



OCLC Research: 2011 Activity Report



We are pleased to provide this activity report to those interested in the work of OCLC Research. It provides highlights of our work, focusing on 2011. The purpose is to dive more deeply into our work, to provide a flavor of important themes, and to point to sources of further information.

Given the nature of the work reported, and our desire to be reasonably comprehensive, this document covers a range of technical and non-technical work. It has a “high acronymic density.” We would be glad to provide more information or clarification about specific areas. Contact us at OCLCResearch@oclc.org.

—Lorcan
Vice President, Research, and Chief Strategist



OCLC Research: 2011 Activity Report

OCLC Research: 2011 Activity Report
OCLC Research

© 2012 OCLC Online Computer Library Center, Inc.
Reuse of this document is permitted as long as it is consistent with the terms of the Creative Commons Attribution-Noncommercial-Share Alike 3.0 (USA) license (CC-BY-NC-SA):
<http://creativecommons.org/licenses/by-nc-sa/3.0/>.
April 2012

OCLC Research
Dublin, Ohio 43017 USA
www.oclc.org

ISBN: 1-55653-446-9 (978-1-55653-446-1)
OCLC (WorldCat): 785423693

Please direct correspondence to:
Lorcan Dempsey
Vice President, Research, and Chief Strategist
OCLCResearch@oclc.org

Suggested citation:
OCLC Research. 2012. *OCLC Research: 2011 Activity Report*. Dublin, O.: OCLC Research.
<http://www.oclc.org/research/publications/library/2012/2011activityreport.pdf>



OCLC Research: 2011 Activity Report

Table of Contents

Introduction	1
1. Shared community R&D—libraries and archives	5
The shared work agenda	5
Collaborative work	10
Prototypes & software	11
2. Advanced development and consultation within OCLC	15
New service introduction	15
Advanced technology development: service support	16
Consulting	18
3. Member and Partner engagement	19
The OCLC Research Library Partnership	20
Publications: contributing to the literature	22
Research grants: supporting library/archival research and education	24
Making connections: Distinguished Seminar Series and other visiting speakers	24
Webinars	26
4. OCLC Research people	27

OCLC Research updates: keeping informed

- *OCLC Research Quarterly Highlights*—a periodic roundup of news, events, publications, prototypes and other activities of OCLC Research designed to serve a broad online readership. <http://www.oclc.org/research/publications/newsletters/quarterlyhighlights/default.htm>
- *Above the Fold*—a weekly current awareness e-newsletter that serves a broad international readership from 5,996 libraries, archives and museums. It consists of a compendium of articles that relate to our work and the information context in which we're all operating—but readers might not see in the course of their regular awareness routines. To subscribe to *Above the Fold*, go to http://www.oclc.org/email/subscribe_abovethefold.htm.
- *OCLC Research Library Partnership Weekly Updates*—a weekly e-mail digest of news and updates for OCLC Research Library Partners that serves more than 1,400 subscribers.

*You research denizens just keep cranking out good stuff—
Good on ya.*

Rick Lugg
Partner, R2 Consulting/Sustainable Collection Services

Introduction

The mission of OCLC Research is to **expand knowledge** that advances OCLC’s public purposes of furthering access to the world’s information and reducing library costs.

OCLC Research is one of the world’s leading centers devoted exclusively to the challenges facing libraries and archives in a rapidly changing information technology environment. Since 1978, we have carried out research and made technological advances that enhance the value of library services and improve the productivity of librarians and library users.

Historically, OCLC Research has been dedicated to dual roles:

1. To act as a community resource for shared Research and Development (R&D), and
2. To provide advanced development and technical support within OCLC® itself.

The aim is to connect those roles for the mutual benefit of OCLC and the community it serves.

OCLC Research has also committed to a third role:

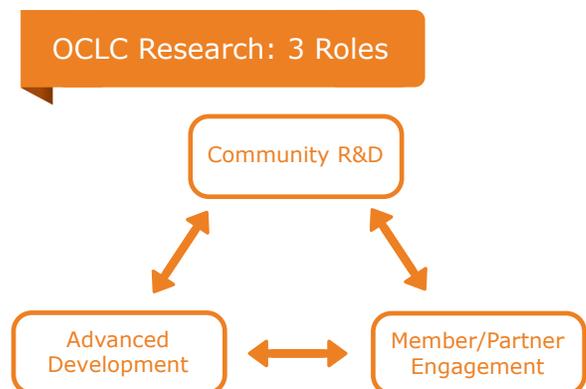
3. To enhance OCLC’s engagement with members and to mobilize the community around shared concerns.

This report pulls together our 2011 work in these three areas.

Shared community R&D

This activity clusters around our **shared work agenda** (see page 5), which gives shape and purpose to our work. This agenda changes over time. Current themes are: Mobilizing Unique Materials, Research Information Management, Metadata Management, System-wide Organization, and User Behavior Studies and Synthesis. We also do community work in Infrastructure and Standards Support.

We collaborate with external partners where it makes sense to amplify our work, and we note some of those collaborations throughout this discussion.



What you can't see so easily is how they really help us effect change at the local level.

David Farneth
Assistant Director, Getty Research Institute

Advanced development and consultation within OCLC

The addition of the **Innovation Lab** in 2011 helps us address our second role of providing advanced development and technical support within OCLC itself. OCLC Research, including the Innovation Lab, is an important source of expertise, development and consultation with the enterprise. The Lab is a technology-based incubator within OCLC Research, providing a framework within which OCLC staff, members and commercial organizations can collaborate to advance ideas for new services, service improvements or technology adoption.

The Innovation Lab has accelerated the introduction of new services within OCLC under broad categories of mobile, social, analytics and services for the smallest libraries. It is an important facilitator of cross-company discussion, helping to align technology and service options, creating a bridge between OCLC Research and product groups, providing advice or rapid consulting, and jump-starting shared initiatives.

We are making leading contributions to the ongoing development of mobile, “big data” infrastructure (Hadoop), bibliographic processing, linked data, analytics, data architecture and other areas. Several OCLC products and services originated in OCLC Research last year, and others are being incorporated into services this year. We are the main focus of expertise about archives in OCLC. We provide specialist expertise where appropriate; for example, we have economic expertise that was leveraged during “Cost Share Model” discussions.

OCLC Member and Partner engagement

The **OCLC Research Library Partnership** was launched in July of 2011, succeeding the RLG Partnership. It has grown in membership and is a leading venue for research institutions to convene, and to connect with OCLC. It is an important contribution to enhancing our third role of engaging with members and mobilizing the community around shared concerns. It provides a venue for research institutions to engage with OCLC to work cooperatively on designing library futures. The OCLC Research Library Partnership has 158 dues-paying members and is a global alliance of like-minded institutions that focuses on making operational processes more efficient and shaping new scholarly services by directly engaging senior managers. It is supported by the full capacities of OCLC Research, informed by an international, system-wide perspective and connected to the broad array of OCLC products and services.



Titia van der Werf

We are deeply interconnected through personal and formal engagement with library and archival institutions and organizations, and we represent OCLC interests and positions in our relations with them. We also represent library and archive interests in internal OCLC discussions, and regularly provide context or introductions for colleagues.

We are very pleased to be expanding our footprint in Europe. Titia van der Werf began working as a program officer in Leiden early in 2012, joining Rob Koopman of the Innovation Lab. We have also hired a research scientist, Shenghui Wang, Ph.D., to be based in Leiden.



Rob Koopman

I think your [report] will both distress and confuse some folks in the profession, but it is thoughtful, provocative, and important.

Mark Greene
University of Wyoming

OCLC Research staff regularly make keynote presentations and publish in a wide range of venues, making a contribution to discussion and the literature in our field. We are pleased to collaborate with ALISE (the Association of Library and Information Science Education) on a grant program to faculty in iSchools/library schools each year.

This document

This report provides highlights of OCLC Research activities, **focusing on 2011**. The purpose is to dive more deeply into our work, to provide a flavor of important themes, and to point to sources of further information. It reviews our internal work and provides an overview of our external shared work agenda. And it presents recent outputs, which take the form of prototype systems or services, as well as published reports, webinars, podcasts, videos and meetings of all types ranging from task group sessions to our three-day symposium, FutureCast. We feel it presents a story of achievement and contribution, representing significant value provided to OCLC, the membership, and the larger community.

I find a very welcoming atmosphere for research libraries of all sizes, and the issues we've discussed are topics my own library is engaged in. I am very optimistic about where we are going with OCLC.

Suzanne E. Thorin
Dean of Libraries and University Librarian, Syracuse University

OCLC Research Library Partnership Working Groups



1. Shared community R&D—libraries and archives

Our shared work agenda, organized in themes, is central here. This was developed to recognize areas where OCLC has standing (e.g., Mobilizing Unique Materials or Metadata Management, as we work to connect bibliographic data and practices to a broader future), and to develop standing in areas the community told us were of interest (e.g., Research Information Management, User Behavior Studies and Synthesis). We have amplified this work in some areas through collaboration.

The shared work agenda



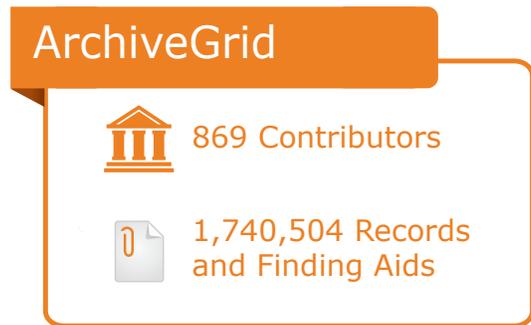
Mobilizing Unique Materials

Can libraries and archives efficiently deliver on the promise of unique materials? We're working to change both parts of the cost/benefit equation associated with special collections and archives. Libraries need to make these materials cheaper to describe and disclose. They need to invest only what is necessary to make them discoverable effectively.

They need to make them more valuable to the institution's teaching and research roles; its faculty, students and citizens; and to be sure the institution's stewardship contributes to its prestige.

Selected activities:

- ▶ Developed an experimental ArchiveGrid® discovery system that provides free access to an increasingly large number of archival collections and provides a real-world environment in which to test changes in archival descriptive practice.
- ▶ Developed a “Well-intentioned practice” for putting digitized collections of unpublished materials online that was endorsed as a standard by the Society of American Archivists.
- ▶ Published four reports (full details listed in publications section):
 - *Taking Stock and Making Hay: Archival Collections Assessment*
 - *Rapid Capture: Faster Throughput in Digitization of Special Collections*
 - *Single Search: The Quest for the Holy Grail*
 - *Scan and Deliver: Managing User-initiated Digitization in Special Collections and Archives.*
- ▶ Held webinars and led working groups on each of the topics above.



Research Information Management

How can libraries effectively support the research missions of their universities and the workflow of their faculty? Libraries can create a cluster of high-value, distinctive services by directly supporting researchers, managing and aggregating the research output, and ensuring that it enhances the reputation of the faculty and the institution. What services do scholars need? How can their output be preserved, discovered and re-used? What can be done to ensure that the proper acknowledgment accrues to the correct entities?

Selected activities:

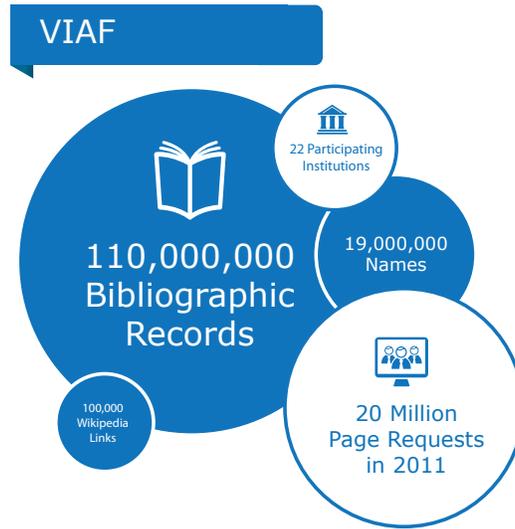
- ▶ Initiated new work items around the sustainability of disciplinary repositories and library roles in data curation. We are also beginning work to explore subject-specific research information liaisons—staff dedicated to a discipline or department who engage with researchers through the entire lifecycle of the research workflow. We will be seeking feedback and Research Library Partner engagement on these topics soon.
- ▶ Published *Supporting Research: Environments, Administration and Libraries* (full details listed in publications section).
- ▶ A report synthesizing studies of researchers' information needs is forthcoming.
- ▶ We are pleased to be joined by Ixchel Faniel, Ph.D., from University of Michigan, a postdoctoral researcher specializing in research data re-use and related workflow, technical and other issues.

Metadata Management

How will we evolve our metadata structures to work effectively in a semantically richer Web or to connect to network hubs? How can we make our data work harder, to release greater value in systems and services? We're working to make the future investment in descriptions of library materials less costly, more responsive to network-level



discovery requirements, and more consistent with user expectations. We're making the accumulated investment in library metadata, terminologies and authorities yield new returns by recombining it to create new relationships and associations that are valued by a wider Web community.



VIAF™: The Virtual International Authority File.

Selected activities:

- ▶ Completed background work on linked data, data modeling and architecture, and vocabularies. This is showing up in OCLC prototypes, but is also influencing OCLC's position in the coming transformation of our bibliographic infrastructure (RDA, linked data, etc.). We are exploring library extensions to Schema.org. We are planning an externally focused work item to move some of this work in community activity ("beyond MARC").
- ▶ Enhanced the Virtual International Authority File and added countries. Agreed, with the national library Principals, on the framework within which it would transfer from being an OCLC Research prototype to being part of the OCLC service array in spring 2012.
- ▶ Released a set of tools and experimental services around FAST, an algorithmically derived simplification and "faceting" of Library of Congress Subject Headings. We released a Linked Open Data version in December 2011, and bulk downloadable versions in February 2012. FAST has generated considerable interest as a subject alternative.
- ▶ Managing the Metadata Crosswalk service, which we developed and which is now internally deployed to support various services. We will work with production to make this a Platform App. It carried out 79,072,619 transforms in 2011.
- ▶ Published three reports (full details listed in publications section):
 - "Publisher Names in Bibliographic Data: An Experimental Authority File and a Prototype Application"
 - *Social Metadata for Libraries, Archives and Museums Part 1: Site Reviews*
 - *Social Metadata for Libraries, Archives and Museums Part 2: Survey Analysis*.
- ▶ A fourth report, *Social Metadata for Libraries, Archives, and Museums Part 3: Recommendations and Readings*, is being published in April 2012.



Identities has a summary page for every name in WorldCat.

System-wide Organization

How is the network reconfiguring how libraries organize to provide their services? We are working with libraries to re-imagine library services as an interdependent system of collections and support. By providing the concepts to guide this reinvention, the evidence on which management decisions can be made and a venue where stakeholders can agree on new directions, we're promoting new shared solutions for collections and services.

Selected activities:

- ▶ Produced a framework to enable academic research libraries to model the costs and opportunities associated with increased reliance on collections managed by the HathiTrust and shared print storage repositories.
- ▶ Identified requirements for sustainable business partnerships among collection-based service producers and consumers in the nonprofit educational sector.
- ▶ Oversee SHARES, a self-selected and self-governing global resource-sharing partnership of 95 research libraries in which participants share their professional expertise as well as research materials. SHARES participants also staff task forces that currently focus on sharing of special collections, creating a Web-based ILL cost calculator, and exploring the international implications of ILL language in e-resource licenses.

Trusted partners make all the difference! Our patrons are consistently amazed by the materials we are able to provide them with and that is due in large part to the dedicated SHARES community.

Aimee Lind
Getty Research Institute

- ▶ Important work in hand that will appear in 2012 includes characterizing a national presence in the collective collection as a whole (testing with Scotland) and analyzing collections and shared print in a North American “megaregions” framework.
- ▶ Performed the most comprehensive study of academic library circulation ever undertaken, and made the data file available for free public download from the OCLC Research website (OhioLINK).
- ▶ Responded to many requests for custom data mining or specialist advice as libraries plan to manage down their print collections and move to shared arrangements.
- ▶ Published two reports (full details listed in publications section):
 - *Cloud-sourcing Research Collections: Managing Print in the Mass-digitized Library Environment*
 - *OhioLINK-OCLC Collection and Circulation Analysis Project 2011.*

Fascinating result of OhioLINK-OCLC Collection and Circ. Analysis Project <http://j.mp/o01XWi>, 80% of circ driven by 6% of collection

Larry Fischer (via Twitter @ larry_fischer)
 Web Content Coordinator, Greene County Public Library
 London, Ohio

User Behavior Studies and Synthesis

What do we know about changing user behaviors? We’re helping libraries calibrate their services around a completely new set of user expectations that have been shaped by consumer technologies. By providing the library community with behavioral evidence on their users’ perceptions, habits and requirements, we can ensure that the design of future library services is all about the user. Our efforts here are amplified by strategic partnerships.

Selected activities:

- ▶ Consolidated our relationship with JISC, who asked us to do more work in this area. Also won another IMLS grant to continue work on information seeking.
- ▶ Published our findings in articles, reports and a video (full details listed in publications section):
 - “Visitors and Residents: What Motivates Engagement with the Digital Information Environment?”
 - “‘If It Is Too Inconvenient, I’m Not Going After it.’ Convenience as a Critical Factor in Information-Seeking Behaviors”
 - *Seeking Synchronicity: Revelations and Recommendations for Virtual Reference*
 - “‘Are We Getting Warmer?’ Query Clarification in Live Chat Virtual Reference”
 - “On Virtual Face-work: An Ethnography of Communication Approach to a Live Chat Reference Interaction”
 - Video: “Digital Visitors and Residents: What Motivates Engagement with the Digital Information Environment? An Update on Current Findings.” Interview with Lynn Silipigni Connaway and David S. White by Lorcan Dempsey. 13 October 2011. 21:23. Posted January 2012. <http://youtu.be/ZCBoLWynsl8>.

Infrastructure and Standards Support

Can we bring a library interest to bear on standards work? Can we make sure important standards activity in the library and archival world has appropriate technical support? Our work here focuses on participating in needed architecture and standards work, supporting adoption and sometimes providing infrastructure support.

Selected activities:

- ▶ OCLC is the maintenance agency for the NISO OpenURL Standard. OCLC Research maintains the OpenURL Registry in support of this standard.
- ▶ We hold leadership positions in NISO, DCMI and W3C and participate in Internet2 and OIXsm Open Identity Exchange. We are closely involved in archival (EAD) and library (FRBR, SRU) standards work. We contribute to the Library Linked Data working group of the W3C.
- ▶ We coordinate monthly NISO and DCMI webinars for all OCLC staff.



~836 Million PURLS Resolved in 2011

Collaborative work

OCLC Research works with a variety of organizations, often as a partner in grant-funded work. This is an important aspect of amplifying our capability, and of developing relationships with important players. We have been a source of expert input and interaction in discussions with HathiTrust. We have proposed some project work in Europeana and DPLA contexts, and hope we can develop some collaboration. We have good personal relationships across a wide spectrum of partners.

Selected collaborations

With IMLS support:

- ▶ Lynn Silipigni Connaway, with Marie L. Radford and Chirag Shah (Rutgers University):
Cyber Synergy: Seeking Sustainability through Collaboration between Virtual Reference and Social Q&A Sites,
Funder: IMLS
- ▶ Ixchel Faniel and Elizabeth Yakel (University of Michigan):
Dissemination Information Packages for Information Reuse (DIPIR),
Funder: IMLS

With JISC support:

- ▶ Lynn Silipigni Connaway and David White (University of Oxford),
Co-Principal Investigators, and Donna Lanclos (UNC Charlotte), Co-
Investigator: Visitors and Residents: What Motivates Engagement with
the Digital Information Environment? Funder: JISC (UK)
- ▶ Brian Lavoie and Neil Beagrie (Charles Beagrie, Ltd.): Digital
Preservation Benefit Analysis Tools, Funder:
JISC (UK)
- ▶ Eric Childress and Ed Chamberlain (University of Cambridge): COMET (Cambridge Open METadata) Project,
Releasing Cambridge Data as Linked Data, Enhanced with VIAF and FAST Links, Funder: JISC (UK)

OCLC Numbers Added...



~ 6 Million



~ 3.7 Million

With RLUK support:

- ▶ Jackie Dooley: Survey of Special Collections and Archives in the UK and Ireland, with Research Libraries UK (RLUK)

Other:

- ▶ Brian Lavoie: Supporting Michele Alexopoulos at University of Toronto, who is analysing bibliographic data to measure innovation diffusion (Canada)
- ▶ Brian Lavoie: The Scottish Presence in the Global Library Resource. Working with the five largest research libraries in Scotland to look at Scottish collection overlaps in Scotland and the Scottish presence in the “diaspora” (UK)
- ▶ User-Centered Design of a Recommender System for a “Universal” Library Catalog (User Behavior Studies and Synthesis), Lynn Silipigni Connaway, OCLC lead, in collaboration with Paul Clough, Barbara Anne Sen, and Simon Wakeling at University of Sheffield (UK), supported by the Arts and Humanities Research Council
- ▶ Constance Malpas: An analysis of holdings in the Research Collections and Preservation Consortium (ReCAP) storage facility to support a planning grant from the Mellon Foundation that will enable them to rationalize their individual collections and develop a prospective strategy for building a stronger, shared collection; ReCAP partner libraries include Columbia, Princeton, and the New York Public Library (NYPL)

Prototypes & software

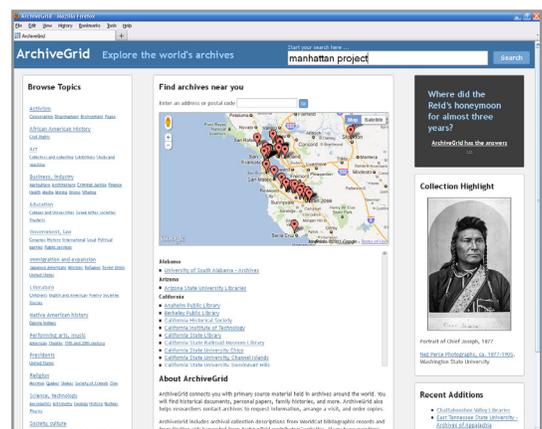
Prototypes are a major output of OCLC Research activities. These serve a number of purposes, from demonstrating the viability of a concept to providing a trial service to validate against sustained use or community expectations. Some of these will transition to OCLC services. Our prototypes are available from <http://www.oclc.org/research/activities/researchworks.htm> and <http://experimental.worldcat.org>. We also occasionally make open source software available. We present a sampling of metrics here, which are for 2011 unless stated otherwise.

Most software is now in maintenance mode. It includes facilities for exchanging museum data (Museum Data Exchange and OAIcat Museum), updated this year, and software to support the Open Archives Initiative and bibliographic retrieval.

ArchiveGrid

OCLC Research is developing a new, freely available discovery system for ArchiveGrid® to replace the current subscription service. This discovery service provides access to detailed archival collection descriptions, making information available about materials held in archives throughout the world.

- ▶ Still in start-up mode as an OCLC Research prototype; transitioning to a free OCLC service managed by OCLC Research
- ▶ Number of contributors: 869
- ▶ Number of records and finding aids: 1,740,504
- ▶ More information:
<http://www.oclc.org/research/activities/archivegrid/default.htm>



Sample page from ArchiveGrid prototype.

Article Exchange

This document-sharing site provides a single, secure location where lending libraries can place requested PDF and TIFF articles and library users can retrieve articles or book chapters obtained for them via interlibrary loan.

- ▶ Transferred to production
- ▶ Available at: <https://share.worldcat.org/ILL/articleexchange/FileUploadLogin>

Classify

This FRBR-based prototype was designed to support the assignment of classification numbers and subject headings for books, DVDs, CDs, and other types of materials.

- ▶ To be integrated into cataloging services this year
- ▶ Total page requests: approximately 17 million
- ▶ Integrated into the popular MARCEdit tool; will be added to Connexion® in mid-year release
- ▶ More information: <http://www.oclc.org/research/activities/classify/default.htm>

Wow, what a time saver [Classify] is to a lone busy high school librarian. Thanks!

Irene Forbes
Librarian, St Andrew's College
Grahamstown, South Africa

WorldCat Genres

Genre profiles allow users to browse genre terms for hundreds of titles, authors, subjects, characters, places, and more, ranked by popularity in WorldCat.

- ▶ 3.5 million works
- ▶ 14.3 million editions
- ▶ More information: <http://www.oclc.org/research/activities/genres/default.htm>

VIAF (Virtual International Authority File)

A joint project with the Library of Congress, the Deutsche Nationalbibliothek, and the Bibliothèque nationale de France, in cooperation with an expanding number of other national libraries and other agencies, VIAF is an OCLC Research activity that explores virtually combining the name authority files of national level authority files into a single name authority service.

- ▶ Transferring to production early 2012
- ▶ Total page requests: more than 20 million
- ▶ 19 million names
- ▶ 110 million bibliographic records
- ▶ More information: <http://www.oclc.org/research/activities/viaf/default.htm>

The screenshot displays a VIAF record for Marc Chagall. At the top, there is a small image of a painting and a list of preferred forms for the name 'Chagall, Marc, 1887-1985' in various languages and scripts, including Hebrew and Arabic. Below this, a 'Preferred Forms' section lists several variations of the name with checkboxes. To the right, a network diagram shows a central node for the VIAF record connected to numerous other nodes representing member libraries, with a legend indicating the location of each library (e.g., RERO in Switzerland).

Part of a VIAF authorities page.

I use the VIAF almost exclusively for authority work when trying to establish foreign (or possibly foreign) authors.

Jennifer Erica Sweda
Original Cataloging, Van Pelt Library, University of Pennsylvania

SearchFAST

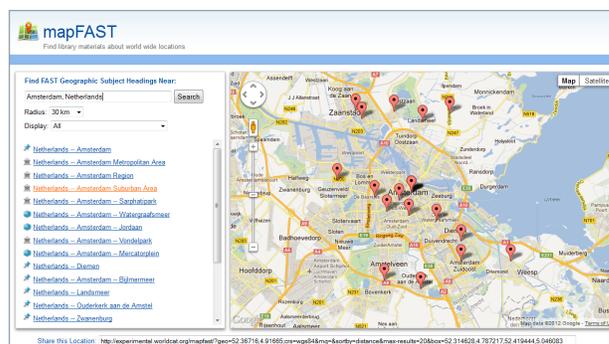
This is a new user interface (as of 11/18/2011) for identifying and accessing FAST (Faceted Application of Subject Terminology) authority records

- ▶ Total visits: 5,139
- ▶ Unique visitors: 2,998
- ▶ Page views: 60,924
- ▶ Top three countries: United States, New Zealand, Australia
- ▶ Available at: <http://fast.oclc.org/searchfast/>

mapFAST

A Google Maps mashup that allows users to access FAST headings for specific locations, as well as FAST headings for surrounding locations that are also FAST subjects.

- ▶ Total visits: 3,281
- ▶ Unique visitors: 2,600
- ▶ Top three countries: United States, England, Germany



Sample mapFAST display.

FAST as Linked Data

This experimental Linked Open Data service provides (from 12/14/2011) another means of accessing the FAST data set.

- ▶ Total visits: 1,578
- ▶ Unique visitors: 2,492
- ▶ Page views: 19,458
- ▶ Pages not viewed (nonbrowser traffic, 26%): 7,034
- ▶ More information: <http://www.oclc.org/research/news/2011-12-14.htm>
- ▶ FAST data set also available for bulk download. For more information: <http://www.oclc.org/research/news/2012-02-15.htm>

RedLaser

RedLaser is a free barcode-scanning shopping application for iPhone, Windows Phones, and Android that uses the WorldCat API to help users find books in a nearby library or store

- ▶ 2011 total usage: 2,984,819 searches for ISBN availability by location

2. Advanced development and consultation within OCLC

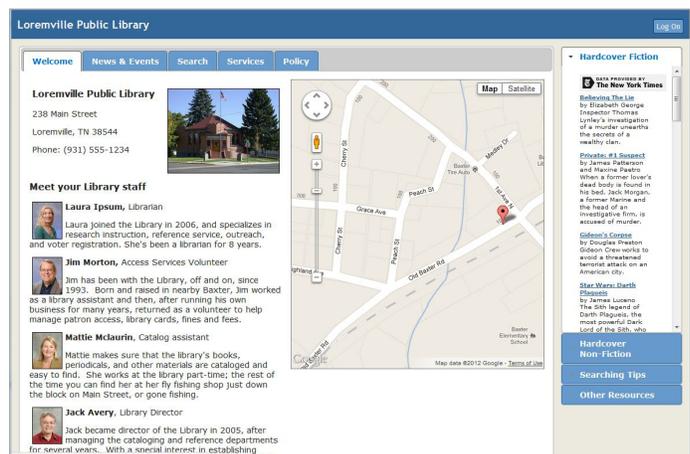
OCLC Research provides development and consultation support across a wide range of OCLC activities, from systems development work to large scale bibliographic processing, tool and widget building, custom statistics and reporting, and consultation. Some of this work originates in support of activities transferred into production. Much of it is generated by request, based on particular expertise. One goal in 2012 is to explore moving more OCLC Research work to the OCLC WorldShare™ platform.

New service introduction

Website for Small Libraries (WSSL) was introduced as a cloud-based service accessible to the small public library segment: <http://beta.worldcat.org/lib>. It is being rolled out in 2012 to libraries, in collaboration with WebJunction colleagues. By the end of 2011, twenty pilot libraries were exploring using it.

Article Exchange was introduced as a safe, cloud-based method of exchanging copyrighted digital content with patrons. It provides an alternative to expensive and difficult to use services, for libraries exchanging copyrighted digital content. It was released as an experimental service at ALA Annual last summer at <http://experimental.worldcat.org/AE>, and has now moved to full production release inside the OCLC ILL Platform.

A URL shortener service was introduced at <http://oc.lc>. This has a variety of potential uses.



Sample website generated by WSSL.

Article Exchange

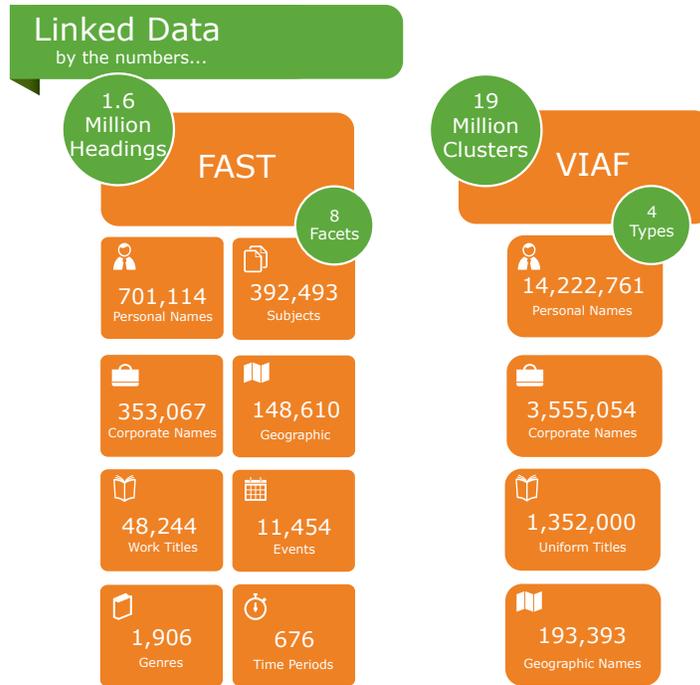
35,526 Uploads
1.6 Megabytes Average File Size
550 Approximate Peak Uploads Per Day
1,750 Average Files On Hand



* Since launch July 13, 2011

Advanced technology development: service support

Linked Data: OCLC Research has pioneered the use of Linked Data in the community and at OCLC. That expertise is now being mobilized as we explore an enterprise Linked Data strategy internally leading to production of architectural designs. We are coordinating modeling of data and other product/service preparatory work.



Big data infrastructure. OCLC Research pioneered the use of clusters and the Hadoop framework for efficiently processing large amounts of data. We assisted in getting Hadoop pilots off the ground. We continue to provide training and support in the use of Hadoop.

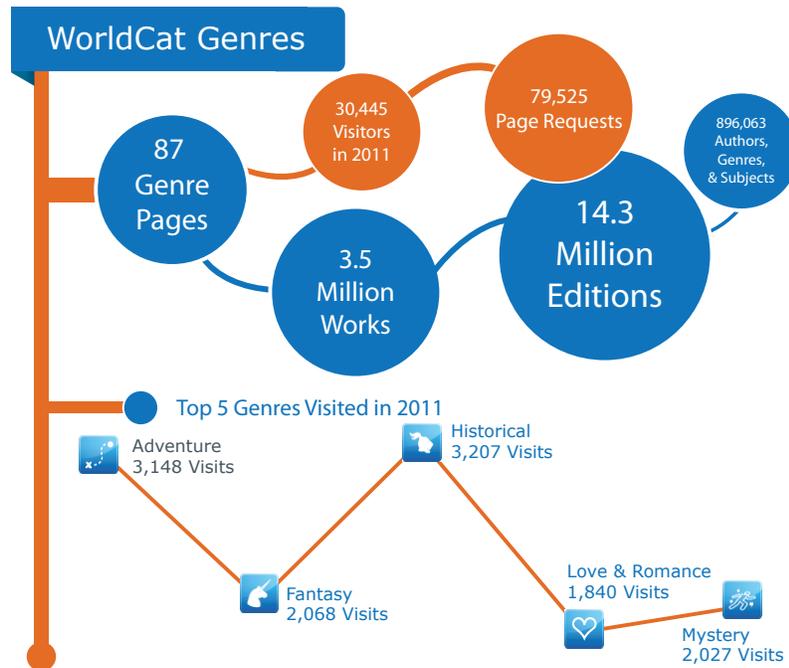
Mobile. OCLC Research staff initially prototyped a proof-of-concept service and then assisted the production team in building a production mobile WorldCat.org™ interface. We continued to coordinate WorldCat® Mobile efforts through the beginning of 2011.

I LOVE how you can email the title and add a note very easily. Thank you again...for keeping libraries in on the cutting edge and in this case, people's hands [with the WorldCat Mobile app].

Tara Bannon Williamson
Senior Librarian, Park Hill Library

Bibliographic data infrastructure.

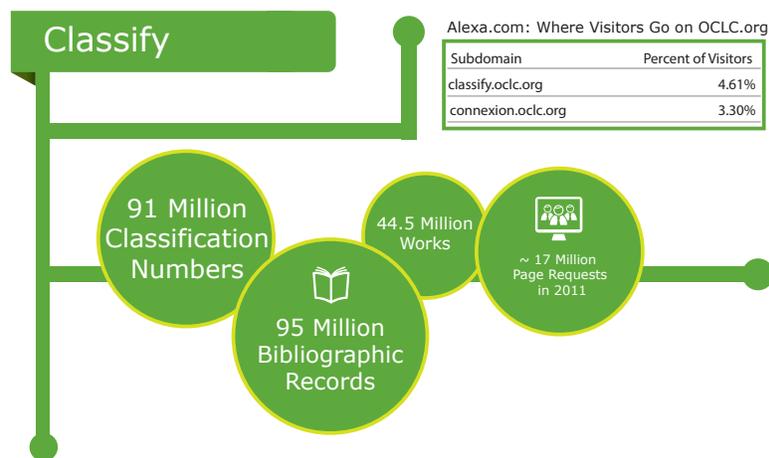
- ▶ **Crosswalk Web Service.** This is now in routine production use. OCLC Research developed it and continues to manage it. We are working to deploy it as a platform service so additional libraries can use it for translating metadata records between formats. It was used to carry out 79,072,619 metadata transformations in 2011.
- ▶ **WorldCat Identities and Genres.** OCLC Research processes data for these services and makes them available to production for deployment in WorldCat.org.



WorldCat Genres allows users to browse fiction and film genres via bibliographic characteristics.

- ▶ **Virtual International Authority File.** This integration of 20 national authority files is maintained in OCLC Research. In 2012 it transitions into production management as an OCLC service. It provides the basis for our support for ISNI (International Standard Name Identifier) work and will be used in various ways to support services.
- ▶ **BISAC-Dewey.** BISAC™ is a classification system used in the book trade. OCLC Research supports use of BISAC and mappings to Dewey® in OCLC services, especially important for interactions with publishers. We have also enriched WorldCat with BISAC headings.

- **Classify.** This is a very popular experimental service that provides workflow support for classifiers. We are preparing for it to be integrated into Connexion, OCLC’s cataloging service, this year.



Classify is an experimental classification Web service.

Analytics.

- **Collection analysis dashboard.** Provided analysis, support and development to “productionize” the ability to rapidly develop collections-based library profiles (processed on Hadoop).
- **Global Library Statistics.** Based on initial work done by OCLC Research, this service aggregates statistics from third-party sources to represent the total global library universe. It has now transitioned to production. <http://www.oclc.org/globallibrarystats>.
- Worked with OCLC Analytics Portfolio team on a pilot project testing a strategy for **registering print archiving commitments** in WorldCat. Pilot participants include: UCLA, UCSD, UC SRLF, Indiana University, University of Oregon, Stanford University, University of Minnesota and CRL.

Name finding system and method
(United States Patent 7877375) issued to Ralph LeVan 25 Jan. 2011

Consulting

We consult regularly, and have areas of deep expertise: retrieval protocols, computational linguistics, economic theory, usage studies, Semantic Web technologies, and bibliographic formats and standards.

OCLC Research provides a very important benefit to the research library community in that there are a lot of things that they work on that aren’t really the province of any one institution and that any one institution can’t handle on their own.

Richard Szary, AUL for Special Collections, UNC at Chapel Hill

3. Member and Partner engagement

OCLC Research staff are richly interconnected with the library and archives community. This includes the developer community, policy/funding agencies and organizations, and individual libraries. We present frequently, visit and engage through working groups and collaborations.

We publish regularly, and provide a venue for interesting and engaging presenters through our distinguished speakers series.

The OCLC Research Library partnership has become a major focus of engagement with research institutions, connecting them to each other and to OCLC capabilities.

Expanding our Global Horizons

We are very pleased that we are developing a European footprint to improve engagement opportunities outside the US. We now have Rob Koopman as a developer in Leiden working with the Innovation Lab, and since the beginning of 2012, Titia van der Werf has been working as a program officer. We also have appointed a research scientist who will join the Leiden office during the second quarter of 2012. This will provide the nucleus for more outreach and collaboration within Europe.



Good rule for RBMS sessions: go to the ones led by OCLC Research folks. #rbms11

Tom Hyry (via Twitter @thyry)
Director of Library Special Collections at UCLA

OCLC and Special Collections



Jackie Dooley



Merrilee Proffitt

OCLC Research provides a central venue for OCLC engagement with the archival community. Recent activities include:

- Developed a new discovery system for ArchiveGrid that provides free, comprehensive access to an increasing number of archival collections to scholars and researchers everywhere. It will replace the current OCLC ArchiveGrid subscription service by June 2012.
- Developed a “Well-intentioned practice” for putting digitized collections of unpublished materials online that was endorsed as a standard by the Society of American Archivists.
- Jackie Dooley serves as Vice-President/President Elect of the Society of American Archivists.
- Merrilee Proffitt serves on the Society of American Archivists Technical Subcommittee for Encoded Archival Description (TS-EAD) and participates in its revision of the EAD standard.
- Merrilee Proffitt also advises and/or contributes to two IMLS-funded Orbis Cascade Alliance archival projects, as well as the Mellon-funded University of Oxford’s FutureArch project to transform the Bodleian Library’s capacity to manage its archives and manuscripts to meet the requirements of the present and future research community.

The OCLC Research Library Partnership

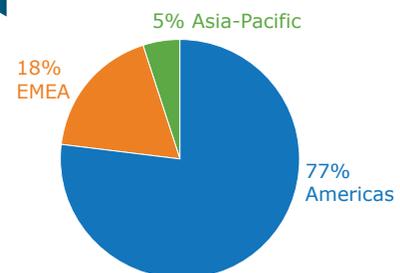
A renewed commitment to the research library community

The OCLC Research Library Partnership was launched in July 2011, succeeding the RLG Partnership. This represented a major commitment by OCLC to research institutions, and resulted in a 44 percent increase in Partners—from 110 RLG Partners at the end of June 2011 to 158 OCLC Research Library Partners by early 2012. Considered globally, 121 Partners are from North America, 29 are from Europe, and eight are from the Middle East and the Pacific Rim.

The 2011 OCLC Research Library Partnership includes:

- 85% of the top 26 institutions in the Times Higher Education World University Rankings (22)
- 50% of ARL members (64)
- 62% of RLUK members (20).

Partnership



We have been very gratified by this response and are building on this success as we move forward. We engage with Partners in multiple ways and have been pleased at their enthusiastic participation. The Partnership influences and contributes to the work agenda. This activity gives OCLC and the Partners a venue for mutual interaction and influence.

I am a member of the Social Metadata Working Group. This group is quite large and there are 23 of us from different institutions in five countries. It has been great to work with other professionals from institutions such as Yale, Stanford, Berkeley and Getty. In the normal course of my day job I have a high level of contact with other national libraries and institutions at the cutting edge of digital technology, but to have this opportunity to research a specific topic in detail and have ongoing discussions with the group over two years was really rewarding.

Rose Holley
Digital Library Specialist, National Library of Australia

FutureCast: Shaping Research Libraries in a Networked Age

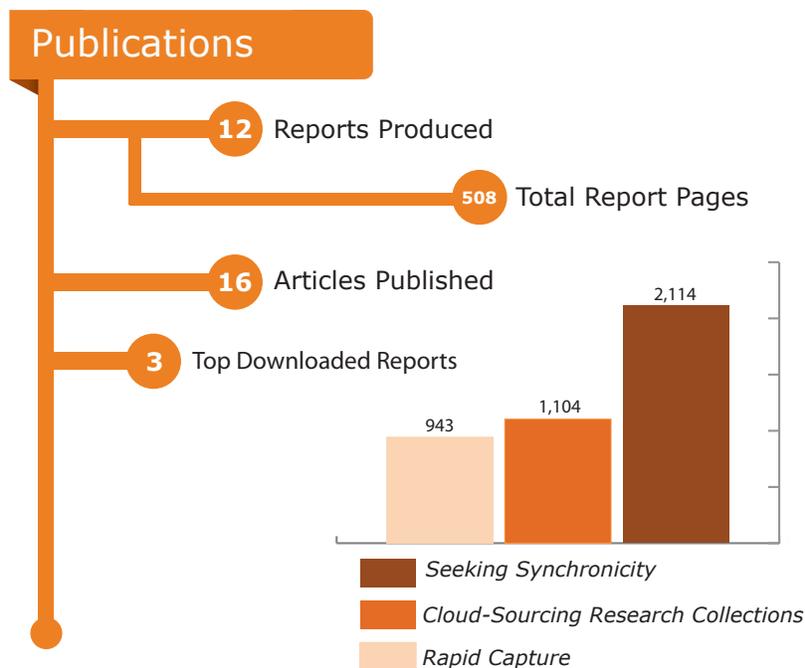
This three-day meeting provided an important glance at the future environment that will impact our work: higher education, electronic publishing and consumer technology. More than 100 OCLC Research Library Partners and invited guests attended the meeting and participated in discussions that led to a collective forecast of what the impact will be on the research library enterprise, how we can prepare for the future, and what we need to do take us from here to there. This enabled us to establish a shared context and identify important areas that provide the foundation on which we can build an action agenda.

Many other OCLC Research Library Partnership activities are represented throughout this document.



100% of attendees thought
FutureCast was worth their time

Publications: contributing to the literature



OCLC Research publications

Bennett, Rick. 2011. "mapFAST: A Research Prototype for Finding Books using a Map." *NextSpace 18* (May). <http://www.oclc.org/nextspace/018/research.htm>.

Connaway, Lynn Silipigni, and Marie L. Radford. 2011. *Seeking Synchronicity: Revelations and Recommendations for Virtual Reference*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/reports/synchronicity/full.pdf>.

Conway, Martha O'Hara, and Merrilee Proffitt. 2011. *Taking Stock and Making Hay: Archival Collections Assessment*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2011/2011-07.pdf>.

Erway, Ricky. 2011. *Rapid Capture: Faster Throughput in Digitization of Special Collections*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2011/2011-04.pdf>.

Lavoie, Brian. 2011. *The Top 25 US Public Libraries' Collective Collection, as Represented in WorldCat*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2011/2011-02.pdf>.

MacColl, John, and Michael Jubb. 2011. *Supporting Research Environments, Administration and Libraries*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2011/2011-10.pdf>.

Malpas, Constance. 2011. *Cloud-sourcing Research Collections: Managing Print in the Mass-digitized Library Environment*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2011/2011-01.pdf>.

OhioLINK Collection Building Task Force, Julia Gammon, and Edward T. O'Neill. 2011. *OhioLINK-OCLC Collection Analysis Project 2011*. Dublin, Ohio: OCLC Research, in collaboration with OhioLINK. <http://www.oclc.org/research/publications/library/2011/2011-06.pdf>.

Prescott, Leah, and Ricky Erway. 2011. *Single Search: The Quest for the Holy Grail*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2011/2011-17.pdf>.

Schaffner, Jennifer, Francine Snyder, and Shannon Supple. 2011. *Scan and Deliver! Managing User-initiated Digitization in Special Collections and Archives*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2011/2011-05.pdf>.

Smith-Yoshimura, Karen, and Cyndi Shein. 2011. *Social Metadata for Libraries, Archives and Museums Part 1: Site Reviews*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2011/2011-02.pdf>.

Smith-Yoshimura, Karen, Carol Jean Godby, Helice Koffler, Ken Varnum, and Elizabeth Yakel. 2011. *Social Metadata for Libraries, Archives, and Museums. Part 2, Survey Analysis*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2011/2011-03.pdf>.

External publications

Bennett, Rick, Edward T. O'Neill, Kerre Kammerer, and JD Shipengrover. 2011. "mapFAST: A FAST Geographic Authorities Mashup with Google Maps." *Code4Lib Journal* 14, 2011-07-25. <http://journal.code4lib.org/articles/5645>.

Connaway, Lynn Silipigni, and Ronald R. Powell. 2010. *Basic Research Methods for Librarians*. Santa Barbara, Calif: Libraries Unlimited.

Connaway, Lynn Silipigni, and Timothy J. Dickey. 2011. "Publisher Names in Bibliographic Data: An Experimental Authority File and a Prototype Application." *Library Resources & Technical Services* 55, 4. Preprint available online at: <http://www.oclc.org/research/publications/library/2011/connaway-lrts.pdf>.

Connaway, Lynn Silipigni, Timothy J. Dickey, and Marie L. Radford. 2011. "If It Is too Inconvenient I'm not Going after It: Convenience as a Critical Factor in Information-seeking Behaviors." *Library and Information Science Research* 33, 3: 179-190.

Dickey, Timothy J. 2011. "Books as Expressions of Global Cultural Diversity: Data Mining for National Collection Analysis." *Library Resources & Technical Services* 55, 3: 148-162.

Dooley, Jackie. 2011. "The OCLC Research Survey of Special Collections and Archives." *Liber Quarterly* 21, 1 (November): 125-137. <http://liber.library.uu.nl/publish/articles/000530/article.pdf>.

Faniel, Ixchel M., and Elizabeth Yakel. 2011. "Significant Properties as Contextual Metadata." *Journal of Library Metadata* 11, 3-4: 155-165. doi: 10.1080/19386389.2011.629959.

Katz, Demian, Ralph LeVan, and Ya'aqov Ziso. 2011. "Using Authority Data in VuFind." *Code4Lib Journal* 14, 2011-07-25. <http://journal.code4lib.org/articles/5354>.

Landry, Patrice, Leda Bultrini, Edward T. O'Neill, and Sandra K. Roe. 2011. *Subject Access: Preparing for the Future*. Berlin: De Gruyter Saur.

Norman, Phil, and Jeff Young. 2011. "The OpenURL Maintenance Agency: Extending and Promoting the Use of OpenURL." *Standard Spotlight. Information Standards Quarterly* 23, 1 (Winter): 26-28. <http://www.niso.org/publications/isq/2011/v23no1/norman/>. doi: 10.3789/isqv23n1.2011.07.

OASIS Search Web Services Technical Committee. 2012. *SRU/CQL Standardization in OASIS. Committee Specification Draft*. [Ralph LeVan is an editor.] <http://www.loc.gov/standards/sru/oasis/>.

Proffitt, Merrilee. 2011. "Something's Got to Give: What Can We Stop Doing in a Time of Reduced Resources?" *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 12, 2: 89-91.

Radford, Marie L., and Lynn Silipigni Connaway. 2012. "Chattin' 'bout my Generation: Comparing Virtual Reference Use of Millennials to Older Adults." In *Leading the Reference Renaissance: Today's Ideas for Tomorrow's Cutting Edge Services*, edited by Marie L. Radford, 35-46. New York: Neal-Schuman.

Radford, Marie L., Gary P. Radford, Lynn Silipigni Connaway, and Jocelyn A. DeAngelis. 2011. "On Virtual Face-work: An Ethnography of Communication Approach to a Live Chat Reference Interaction." *The Library Quarterly* 81, 4 (October): 431-453. <http://www.oclc.org/research/publications/library/2011/201109-lq.pdf>.

Radford, Marie L., Lynn Silipigni Connaway, Patrick A. Confer, Susanna Sabolcsi-Boros, and Hannah Kwon. 2011. "Are We Getting Warmer? Query Clarification in Live Chat Virtual Reference." *Reference and User Services Quarterly* 50, 3: 259-279.

Washburn, Bruce. 2011. "Library Mobile Applications: What Counts as Success?" *Information Outlook*, 15, 1 (January/February). Preprint available online at: <http://www.oclc.org/research/publications/library/2011/washburn-io.pdf>.

Research grants: supporting library/archival research and education

OCLC/ALISE Library and Information Science Research Grants

ALISE, the Association for Library and Information Science Education, is a nonprofit organization that serves as the intellectual home of university faculty in graduate programs in library and information science in North America. This grant program annually awards three grants of up to \$15,000 each to support one-year research projects by full-time faculty in LIS schools around the world. When possible, priority is given to proposals from junior faculty and applicants who have not previously received funds through this program.

Awards made in 2011—work to be done in 2012

- ▶ Abdulhussain Mahdi and Arash Joorabchi, University of Limerick (Ireland): “A New Unsupervised Approach to Automatic Topical Indexing of Scientific Documents According to Library Controlled Vocabularies”
- ▶ Carolyn Hank, McGill University (Canada), and Cassidy Sugimoto, Indiana University Bloomington: “The Biblioblogosphere: A Comparison of Communication and Preservation Perceptions and Practices between Blogging LIS Scholar-Practitioners and LIS Scholar-Researchers”
- ▶ Laura Saunders and Mary Wilkins Jordan, Simmons College: “Reference Competencies from the Practitioner’s Perspective: An International Comparison”

Projects funded in 2010—work conducted in 2011

- ▶ Bei Yu, Syracuse University: “Text Classification of Digital Reference Interviews: an Investigation of Information Seeking Behavior in the Social Web Environment”
- ▶ Chirag Shah, Rutgers University: “Modalities, Motivations, and Materials—Investigating Traditional and Social Online Q&A Services”
- ▶ Cristina Pattuelli, Pratt Institute: “FOAF in the Archive: Linking Networks of Information with Networks of People”

Making connections: Distinguished Seminar Series and other visiting speakers

- ▶ Gary Marchionini, Dean and Cary C. Boshamer Professor, School of Information and Library Science, University of North Carolina at Chapel Hill: “Libraries as Read/Write Services”
- ▶ Nancy K. Roderer, Director, Welch Medical Library, Director, Division of Health Sciences Informatics, Professor, Schools of Medicine and Public Health, Johns Hopkins University: “From Books and Buildings to Information and Services”

- ▶ Scott Walter, Associate University Librarian, Associate Dean of Libraries and Professor of Library Administration, University of Illinois at Urbana–Champaign: “The ‘Service Turn’ and the Future of the Academic Library”
- ▶ Johan Bollen, Associate Professor, Indiana University School of Informatics and Computing: “Mobilizing Usage Data: Scholarly Metrics and Recommender Systems”
- ▶ Paul Courant, University Librarian and Dean of Libraries, Arthur F. Thurnau Professor, Harold T. Shapiro Collegiate Professor of Public Policy, Professor of Economics and of Information, University of Michigan: “Economic Perspectives on Academic Libraries”
- ▶ Jane Greenberg and Todd Vision, University of North Carolina at Chapel Hill: “The Dryad Repository: A New Path for Data Publication in Scholarly Communication”
- ▶ Joy Palmer, Senior Manager for Library and Archival Services, Mimas (a UK education service provider based at the University of Manchester): “COPAC: Reengineering the UK National Academic Union Catalogue to Serve the 21st Century Researcher”



Webinars

We held 15 webinars in 2011 that together totaled 893 minutes and were attended by 1,317 OCLC Research Library Partners and members. Recordings are openly available on the OCLC Research website, from <http://www.oclc.org/research/>, for online viewing at any time.

The top three webinars by attendance were:

- ▶ TAI CHI HTML5 & CSS3 Webinar (304 attendees)
- ▶ Scan and Deliver: Creative User-initiated Digitization in Special Collections and Archives (212 attendees)
- ▶ ArchiveGrid (178 attendees)

For further information about our current work, visit the OCLC Research website: www.oclc.org/research.



4. OCLC Research people

Rick Bennett
Consulting Software Engineer

Robert Bolander
Senior Communications Officer

Jeremy S. Browning
Senior Software Engineer

Eric R. Childress
Consulting Project Manager

Patrick Confer
Web Content Coordinator

Lynn Silipigni Connaway, Ph.D.
Senior Research Scientist

Alyssa Darden
Research Assistant

Lorcan Dempsey
Vice President, Research, and Chief Strategist

Karen Disbrow
Administrative Coordinator

Jackie Dooley
Program Officer

Nancy Elkington
Director, Partner Relations

Ricky Erway
Senior Program Officer

Ixchel Faniel, Ph.D.
Post-Doctoral Researcher

Jean Godby, Ph.D.
Senior Research Scientist

Carol Hickey
Research Support Specialist

Thomas B. Hickey, Ph.D.
Chief Scientist

Erin Hood
Research Assistant

Andrew Houghton
Consulting Software Engineer

Tip House
Chief Architect, Global Infrastructure

Kerre A. Kammerer
Consulting Software Engineer

Rob Koopman
Senior Developer / Team Leader

Susanne Krouse
Executive Assistant to the Vice President

Brian Lavoie, Ph. D.
Research Scientist

Ralph LeVan
Senior Research Scientist

Daniel Londono
Diversity Fellow

Constance Malpas
Program Officer

Janet Mason
Administrative Coordinator

Dennis Massie
Program Officer

Jeanette McNicol
Executive Assistant

Jim Michalko
Vice President, OCLC Research Library
Partnership

Edward T. O'Neill, Ph.D.
Senior Research Scientist

Merrilee Proffitt
Senior Program Officer

Melissa Renspie
Senior Communications Officer

Jennifer Schaffner
Program Officer

JD Shipengrover
Senior Web and User Interface Designer

Devon Smith
Consulting Software Engineer

Karen Smith-Yoshimura
Program Officer

Mike Teets
Vice President, Innovation

Roy Tennant
Senior Program Officer

Roger H. Thompson, Ph.D.
Consulting Software Engineer

Jenny Toves
Software Architect

Titia van der Werf
Senior Program Officer

Diane Vizine-Goetz, Ph.D.
Senior Research Scientist

Harry R. Wagner
Consulting Software Engineer

Bruce Washburn
Consulting Software Engineer

Jeff Young
Software Architect

With interest I read the OCLC paper and watched your webinar about scan and deliver. Great work and I really like your approach!

Marc Holtman
Project Leader Digitization, Amsterdam City Archives



OCLC Research
6565 Kilgour Place
Dublin, Ohio 43017-3395 USA