly record; provide opportunities for members to reallocate physical space; and provide access to archived materials. CDL, as the administrative host for the project, analyzes the data files contributed by members and creates an overlap analysis using the Print Archives Preservation Registry (PAPR).

Portrait of a bride and groom, Santa Ana, California. Edward W. Cochems Photographs, Courtesy of University of California, Irvine Special Collections.

http://content.cdlib.org/ark:/13030/c83r0s4b/?query=Cochems&brand=calisphere

Retention commitments of WEST members are recorded in PAPR, which was developed by CDL's Discovery and Delivery team in partnership with the Center for Research Libraries.

Moving back-end support to the network level: Noteworthy this year was that both our Interlibrary Loan (VDX) and citation linking services (SFX/UC-eLinks) were moved from CDL to external hosts at OCLC and Ex Libris respectively. These migrations to hosted services freed up both CDL staff and hardware to enable work on future projects.

MEET Joe Ferrie

Discovery & Delivery Group

Problem: CDL wanted to move the complicated underlying technology of the UC's libraries' interlibrary loan (ILL) service, VDX, from in-house to its vendor, OCLC. Whom could we count on to do the job? CDL's Joe Ferrie is the ideal person to document and package up UC library VDX practices for OCLC. Why is Joe so perfect for this job? Because he understands the research needs of scholars, having been one himself. Joe's penchant for software development began in his own years as a graduate student at UC Berkeley, studying patterns of iambic pentameter in the poetry of Alexander Pushkin. In the process, Joe discovered he needed to understand programming to advance his studies—he was an early computational digital humanities scholar—and this interest led him to a career as a software developer, eventually taking a hiatus from academe to work for PeopleSoft and Oracle. He has come full circle now, back to the University of California and the libraries, where he works on projects that will ease and enhance the work of other researchers. Moving VDX to OCLC has enabled campus ILL staff to work more directly with OCLC; another advantage is speeding the upgrade process for the UC libraries, bringing in bug fixes and new features without delay. It also frees up server resources and staff at CDL to work on other pressing needs.

Joe's joys in working at CDL are many—he loves working with very smart people who have a deep history in information science and making contributions that help research scholars. Plus he finds himself doing a rich variety of software development he wouldn't have been able to accomplish in industry—designing and building software products from the bottom up. His most exciting recent development is as technical lead creating a new code base for Request (another portion of UC's ILL system) so that the system will eventually evolve into a structure that better meets the needs of researchers. Along with CDL colleagues Michael Thwaites and Priyank Doshi, he took the current underlying code—well over a decade old—and moved it into a modern framework. By rewriting the code, he has laid the foundation for the next round of changes that will positively impact our users borrowing materials from throughout the world. Joe's past bodes well for the future work of researchers and scholars.

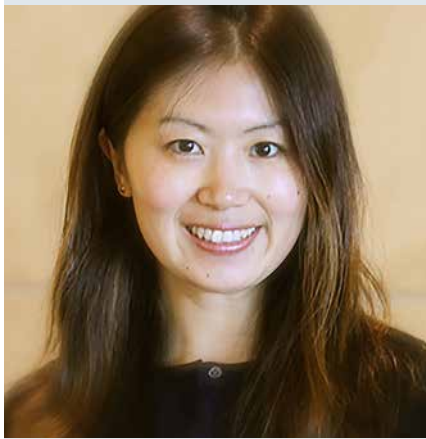## ADOPTING A STRATEGIC FINANCIAL PLAN: EFFICIENCY AND VALUE

Negotiating 6 major renewals for licensed content: On behalf of the 10 UC campuses, CDL embarked on renegotiating licenses for 6 major vendors: American Chemical Society (ACS), CRC netBASE, Thomson/ISI, Taylor & Francis, Springer and Wiley. The majority of these resulted in multi-year contracts, resulting in further financial savings that allowed us to meet our fiscal goals. These agreements included access for our scholars and researchers to large ebook collections as well as to ejournals.

Using metrics for decision-making: Severe budget challenges combined with continuing escalation in the cost of scholarly journals have required UC Libraries to seek economies in systemwide online journal expenditures. Since 2010, one package has been discontinued and several other packages have been scaled back to control costs.

A strategic and holistic approach to journal evaluation was undertaken in 2012 using objective, quantifiable metrics that took into account multiple vectors of value. CDL and the UC libraries made comparisons of quality, utility and cost effectiveness across all journal publisher packages that UC licenses, focusing on packages that were up for renewal in 2013.

UC Libraries used these metrics to facilitate fiscally responsible decision-making. Our approach sought to re-negotiate the cost of journals, where appropriate, to better align cost to value, and/or to restructure or scale back publisher packages through carefully targeted cancellations to match our available budget. The evaluation methodology intentionally accounted for disciplinary differences to ensure that all disciplines of study at the university would be adequately supported. These results can also be used in future negotiations, with our data in hand.

MEET Chan Li

Collection Development &
Management Group

Chan Li admits she loves statistics and that's a good thing because she is steeped in them! Chan has been a key player in the development of CDL's weighted value algorithm to assess the significance of UC systemwide electronic journals, across disciplines, to our users. This CDL-developed methodology compares each UC ejournal title licensed for systemwide use at the University of California against other UC-licensed titles within the same subject category according to a variety of objective value indicators to arrive at a comparative value for each journal within the UC shared licensing portfolio. The algorithm moves away from more traditional methods of measuring value such as cost and usage, often comparing the results across all disciplines collectively. In this new model, *utility* (usage and numbers of UC citations), *quality* (citation impact factors from two different sources), and *cost effectiveness* (based on two factors) are all considered within each discipline separately.

Chan has worked intensively with CDL and UC campus colleagues to further refine this system, taking into account not only statistical, quantitative data gathered in the study, but also the qualitative experience of more than 30 bibliographers across the UC system. These subject specialists examined the quantitative data from 7,600 ejournals, overlaying their knowledge of their local faculty research needs, growing research areas, and niche specialties of a particular department or campus.

Not only has this algorithm helped in cutting less important titles for our researchers and scholars, allowing us to purchase additional titles effectively, it has also facilitated fact-based discussions with journal vendors. With the results of our study in hand, CDL can demonstrate the relative importance of a vendor's ejournals when compared to others in the same discipline. This allows us to effectively negotiate for favorable pricing. Librarians also save time by easily recognizing high value journals, and can focus on the more problematic titles.

The project also has made waves in the broader library community. Many libraries have cited CDL's approach in publications and several have replicated our approach to do their collection return on investment (ROI) analysis. Examples include the Canadian Research Knowledge Network (CRKN) and the University of Memphis.

Next on the docket for Chan are a further refinement of the algorithm in consultation with UC faculty in statistics and economics— plus formal publication of CDL's groundbreaking advancements in this important enterprise. Chan, lover of statistics, credits "brilliant, hardworking colleagues and luck to have landed at CDL" for stoking her passion.

A pear sitting on ice, Milton Halberstadt, photographer, D-468, Special Collections, UC Davis, © University of California Regents.

http://content.cdlib.org/ark:/13030/kt3n39s043/?query=Halberstadt&brand=calisphere

Creating a sustainable organization for the DMPTool:

CDL made strides in helping our researchers to create, edit, share, and save data management plans to meet funder requirements with its data management planning tool, DMPTool. In fact, this year the California Digital Library and its partners were awarded a $590,000 grant from the Alfred P. Sloan Foundation to fund further development of the popular Data Management Planning Tool (DMPTool) in 2013. The bulk of the grant will go to the UC Curation Center (UC3) at the CDL to fund improvements to the DMPTool including expanded functionality, training modules, documentation and the creation of an open-source community to sustain the DMPTool in the future.

- Josh Hubbard and Bhavitavya (Bhavi) Vedula have joined the CDL as developers on the grant-funded DMPTool2 project.

- Project partners are the University of Virginia Library, University of Illinois at Urbana-Champaign Library, and DataONE.

- Additional DMPTool webinars have been funded by an Institute of Museum and Library Services (IMLS) grant. This partnership and grant funding help us achieve our goal of gaining financial efficiencies.

University of California Curation Center (UC3)

For the past two years, Perry Willett has served as product manager for the DMPTool, a service aimed at helping researchers create ready-to-use data management plans required of specific funding agencies; meet their requirements; obtain instructions and guidance for the plans; and learn about resources at their institutions to fulfill their grants' data management requirements. Perry's background in the humanities (B.A., English Literature and German; M.A., Comparative Literature; M.L.S.), talents in organizing groups, listening to users, and troubleshooting technical issues make him a valuable asset to CDL's University of California Curation Center (UC3). Perry coordinates the project and interacts with partners at the University of Virginia, University of Illinois, the Smithsonian Institution, UCLA, UCSD, DataOne, and the UK Digital Curation Centre, among others. This is an important collaborative project for CDL and the UC libraries. Perry communicates with librarians and IT staff at the nearly 100 institutions who have thus far configured and customized the DMPTool. Communications with users involves much reciprocal discussion—explaining to users what's possible to do in the current version of the tool. Perry describes one aspect of his work as "matchmaking"—hearing from several people on the same campus with similar questions (or answers), unaware that their colleagues are asking similar things, then bringing them together to solve problems.

This year the DMPTool project garnered grants from the Alfred P. Sloan Foundation and the Institute of Museum and Library Services (IMLS) to create a version 2 aimed at adding functionality for institutional administrators to customize the tool independently. Perry brings his exhaustive end-user experience with version 1 of the tool to inform and improve the new version. But the DMPTool is only one of Perry's many challenges. He is also the product manager for UC's Merritt repository service. Building a digital repository from scratch is daunting. CDL staff learned from work on the original Digital Preservation Repository, and knew we required a repository that could do more, particularly by providing public access to the data stored. In thinking through Merritt's requirements, staff also worked more closely with researchers, understanding their unique workflows, requirements and idiosyncrasies. Among the gnarliest preservation issues encountered are the issues of storing huge numbers of digital objects, coupled with the enormous size of some objects themselves—100 gigabytes or more. Eventually, the DMPTool will meet up with Merritt—a good solution for storing and sharing researchers' data. Perry brings an understanding of general research principles to the formidable challenges in creating CDL digital services; with his end-user experience, he represents well the needs of researchers and scholars in a collaborative environment. CDL and the UC libraries are lucky to have Perry's talents in these endeavors!

253
COLLECTIONS

469
THOUSANDS OF OBJECTS

18.43
TERABYTES OF STORAGE
DIGITAL COLLECTIONS AND
OBJECTS IN THE UNIVERSITY
OF CALIFORNIA CURATION
CENTER'S MERRITT
REPOSITORY

Achieving cost recovery for EZID: For tracking and citation management, EZID allows researchers to create and manage persistent identifiers. Its subscribers include academic institutions, government agencies, non-profit organizations and commercial entities. This service reached its initial goal toward becoming self-supporting, and was honored with the DataCite 2012 Gold Award for assigning more than 250,000 DOIs (digital object identifiers) in one year. This accomplishment reflects EZID's growth in 2012, as well as the hard work of our clients and the CDL team in promoting responsible management of research data sets.

Completing virtualization of infrastructure: During FY2012-13, CDL completed the planned infrastructure virtualization activity, which was the first part of a 3-phase project to move away from physical Sun/Solaris systems. By doing so, CDL improves operations, provides easier access to information, and achieves additional flexibility. Phase 1 involved moving more than 30 (of about 50) applications into our virtual environment, allowing us to retire the oldest systems in our environment and reduce the associated support costs. During FY2013-14, CDL is scheduled to complete the second phase of the project, which is to retire Sun/Solaris systems using x86 processors. The final phase, retiring all remaining physical Sun/Solaris systems, is scheduled for FY2014-15.

Implementing the "Paid Up" model for storage costs: Without effective, affordable, and proactive curation management of digital asset, today's resources will not remain viable and useful to future scholars. The Merritt repository, operated by CDL's UC Curation Center (UC3), will shortly be moved to a partial cost recovery operating mode. For the UC community, the price for using the repository is based solely on the amount of digital storage consumed. By taking advantage of a new cloud storage service from UCSD/SDSC, UC3 is able to offer storage prices far below commercial alternatives. Two payment options are available: (1) annual or pay-as-you-go billing, with a baseline rate of $0.40/GB/year; and (2) one-time or paid-up billing, with a baseline rate of $2.93/GB for a 10 year term.

The pay-as-you-go option is appropriate when customers have access to predictable ongoing funding sources; in situations where this is not the case, for example, limited term, grant funded projects, the paid-up option may be more attractive as it provides an opportunity for principal investigators to build into grant budgets the one-time costs sufficient to fund extended use of Merritt for content management, preservation, sharing, and access. Stephen Abrams, associate director

of the UC Curation Center at CDL, developed the paid-up model using a multi-factored formula to estimate various costs over time. The model was reviewed by UC libraries and the larger library and preservation community as a logical, rigorous and welcome alternative to the pay-as-you-go option.

Implementing a Customer Relationship Management (CRM) system: Many CDL stakeholders identified a need for service managers to better track leads for fee-based services and to better utilize online tools for marketing and outreach purposes. It was also acknowledged that such a system may further benefit CDL by either connecting with or replacing existing systems used to communicate with CDL clients and end-users. As a result of an investigation of various CRMs, CDL licensed Salesforce, which is currently being used by several services: EZID, eScholarship, Web Archiving Service, and with other services coming on board in the future. The system provides a more efficient way for service managers to keep track of all interactions with clients as well as enable a view of clients' use of multiple services at CDL.
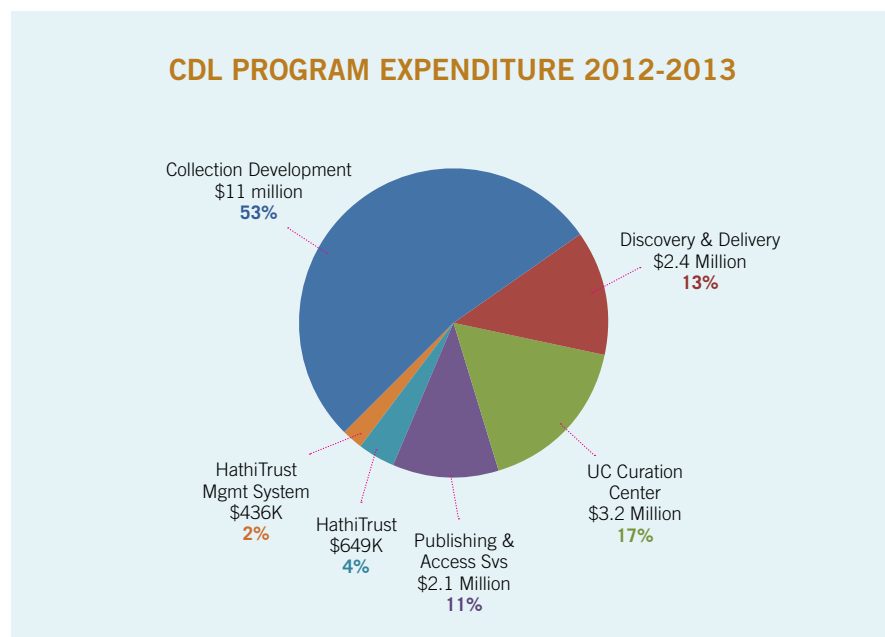

Artichoke, Milton Halberstadt, photographer, D-468, Special Collections, UC Davis, © University of California Regents.
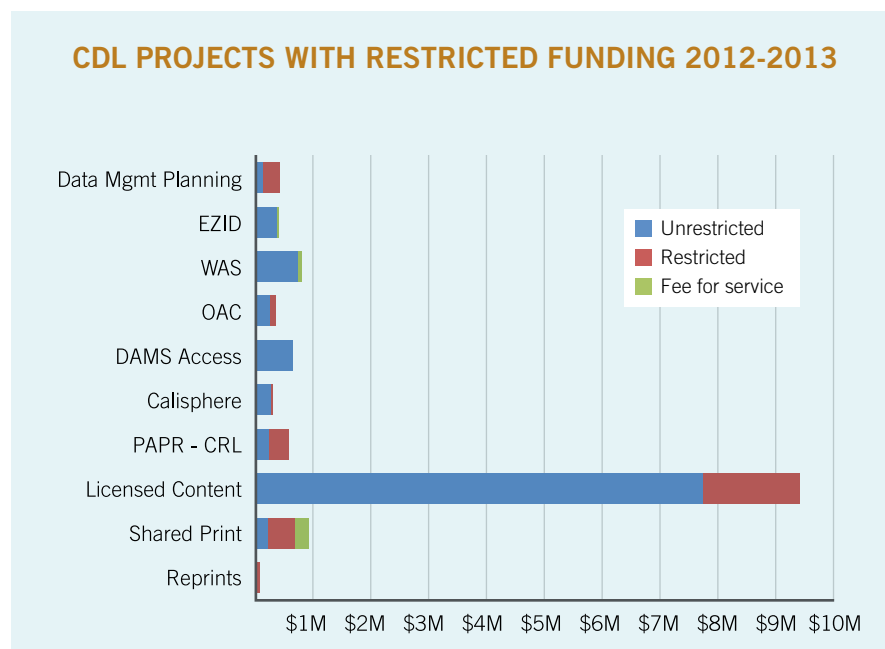http://content.cdlib.org/ark:/13030/kt9b69s5h3/?query=Halberstadt&brand=calisphere

FUNDING

CDL has a budget of approximately $16M from the campus assessment fund, augmented by $1.6M from the state lottery fund as well as $1.5M from grants, fee-for-service or other restricted funds. Following is a chart showing CDL expenditures by program:

CDL PROGRAM EXPENDITURE 2012-2013

- Collection Development $11 million **53%**
- Discovery & Delivery $2.4 Million **13%**
- UC Curation Center $3.2 Million **17%**
- Publishing & Access Svs $2.1 Million **11%**
- HathiTrust $649K **4%**
- HathiTrust Mgmt System $436K **2%**

Following is a chart depicting external funding by CDL program for unrestricted funds, restricted funds (grants), and for fee for services (cost recovery).

CDL PROJECTS WITH RESTRICTED FUNDING 2012-2013

Data Mgmt Planning, EZID, WAS, OAC, DAMS Access, Calisphere, PAPR - CRL, Licensed Content, Shared Print, Reprints

Legend: Unrestricted, Restricted, Fee for service

$1M $2M $3M $4M $5M $6M $7M $8M $9M $10M

MEMBER ORGANIZATIONS

Coalition for Networked Information (CNI)

DataCite

DPN (Digital Preservation Network)

HathiTrust

IMPACT (Improving Access to Text) Project

International Coalition of Library Consortia (ICOLC)

International Internet Preservation Consortium (IIPC)

Library Publishing Coalition

Name Assigning Authority Number (NAAN)

National Digital Stewardship Alliance (NSDA)

National Information Standards Organization (NISO)

Open Access Scholarly Publishers Association (OASPA)

OCLC Research Library Partnership

Preservation and Archiving Special Interest Group (PASIG)

Public Knowledge Project (PKP)

Portico

Scholarly Publishing and Academic Resources Coalition (SPARC)

Western Regional Storage Trust (WEST)

PARTNERS/SERVICE AREA

CERN (SCOAP3)

DataOne (NSF DataNet)

Data Curation Center (DCC)

EDINA (DataCite)

EDUCOPIA (Scholarly publishing)

FORCE11 (Scholarly publishing)

Information Center for the Environment (ICE), **University of California, Davis** (DataONE)

Stanford University Library (JHOVE2)

JISC (EZID)

Lulu (eScholarship PLUS)

Moorea Biocode Project (DataONE)

OCLC (Melvyl, VDX)

Open Context, Alexandria Archive Institute (Merritt)

OSTI (Office of Scientific and Technical Information, Department of Energy) (DataCite)

Portico (JHOVE2)

Peer REview for Publication & Accreditation of Research Data in the Earth Sciences (PREPARDE) (Data publication)

Purdue University Libraries (DataCite)

Public Knowledge Project (eScholarship)

ResearchHub, UC Berkeley (Merritt)

RDA (Research Data Alliance) (Data curation)

SDSC (San Diego Supercomputer Center) (Merritt, WAS)

Stanford University Library (JHOVE2)

USGS (U.S. Geological Survey) (DataONE)

UCSF Clinical & Translational Science Institute (DataShare)

University of California Berkeley's Richard B. Gump South Pacific Research Station (DataONE)

University of Virginia Library (DMPTool)

University of Illinois at Urbana-Champaign Library (DMPTool)

SELECTED CDL STAFF SERVICE

Stephen Abrams

Program Committee, Joint Conference on Digital Libraries (JCDL) (http://jcdl2013.org)

Program Committee, Society for Imaging Science and Technology (http://www.imaging.org/ist/conferences/archiving)

Advisory Board, EU-funded 4C project (Collaboration to Clarify the Cost of Curation) (http://www.4cproject.eu)

Advisory Board, Federal Agency Digitization Guidelines Initiative (FADGI) Still Image Digitization Working Group (http://www.digitizationguidelines.gov/still-image)

Ivy Anderson

Chair, HathiTrust Collections Committee (http://www.hathitrust.org/wg_collections_charge)

Member, HathiTrust Program Steering Committee (http://www.hathitrust.org/psc)

Steering Committee and Chair, SCOAP3 Technical Working Group (http://scoap3.org/committees.html)

Advisory Board, Berkeley Digital Library Copyright Project, UC Berkeley Law School (http://www.law.berkeley.edu/librarycopyright.htm)

Advisory Group, JSTOR Library (http://about.jstor.org/content/advisory-group)

Library Advisory Council, Oxford University Press (http://global.oup.com)

Advisory Board, PubMed Central (http://www.ncbi.nlm.nih.gov/pmc/about/nac)

Patricia Cruse

Advisory Board, CRADLE (Curating Research Assets and Data) (IMLS) (http://www.imls.gov)

Vice President, DataCite Board (http://www.datacite.org/board)

Advisory Committee, Portico (http://www.portico.org/digital-preservation/about-us/advisory-committee)

DataONE Leadership Team (http://www.dataone.org/leadership-team)

Co-chair, DataCite Governance and Sustainability Working Group (http://www.datacite.org)

Advisory & Editorial Board, Scientific Data, Nature Publishing Group (http://www.nature.com/scientificdata/about/advisory-and-editorial-board)

Advisory Panel, Collaborative Stewardship: Providing Sustainable Digital Heritage Training for Tribal Libraries, Archives, and Museums (IMLS) (http://www.imls.gov)

Laine Farley

UC Libraries Representative, HathiTrust Board of Governors (http://www.hathitrust.org/board_of_governors)

Board of Directors, UC Press (http://www.ucpress.edu)

Executive Committee, Western Regional Storage Trust (WEST) (http://www.cdlib.org/services/west/about/governance.html)

SELECTED CDL STAFF SERVICE *(Continued from page 25)*

Barbara Hui

Chair, Technical Committee, Public Knowledge Project (http://pkp.sfu.ca)

Member, LibTech Committee, Open Library of Humanities (https://www.openlibhums.org)

John Kunze

Member, Advisory Board, Dublin Core Metadata Initiative (http://dublincore.org)

Chair, Dublin Core Kernal Working Group (http://dublincore.org/groups/kernel)

Chair, Dublin Core Preservation and Metadata Working Group (http://dublincore.org)

Patricia Martin

Member, Program Committee, Northwest Interlibrary Loan and Resource Sharing Conference (http://nwill.org)

Member, Program Committee, Joint Conference on Digital Libraries 2013 (http://www.jcdl2013.org)

Member, DLF Planning Committee (www.diglib.org)

Lisa Schiff

Co-Chair, ORCID Business Steering Group (http://orcid.org/about/community)

Editorial Board, Journal of Librarianship and Scholarly Communication (http://jlsc-pub.org/jlsc/editorialboard.html)

Joan Starr

Chair, DataCite Metadata Working Group (http://www.datacite.org/node/97)

Carly Strasser

Member, DataOne Working Group on Community Education and Engagement (http://www.dataone.org/working_groups/community-education-and-engagement)

Margery Tibbetts

Member, ELUNA SFX Product Working Group (http://el-una.org/product-groups/sfx-product-group/sfx-product-group-leaders)

Ken Weiss

Editorial board (Technical Editor), Dermatology Online Journal (http://anagen.ucdavis.edu)

Perry Willett

Chair, Council on Library Resources (CRL) Certification Advisory Panel (http://www.crl.edu/archiving-preservation/digital-archives/portico-hathitrust/advisory-panel)

Historical Photographs of Four Ethnicities on California's Central Coast, Black Gold Cooperative Library System. Digitized and made available with Local History Digital Research Project (LHDRP) funds, supported by a Library Services and Technology Act grant administered by the California State Library and coordinated by the Califa Library Group.