

## Jan/Feb

### Announcements

Upcoming Minitex/OCLC  
Training ..... 1

### General

ALA Midwinter Summary..... 2

### Cataloging & Metadata

ALCTS Continuing Education  
Opportunities..... 6

OCLC Research Webinar:  
Libraries as Read/Write  
Services..... 6

Minitex RDA Mailing List..... 7

Genre/Form Webinar..... 7

RDA Training..... 7

Connexion Client Update..... 8

### Resource Sharing

Article Exchange Integrated  
Into WorldCat Resource  
Sharing..... 9

### Digitization & Preservation

ALA Midwinter CONTENTdm  
User Group Meeting..... 9

### General

## Upcoming Minitex/OCLC Training

*Sara Ring, Minitex/BATS*

Minitex has scheduled the following **free** OCLC courses through June 2012. (**Note:** the links to registration below will take you to OCLC's registration form for each class. As of January 2012, registration for OCLC courses offered by Minitex will be through OCLC and their WebEx web conferencing software.)

If you don't see a class listed that you need, view the courses offered by OCLC and other training partners by visiting the OCLC Training Portal:

<http://training.oclc.org/training>

### Copy Cataloging with CatExpress

This online course provides instruction on copy cataloging using the CatExpress interface. This course covers the copy cataloging workflow including searching, holdings maintenance, exporting and downloading records.

**Wednesday, February 15, 11:00 AM - 12:00 PM**

<http://z.umn.edu/6g7>

**Thursday, March 8, 1:00 PM - 2:00 PM**

<http://z.umn.edu/6g8>

**Tuesday, April 17, 10:00 AM - 11:00 AM**

<http://z.umn.edu/6g9>

**Tuesday, May 15, 11:00 AM - 12:00 PM**

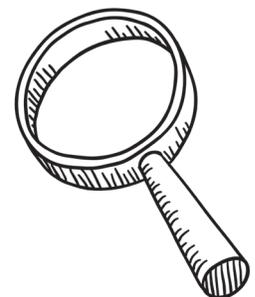
<http://z.umn.edu/6ga>

**Tuesday, June 19, 1:00 PM - 2:00 PM**

<http://z.umn.edu/6gc>

### Connexion Client Module 2: Basic Searching

This online module provides instruction on performing basic bibliographic searches using OCLC's Connexion Client cataloging software. After completing this web-based training, learners will be able to conduct basic searches for bibliographic records in WorldCat, and evaluate whether a given record matches the item to be cataloged. This is the second part of a three module series designed to enable basic copy cataloging work in Connexion Client.



**Monday, April 9, 1:00 PM - 3:00 PM**

<http://z.umn.edu/6fz>

**Wednesday, May 2, 1:00 PM - 3:00 PM**

<http://z.umn.edu/6g0>

### Connexion Browser Module 2: Basic Searching

This online module provides instruction on performing basic bibliographic searches using OCLC's Connexion Browser cataloging software. After completing this web-based training, learners will be able to conduct basic searches for bibliographic records in WorldCat and evaluate whether a given record matches the item to be cataloged. This is the second part of a three module series designed to enable basic copy cataloging work in Connexion Browser.

**Tuesday, March 13, 1:00 PM - 3:00 PM**

<http://z.umn.edu/6g4>

**Wednesday, April 11, 1:00 PM - 3:00 PM**

<http://z.umn.edu/6g5>

**Tuesday, May 15, 1:00 PM - 3:00 PM**

<http://z.umn.edu/6g6>

### CONTENTdm Basic Skills 1: Getting Started

This course provides instruction on the structure of CONTENTdm, navigating and using CONTENTdm Collection Administration to create and manage a digital collection, as well as using the CONTENTdm Project Client to prepare digital items for addition to a collection. After completing this web-based training, learners will have the foundational knowledge required to begin using CONTENTdm to manage digital items. With this basic understanding of the software, learners will be prepared for the next module, "Working with Text in CONTENTdm." Learners will be able to navigate CONTENTdm Collection Administration as well as the Project Client.

**Wednesday, May 30, 1:00 PM - 3:00 PM**

<http://z.umn.edu/6g1>

### CONTENTdm Basic Skills 2: Working with Text

This course provides instruction on the structure of CONTENTdm, navigating and using CONTENTdm Collection Administration to create and manage a digital collection that includes text-based materials, as well as using the CONTENTdm Project Client to prepare digital items for addition to a collection. After completing this web-based training, learners will have the foundational knowledge required to begin using CONTENTdm to manage textual materials. Following completion of this webinar, in addition to "Getting Started with CONTENTdm," learners will be prepared for the next module, "Maintaining Collections in CONTENTdm."

**Friday, June 1, 1:00 PM - 3:00 PM**

<http://z.umn.edu/6g2>

### CONTENTdm Basic Skills 3: Maintaining Collections

This is the third webinar in the basic series for users new to CONTENTdm. This course provides instruction on editing collection items through CONTENTdm administration and the CONTENTdm Project Client. Additionally, the course explores workflows that make it possible through the use of tab-delimited data to import significant numbers of compound objects and/or legacy metadata in one operation. After completing this web-based training, learners will have the foundational knowledge required to begin using CONTENTdm to manage digital collections.

**Monday, June 4, 1:00 PM - 3:00 PM**

<http://z.umn.edu/6g3> ■

General

---

## ALA Midwinter Summary

*Sara Ring, Minitex/BATS*

I attended a number of thought-provoking sessions at the midwinter meeting of the American Library Association in Dallas, TX, Jan. 20-24. One topic discussed (about which we will hear much more in the coming years) was the concept of the Semantic Web and library linked data. I'll define these concepts first, and, then, follow with a summary of what I learned at a preconference I attended on the topic. If you are already familiar with these concepts, skip ahead to the preconference summary.

Though you will find various definitions depending on where you look, I will quote a definition by the Worldwide Web Consortium (W3C), as the Semantic Web effort has been led by them: "The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries..." The web, as we know it today, concentrates on the sharing and retrieval of documents. The Semantic Web is about using common formats (structured data) to be able to bring together data from all sorts of sources on the web. The W3C gives this example on their website:

*There is lots of data we all use every day, and it is not part of the web. I can see my bank statements on the web, and my photographs, and I can see my*

*appointments in a calendar. But can I see my photos in a calendar to see what I was doing when I took them? Can I see bank statement lines in a calendar?*

We cannot, because data is controlled by applications, and that data is not open.

The Semantic Web is based on the Resource Description Framework (RDF). RDF is a model for the interchange of data on the Web. RDF uses Uniform Resource Identifiers (URIs) for naming relationships between things. Think of RDF as a way to structure metadata as statements (not records), composed of a Subject (resource) Predicate (property) and Object (value). Illustration of this concept will come up later in this article.

I need to define one more term that will be used heavily in the summary article below: linked data. If RDF is the data model for the Semantic Web, linked data is a “method of publishing structured data so that it can be interlinked and become more useful. It builds upon standard Web technologies such as HTTP and URIs, but rather than using them to serve web pages for human readers, it extends them to share information in a way that can be read automatically by computers. This enables data from different sources to be connected and queried.”<sup>1</sup> While the term, linked data, was coined by Tim Berners Lee, the concept of publishing structured data and linking it is very old. In libraries, for example, think of the way we use controlled headings in library catalogs.

One question you may ask up front: “Why should I care about the Semantic Web and library linked data?” The short answer is: If libraries (and library data) are to be of value to our users now and into the future, we should start paying attention to the Semantic Web. I intend to continue the discussion about library linked data with Minitex participating libraries, so you will be hearing more on this topic in the future. I hope the summary below is a starting point for your own exploration.

## Libraries, Linked Data and the Semantic Web: Positioning Our Catalogs to Participate in the 21st Century Global Information Marketplace

The presentations from each speaker can be found here: <http://alamw12.scheduler.ala.org/sessions/handouts>

### Speaker 1: Eric Miller, Zepheira

*Libraries in the Web: Weaving a Web of Data*

Eric Miller began his talk with a screen capture of a document by Tim Berners Lee (founder of the World Wide Web). The 1989 document was titled *Information*

<sup>1</sup> See Wikipedia: [http://en.wikipedia.org/wiki/Linked\\_data](http://en.wikipedia.org/wiki/Linked_data)

*Management: A Proposal* (<http://info.cern.ch/Proposal.html>). The document was an attempt to show a global hypertext system...later to be called the World Wide Web. There was a handwritten note from his boss at the top of the document that said “vague but exciting.” At that time the World Wide Web was just a concept, only later to be realized. I think the point was that the Semantic Web is in its infancy, the potential not fully realized.

Miller showed an example of what humans see when they view a resource on a webpage, vs. what browsers see when they view a webpage. We (humans) would see, for example, a Title, Author, Publication Date, and article content. A browser would read it as Headline, Subheadline, Italics, and text. Computers today blindly retrieve information that we request, and they don’t understand the meaning behind the information they are showing us.

The more organizations that provide their data as linked data, the more valuable their data becomes. The web perspective is not to lump data into one place, but, instead, connect to each others’ linked data sets.

To show the power of human computing, he spoke about CAPTCHA (<http://en.wikipedia.org/wiki/CAPTCHA>). You see them everywhere on the web, at least, anytime you have to prove you’re a human and not a machine. There are 20 million people filling them in every 10 seconds. The words come from scanned books, and by typing them in, we are all helping to digitize old books.

Miller spoke about the power of web identifiers as well. Creative Commons is a good example. People identify their objects with a Creative Commons license. Creative Commons didn’t have to build a whole database to say: Show me all the things that have this particular license. Because people are putting specific identifiers on their objects, one can search across the web for any resource with a Creative Commons license.

To fully realize the Semantic Web, it will take big application changes. Libraries are not /N the web. For example, the Library of Congress NDIIPP (National Digital Information Infrastructure Program) has 180 partners. These institutions could all benefit from sharing data, but it exists in silos (mentioned the new resource Viewshare as a way of connecting data: <http://viewshare.org>), rather than linked together.

There are a few projects going on that Miller referred to as “Linked Data Accelerants.” There is increased industry and organization adoption (Microsoft, Google, Yahoo!, Schema.org) of linked data as Google and Yahoo! announced they will index RDFa. **Note:** RDFa is a kind of ‘lite’ version of RDF, which adds metadata to HTML (or XHTML)

webpages. The Library of Congress has published some of their authorities and vocabularies as linked data already (such as LC Subject Headings and the LC Name Authority File). See: <http://id.loc.gov>

The EU Commission has launched an Open Data Strategy for Europe, and they expect this to result in a 40 billion euro boost to the economy. <sup>2</sup>

There is increased understanding of the return on investment (ROI) and benefits of using linked data. Miller gave the example of Best Buy. Since Best Buy moved to linked data, they have seen a 30% increase in search traffic to their website. Data, such as store name, address, store hours and GEO data, is being marked up using RDFa, so search engines are now able to identify each of the data components more easily and put them into context.

Despite these linked data accelerants, Miller noted that linked data activity is happening mostly outside of library community. Below are a few examples.

- **Viewshare** (<http://viewshare.org>) is a free platform for generating and customizing views (interactive maps, timelines, facets, tag clouds) that allow users to experience your digital collections. Viewshare is built on a platform called Recollection, which is open source and uses RDF.
- **Exhibit 3.0** (<http://www.simile-widgets.org/exhibit>) is an Open Source tool from MIT's Simile that allows you to easily create web pages to visualize large sets of data. There are some neat examples on the website linked to above.
- **Google Refine** (<http://code.google.com/p/google-refine>) is a power tool for working with messy data, cleaning it up, transforming it from one format into another, extending it with web services, and linking it to repositories like Freebase (A repository of Creative Commons licensed structured data: [http://wiki.freebase.com/wiki/What\\_is\\_Freebase%3F](http://wiki.freebase.com/wiki/What_is_Freebase%3F)).

## Speaker 2: Ross Singer, Talis

*Thinking Beyond our Collections: Making our Models Linked and Linkable*

Ross Singer began his talk by explaining the concept of linked data. It is using URIs (uniform resource identifiers) as names for things, and using HTTP URIs so that people can look up those names (URLs). When someone looks up a URI, it provides useful information using standards (RDF) and links to other URIs to discover even more things.

What is RDF? RDF stands for Resource Description

<sup>2</sup> European Commission press release: <http://z.umn.edu/6hr>

Framework. It is a data model that uses statements (triples) to describe data. The first part of a triple is the subject, followed by the predicate, followed by the object. Let's use the example of a resource called *Weaving the Web*. This is a book by Tim Berners-Lee that has an ISBN of 0062515861. If we represent some of this information as a triple, one statement looks like this:

Weaving the Web (*Subject*)  
has author/creator (*Predicate*)  
Tim Berners Lee (*Object*)

We can then use URIs (URLs) in place of the names listed above. Here's one example. Note that the first URI is completely made up, just to show that *Weaving the Web* exists somewhere on the web.

<http://locationofresource.org>  
This is a made up URI for *Weaving the Web*

<http://purl.org/dc/terms/creator>  
URI for the Dublin Core *Creator* term

<http://viaf.org/85312226>  
URI for Tim Berners Lee in the Virtual International Authority File

In addition to following the RDF data model, the data still has to be expressed/published in some way so that it can be shared and read by other machines on the web. There are a few different ways RDF is expressed: as XML (RDF/XML), Notation 3 (or N3), Turtle and N-Triples.

Why use RDF to make our library collections linked and linkable? RDF is versatile, and in particular:

- schema-less
- properties can be assigned from a number of vocabularies
- description can be both generalized as well as domain or audience-specific

RDF does bring some challenges. For example, logic prevails (no schema to validate your data against). There is no provenance of data in RDF (no clear way to address conflicting data).

He mentioned a few other datasets to consider modeling around:

- **DBpedia** (everyone builds around this, and it is the hub of the Semantic Web). However, the data is messy. He showed an example from DBpedia of the page for a Sonata (<http://dbpedia.org/page/Sonata>), which includes a reference to a Flickr page of photographs of

- a heavy metal music performance.
- [GeoNames.org](#) (geographic and administrative data) contains more than 8 million resources.
- [Musicbrainz.org](#) is an open music encyclopedia
- [Openlibrary.org](#) (one webpage for every book) has works, editions, authors, and subjects
- [Schema.org](#): Google, Bing, and Yahoo! have agreed to index websites that mark up their pages using their schemas. See the [Schema.org book \(http://schema.org/Book\)](#) and article ([http://schema.org/Article](#)) descriptions.

Singer ended his talk by saying that linked data gives us the potential to integrate into the larger web. Reuse of our library data increases our relevance, and we can reuse other's data.

### Speaker 3: Karen Coyle, Consultant *Are We There Yet?*

Karen Coyle first talked about library data, showing a screen shot of a catalog card contrasted with a screen shot of the same data presented in an online catalog. Her point was the data doesn't look that different. So, how do we move from where we are now toward linked data? Linked data is:

- data, not text
- identifiers for things
- statements, not records
- machine-readable schema
- machine-readable lists
- open access on the web

There is a lot of text in our MARC records. For example, if you look at the ISBN MARC field (020), you will find a number, but often you will see text like "(hardcover)" after an ISBN. Or, you may see a description MARC field (300) that states the size of the item as "23 cm." This is text, not data that can easily be read by machines.

Coyle presented examples of data (not text) in library MARC records. Real "data" is in the fixed field of a MARC record. MARC records are fragile. If we remove a 100 field from a record, that information doesn't make sense on its own (another reason to use statements to describe items rather than records).

What are the things in our library metadata?

- people, corporate bodies, families
- places (subjects, locations)
- events (conferences)

- topics (classification, subject headings)
- resources (books, sound recordings)
- physical formats

How many things listed above are unique to library data? The answer is none of them. I think her point was that there are communities outside the library world with which we could be working that are also describing people, places, and things.

She gave some examples of controlled vocabularies and authorities used in libraries that are presented as linked data and that are linking to other sources, such as:

- Virtual International Authority File (VIAF) links to Wikipedia (in Wikipedia, one can even enter in a VIAF number).
- MARC Genre/form headings, Relators, Countries, Geographic Areas, Languages. The Library of Congress has exposed these as linked data and provided open access to them.

### Speaker 4: Corey Harper, New York University Libraries *Of Cataloging & Context: Metadata and Metadata Experts on the Linked Data Web*

Who is producing library linked data? CKAN (open-source data portal) lists 52 Library Data Sets (such as the British National Bibliography and Cambridge University) at: <http://ckan.net/group/lld>

Harper also discussed Europeana (15 million items from 1500 institutions). They are aggregating cultural heritage objects from throughout Europe, creating linked data for institutions, and creating virtual exhibitions. More information and this project can be found at:

<http://www.linkedheritage.org>

Who is consuming library linked data? Resources mentioned included:

- [Thinkbase](http://thinkbase.cs.auckland.ac.nz/start.jsp) (<http://thinkbase.cs.auckland.ac.nz/start.jsp>) Thinkbase uses a visualization tool (Thinkmap) to create an interactive visual representation of the contents and semantic relationships in Freebase. Freebase is an open, shared database of the world's knowledge.
- [Pelorus](http://datagov.clarkparsia.com) (<http://datagov.clarkparsia.com>) An application to enable faceted browsing of the data.gov catalog. This is possible because data.gov provided their data as linked data. This application also relates government datasets to books that are about the same subject matter (via Library of Congress Subject Headings).

Other interesting linked data tools mentioned:

- **Free Your Metadata** ([freeyourmetadata.org](http://freeyourmetadata.org))  
Use Google Refine to clean-up, polish, and publish your metadata as linked data.

**Speaker 5: Peter Brantley, Director, BookServer Project, Internet Archive**

*Breaking the Catalog: Navigating Books on Shelves*

Peter Brantley stated that bibliographic data does not give you a sense of the “awesomeness” of the materials. All bibliographic data underperforms in this way, no matter how we describe it. Discovery is metadata, contextualized by user desire. One of linked data’s challenges is to contribute to discovery.

Consider the website Small Demons ([smalldemons.com](http://smalldemons.com)). It can give you a sense of the “awesomeness” of the books mentioned. Small Demons states on their homepage that it is “...the people, places and things from books, and everywhere they can take you.”

Linked data will not solve all our problems, but it can help. For example, if I want to find out what to watch or read, I would want to go to the largest aggregation. He thinks that library linked data has to be at network scale, on a single platform, to be end-user attractive. ■

Cataloging & Metadata

---

## ALCTS Continuing Education Opportunities

*Sara Ring, Minitex/BATS*

The Association for Library Collections & Technical Services recently scheduled new webinars, listed below. And, don’t forget about the next free e-Forum, *Transforming Collections* (Feb. 22-23), that will focus on all the ways our library collections are changing and will discuss topics such as handling new formats, preservation methods, repository services, planning for the future, best practices for moving forward, and budgeting for changing times. More information about this e-Forum can be found here:

<http://www.ala.org/ala/mgrps/divs/alcts/confevents/upcoming/e-forum/022212.cfm>

## (Webinar) Libraries & MARC Holdings: From Works to Items

*Wednesday, March 21, 1:00 PM (Central)*

<http://z.umn.edu/6gu>

This session provides a brief introduction to and overview of the MARC21 Format for Holdings Data (MFHD) and ANSI/NISO Z39.71. Discussion topics include the importance of standardized holdings for such mission-critical library efforts as Item-Level Inventory Control, InterLibrary Loan (ILL) & Resource Sharing, and OffSite Storage.

**Who Should Attend:** Librarians and library staff who create or maintain holdings information. Introductory session provides good background for anyone interested in the two-day SCCTP Serials Holdings Workshop that provides a rich knowledge of the intricacies of coding MARC holdings.

## (Webinar) Steps in a Digital Preservation Workflow

*March 7, 1:00 PM (Central)*

<http://z.umn.edu/6gt>

This presentation will outline generic considerations and processes for building and managing a digital preservation workflow. We will consider the workflow within the larger context of a digital content life cycle, which runs from information creation through to ongoing discovery and access. We will focus upon generalized steps that institutions can use to acquire, preserve and serve content. The presentation will describe distinct workflow stages in conjunction with sample procedures, policies, tools and services, stressing the dynamic nature of workflows over time, including the use of modular components and ongoing work to enhance automation and cope with issues of scale.

**Who should attend:** Librarians and archivists involved with preservation of digital materials. ■

General

---

## OCLC Research Webinar: Libraries as Read/Write Services

*OCLC; edited*

OCLC Research hosts a Distinguished Seminar Series Presentation: *Libraries as Read/Write Services* with Gary Marchionini, Ph.D., University of North Carolina at Chapel

Hill. This free online presentation will take place on Tuesday, February 21, 8:00 AM (Central)

The presenter will discuss two trends that are driving foundational shifts in libraries of the 21st century. First, the development of digital libraries has expanded the nature of service communities beyond local geography and influences the need to collect the works of global interest. This, in turn, is causing libraries to distinguish themselves by focusing their digital collections on local materials that are unique, becoming more like archives that curate singular materials. Second, people increasingly work, play, and live with the aid of electronic technologies that create traces of life's behaviors. People consciously and unconsciously create documents, photos, and files as well as streams of interactions with other people and with myriad electronic systems. Individual traces of our lives in cyberspace in aggregate represent what Gary Marchionini terms "profections" of our personal identities.

Libraries have trusted roles to play in helping people manage their personal digital libraries—to provide "write" as well as "read" services to their patrons. Inevitably, some of the personal assets will become part of the public collection with patron permission. Libraries must develop strategies to support deposit (writing) of patron assets and to manage the blurring of boundaries between personal and public digital libraries. The School of Information and Library Science at UNC's LifeTime Library Project aims to develop and teach such strategies. Register at the following url:

<http://z.umn.edu/6gf> ■

## Cataloging & Metadata

---

### Minitex RDA Mailing List

*Mark K. Ehlert, Minitex/BATS*

The Minitex Bibliographic and Technical Services (BATS) unit is happy to introduce a new e-mail discussion list dedicated to conversation and announcements on Resource Description and Access (RDA), the new set of cataloging guidelines likely to succeed AACR2 in early 2013. This list, dubbed mt-x-rda, is open to all library staff and other interested parties in the Minitex region (including Wisconsin). Those outside the region may also participate.

To subscribe, point your browser to the following website and fill out the form in the middle of the page:

<http://lists.minitex.umn.edu/mailman/listinfo/mt-x-rda>

On the same page you will find a link to the archives as well as a form near the bottom that will give you access to password, subscription, and message delivery options. The intent behind the mt-x-rda list is to provide a regional forum for subscribers to discuss both practical and theoretical matters concerning RDA: rule interpretations, workflow changes, record indexing and display. In the spirit of Minitex's collaborative history, we look forward to an active and engaged community assisting one another with learning and applying this new cataloging code. ■

### Genre/Form Webinar

We still have open seats for Janis Young's leap day webinar on the LC genre/form headings.

#### Defining the Genre; What's the Form?: LC Genre/Form Terms for Library and Archival Materials

*February 29, 9:00 AM - 11:00 AM*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=367>

Janis L. Young, Senior Cataloging Policy Specialist at the Library of Congress, will introduce participants to the Library of Congress Genre/Form Terms for Library and Archival Materials (LCGFT), a new thesaurus of genre/form vocabulary that catalogers have recently begun applying to special library formats such as videos, sound recordings, and maps. Through lecture and exercises, Janis will define LCGFT core principles and methods of application, as well as show how these terms complement subject headings to provide catalog users a fuller understanding of the library materials that interest them. ■

## Cataloging & Metadata

---

### RDA Training

*Mark K. Ehlert*

We still have plenty of seating available for our February and March RDA training sessions, which are listed below. If you cannot make it to these webinars, rest assured we will be presenting them again throughout the year. All times listed below are Central.

## RDA for the Solo/Copy Cataloger

March 5, 7, 9, 9:30 AM - 12:00 PM

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=351>

With RDA implementation likely in early 2013, now is a good time for catalogers to become familiar with recognizing, reviewing, and creating records based on the new cataloging standard. Directed to catalogers in small libraries and those copy catalogers with basic experience, this 3-part, online workshop describes the fundamentals of RDA cataloging for most formats. Though emphasis will be placed on bibliographic records, there will also be a brief look at what's new for RDA authority records.

## RDA for the Seasoned Copy Cataloger: Audiovisual Materials

March 20, 22, 26, 28, 9:30 AM - 12:00 PM

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=330>

Experienced copy catalogers of audiovisual materials who have an interest in learning more about RDA are welcome to participate in this 4-part, online workshop. This webinar series will focus on the descriptive portions of RDA-MARC bibliographic records, with a brief tour of changes made to MARC authority records. ■

## Cataloging & Metadata

### Connexion Client Update

OCLC; edited

Connexion client 2.40 is expected to be released in the March/April 2012 timeframe. Below is a summary of the planned enhancements. This information is also posted on the OCLC website:

<http://www.oclc.org/connexion/interface/client/enhancements/future.htm>

### Changes to Controlling Headings

(Expected: February 2012) The following applies to all currently supported versions of Connexion client (versions 2.10, 2.20, and 2.30).

For Edit --> Control Headings --> All, the system:

- Uncontrols invalid controlled headings (descriptive name headings and/or subject headings).
- Automatically checks for an updated version of the heading text and updates the text if needed.
- Automatically updates a previously fully or partially

controlled heading with punctuation-only changes.

- Automatically re-evaluates a partially controlled heading and attempts to fully control it

### Connexion Client Version 2.40

(Expected: April/May 2012) Version 2.40 will include the following enhancements:

#### GLIMIR

Set an option to show search results in GLIMIR (Global Library Manifestation Identifier) clusters of WorldCat bibliographic records that have different languages of cataloging for the same work (called "parallel" records). For each cluster, GLIMIR search results show the number of records that your library holds, the total number of holdings, and the total number of records. GLIMIR clusters make it easier to identify and select the exact record you need for cataloging. In the client only, you can also show the GLIMIR cluster for any displayed bibliographic record by using a new menu item Cataloging --> Show --> All GLIMIR Cluster Records. For more information on GLIMIR see:

<http://www.oclc.org/reports/worldcatquality/default.htm>

#### Classify

Use a new command Classify on the Tools --> Options menu to automatically supply a Dewey or Library of Congress classification number in a displayed record for a book, DVD, CD, or other material type in a field you specify. The new command uses the FRBR-based Classify Web service to provide a classification number from a WorldCat record with a matching standard number in this order: OCLC number, ISBN, ISSN, UPC, LCCN. For more information on Classify see:

<http://classify.oclc.org/classify2>

#### RDA Workforms

Set an option in the client using Tools --> Options --> RDA tab (formerly the RDA Toolkit tab) to use RDA versions of the existing AACR2 workforms to create records. Set the option separately for bibliographic and/or authority workforms. Existing workforms open by default when you create new records unless you set the RDA workform option(s).

#### 264 Field in WorldCat Search Results

Record descriptions in a brief or truncated list of WorldCat search results include data from the 264 field, if present in the record and if no field 260, 261, or 262 is present. Field 264 contains RDA production, publication, distribution, and manufacture information.

## MARC Update 2011

New indexes are added to the drop-down index lists in search and/or browse windows, as appropriate. For details, see *Technical Bulletin 260 OCLC-MARC Format Update 2011*:

<http://www.oclc.org/support/documentation/worldcat/tb/260/default.htm>

Language names are updated in:

- The “Language” and “Language of Cataloging” drop-down lists in the Search WorldCat window.
- The “Language” drop-down list in the Search Online Heritage Printed Book (HPB) Database (if the HPB menu is activated in Tools --> Options --> Heritage Printed Book tab).

## Macro Changes

A new macro command *GetClassNumber* is available for the new Edit --> Classify functionality mentioned previously. Two new OCLC-supplied macros are available in the OCLC.mbk macro book:

- ClearELv13Workform—This macro clears candidate fields from an encoding level 3 bibliographic record and replaces them with workform prompts. The prompts are a blend of two sets of workforms—AACR2 versus RDA.
- GenerateERRecord—This macro creates a workform to represent an online version of an item based on the print, microform, or another format version of a bibliographic record for the same title. (Note: The GenerateERRecord macro replaces the two OCLC macros: GenerateESerial and GenerateESerialCONSER). ■

## Resource Sharing

### Article Exchange Integrated Into WorldCat Resource Sharing

*OCLC; edited*

Following a successful five-month trial, the Article Exchange feature is now available within the WorldCat Resource Sharing service and, where available, through an ILLiad add-on. This new feature provides a single, secure location where lending libraries can place and borrowing libraries and their users can retrieve requested articles obtained through interlibrary loan. Lending libraries scan documents and upload them to the site and notify the borrowing libraries of the documents' availability. The site generates a unique URL and password for each

document. An e-mail button enables the lender to insert the borrower's e-mail address and send the link directly to the library. Visit the full news release here:

<http://z.umn.edu/6go> ■

## Digitization & Metadata

### ALA Midwinter CONTENTdm User Group Meeting

*Sara Ring, Minitex/BATS*

**Part One: CONTENTdm Update** by Geri Bunker Ingram, OCLC Digital Collection Services

Geri Bunker Ingram highlighted Denver Public Library (<http://digital.denverlibrary.org/cdm>), which has done quite a lot of customization with their CONTENTdm 6.1 site. They have created an iPhone mobile app with information and photographs of Denver's historic buildings. With it, users can explore 7 historic neighborhoods and add comments, upload photos, and stories about the places and events featured in these neighborhoods. The iPhone app is eventually going up in the Applications part of OCLC WorldShare (<http://www.oclc.org/developer/applications>). In the meantime, I did find more information about this project with links to the iPhone app here:

<http://creatingcommunities.denverlibrary.org>

CONTENTdm 6.2 will be released in Summer 2012. My Favorites and optimization for Google will be among the things included in this release.

A new CONTENTdm Product Manager started work in January, Christian Sarason (has background in management of big data in the oceanography field).

*There are changes to CONTENTdm User Groups. Minitex has agreed to host a national virtual meeting in later fall, 2012. Our group was formerly called the Upper Midwest CONTENTdm User Group. Now we will be known as the Northern group.*

See the CONTENTdm Blog announcement for more information about these changes (Note: you have to have access to the CONTENTdm User Support Center):

<http://www.contentdm.org/USC/blog/blogs/blog1.php/2012/01/19/contentdm-user-group-meetings-in-2012>

**Part Two: SEO for Digital Repositories** by Kenning Arlitsch (University of Utah)

Arlitsch and his team have spent about 22 months of

research on Search Engine Optimization for the University of Utah's digital collections. That includes a large CONTENTdm collection of around 1.5 million items (they use CONTENTdm for their institutional repository as well). When they analyzed their digital collections, they found Google index ratios as low as 2% and a very poor showing in Google Scholar (to determine the index ratio, they divided the number of URLs submitted to Google by the number of actual URLs indexed by Google).

After making numerous changes, they found they increased their Google index ratio across the board from July 2010 to October 2011 (ending with an 87% Google Index Ratio on average). In their literature review, they found most SEO studies to be dated, and most focused on general websites. Few dealt with digital collections in databases.

They used Google's webmaster tools to submit a Sitemap to Google (an XML file that contains a url for every digital object in the collection). This increased Google referrals by 200% and total visitors by 79%. Why spend time on SEO? He cited OCLC's Perceptions of Libraries, 2010 report to show that 83% of college students begin their research via a Search engine. Google is at the top with roughly 65 % of the traffic compared to other search engines.

Know your stakeholders and what they value. For example, Faculty value publication page views, downloads, and citations. Donors/archivists value digital collection pages indexed, page view, visitors, and reproductions ordered.

Educate your staff on what the Search Engines value and their policies and practices. Search Engine rules and enforcement levels change. They are insensitive to standards valued by librarians. For example, in the Google Scholar index, they advise you to "Use Dublin Core tags (e.g., DC.Title) as a last resort." Google Scholar wants Highwire Press, PRISM, BePress, and Eprints metadata schema. They do not want Dublin Core because it doesn't have adequate fields to support citation metadata. For example, citation information is usually placed all in one lump of text in a dc.relation or dc.source field, and there is not consistency in how it is formatted in these fields. It makes it difficult for Google Scholar to parse and index the data into their own bibliographic components (journal name, volume, issues, page numbers, etc).

Set up Google webmaster tools and ask questions. For example, they found a lot of crawl errors in their robots.txt file. Two different staff set up their Sitemap and robots.txt file and they weren't talking to each other. So, their Sitemap was basically saying to Google "come on in," and the robots.txt file was saying "stay out." The rest of his presentation focused primarily on Google Scholar. This

research was published as an article in the publication *Library Hi Tech*, vol. 30, issue 1.

Google Scholar has a completely separate index from the main Google index, and different rules. First they asked themselves, does Google Scholar really matter? They found that lots of people use Google Scholar through their literature review. Part of their methodology included surveying other institutional repositories to find their GS index ratio. They used the command: site: "insert your site url here" However, this only searches the primary versions of academic papers. A paper formally published in a journal would be considered the primary version. Other versions (like preprints) might be included in the index, but are considered other versions (you have to click through to view the "other versions"). In the second survey they searched Google Scholar for known repository items title by title. This proved more accurate.

Google Scholar wants the right metadata tags used consistently and accurately. Looking at a mix of different repositories across the U.S, they found on average only 30% of the content appeared in Google Scholar! In the study, there were higher index ratios for those institutions that followed the crawling guidelines.

Their challenge was to present bibliographic citations in a way that Google Scholar can identify, parse, and digest. They began to align their Dublin Core fields with Highwire Press and parse each citation into HTML meta tags Google Scholar can read.

### Pilot 1

6,482 URLs in Sitemaps were submitted via Google Webmaster Tools. Any errors generated during Google crawls were analyzed and addressed. They updated and corrected metadata for 20 pilot articles. Their Google index ratio increased, but their Google Scholar index ratio remained at 0%.

### Pilot 2

Utilized OCLC's relationship with Google Scholar (they were able to talk to Google Scholar engineers on the phone). They brought their Google Scholar index ratio up to 62%, then the server crashed.

### Pilot 3

They increased sample size to 56 articles. In this pilot, they brought their Google Scholar Index Ratio up to 90%, with the server up 100% of the time.

They received an IMLS Grant to support their SEO research for the next 3 years. They will be working with Jean Godby from OCLC Research on the project. OCLC will

develop formal crosswalks between Dublin Core and one or more of the publishing industry schemas recommended by Google Scholar. With this grant they hope to expand their research, publish a toolkit (that includes SEO recommendations, metadata transformation mechanisms, tools for monitoring and recording), and to disseminate their findings through conferences, webinars, and papers.

Arlitsch ended his talk with a few pointers on optimizing digital collections for search engines. If you have a digital collection, you should:

- Establish baseline data (Configure Google Analytics, Set up Webmaster Tools)
- Submit Sitemaps and configure a robots.txt file
- Monitor and address errors
- Inform staff and assign ownership
- Find out what your staff know
- Clean up your metadata
- Upgrade repository software ■

## Minitex/OCLC Mailing Contact Information

Minitex Bibliographic and Technical Services (BATS)  
612-624-4002, 800-462-5348, [mino@umn.edu](mailto:mino@umn.edu)  
Sara Ring, Carla Dewey Urban, Mark Ehlert  
*OCLC Cataloging, ILL, Digitization and Preservation products and services.*

Minitex Contract Cataloging Program (ConCats)  
612-624-4002, 800-462-5348, [ConCats@umn.edu](mailto:ConCats@umn.edu)  
Mark Ehlert

Minitex Cooperative Purchasing & Electronic Resources Services (CPERS)  
Rita Baladad 612-626-8252, [balad001@umn.edu](mailto:balad001@umn.edu)  
*OCLC Reference products and services*

The *Minitex/OCLC Mailing* is an informational bulletin sent monthly to Minitex/OCLC libraries. Permission to reprint with appropriate acknowledgement is granted. All articles should be attributed to Minitex unless otherwise credited. This publication is available in alternate formats upon request. Please call Kay Kirscht, Minitex 612-624-4002 for further information.

Minitex is a publicly supported network of academic, public, state government, and special libraries working cooperatively to provide and improve library service to patrons in Minnesota, North Dakota, and South Dakota.

For address and name changes, please send a message to [mino@umn.edu](mailto:mino@umn.edu).

The University of Minnesota is an equal opportunity educator and employer.

♻️ Printed on recycled and recyclable paper with at least 30 percent postconsumer material.

**Minitex**  
DEDICATION. EXPLORATION. INNOVATION.

*An Information and Resource  
Sharing Program of the  
Minnesota Office of Higher  
Education and the University  
of Minnesota Libraries*

**Minitex  
University of Minnesota  
15 Andersen Library  
222 21st Avenue South  
Minneapolis, MN 55455-0439**

**Minitex**  
DEDICATION. EXPLORATION. INNOVATION.

**JANUARY/FEBRUARY 2012**

**MINITEX/OCLC MAILING**

A Publication of the Minitex Bibliographic and Technical Services Unit

# MINITEX CALENDAR

This calendar primarily lists events scheduled by Minitex, although other events are included. This is an informational posting only, registration materials are sent

separately. If you would like your event included in the calendar, please call Kay Kirscht at 612-624-3532.

## FEBRUARY

---

15

### Copy Cataloging with CatExpress

11:00 a.m. - 12:00 Noon, CST

*Online OCLC Training*

<http://z.umn.edu/6g7>

22

### Britannica's Interactive Learning Materials

11:30 a.m. - 12:00 Noon, CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=369>

23

### Using Maps in Student Resources in Context

12:30 p.m. - 1:00 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=370>

23

### ELM After School - Teaching Resources

3:30 p.m. - 4:00 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=374>

29

### "Defining the Genre; What's the Form?": LC Genre/Form Terms for Library & Archival Materials

9:00 a.m. - 11:00 a.m., CST

*Online Training Session*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=367>

## MARCH

---

2

### Creating Timelines in Britannica

10:00 a.m. - 10:30 a.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=368>

5

### RDA for the Solo / Copy Cataloger (Part 1 of 3)

9:30 a.m. - 12 Noon, CST

*Online, 3-part Training Session Series*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=351>

5

### Getting to Know the Presidential Candidates

3:00 p.m. - 4:00 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=383>

6

### Smart Resources for the Smart Board

3:30 p.m. - 4:00 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=375>

7

### RDA for the Solo / Copy Cataloger (Part 2 of 3)

9:30 a.m. - 12 Noon, CST

*Online, 3-part Training Session Series*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=351>

8

### Copy Cataloging with CatExpress

1:00 p.m. - 2:00 p.m., CST

*Online OCLC Training*

<http://z.umn.edu/6g8>

9

### RDA for the Solo / Copy Cataloger (Part 3 of 3)

9:30 a.m. - 12 Noon, CST

*Online, 3-part Training Session Series*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=351>

9

### Britannica's Interactive Learning Materials

3:00 p.m. - 3:30 p.m., CST

*Online Reference Outreach & Instruction Training session*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=369>

12

**ELM & Wikipedia**

3:30 p.m. - 4:00 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=379>

13

**Connexion Browser Module 2: Basic Searching**

1:00 p.m. - 3:00 p.m., CST

*Online OCLC Training*

<http://z.umn.edu/6g4>

14 - 15

**5th Annual Library Technology Conference 2012**

**Macalester College, Saint Paul, MN**

*Conference*

<http://www.macalester.edu/libtechconference/index.html>

16

**ELM: More Than Just Articles**

12:30 p.m. - 1:30 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=362>

20 - 28

**RDA for the Seasoned Copy Cataloger - Audiovisual Materials**

9:30 a.m. - 12 Noon CST

*Online, 4-part Training Session Series*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=330>

20

**Smart Resources for the Smart Board**

3:30 p.m. - 4:00 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=375>

21

**Making ELM Access Easy**

9:00 a.m. - 9:30 a.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=358>

27

**ELM: More Than Just Articles**

10:30 a.m. - 11:30 a.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=362>

29

**Using Maps in Student Resources in Context**

1:00 p.m. - 1:30 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=370>

**APRIL**

---

3 - 5

**FRBR & FRAD - Keys to RDA**

9:30 a.m. - 12 Noon CST

*Online, 2-part Training Session Series*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=350>

4

**Just Search - EBSCOhost**

1:00 p.m. - 1:45 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=380>

5

**ELM and Local Government**

3:30 p.m. - 4:00 p.m., CST

*Online Reference Outreach & Instruction Training session*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=377>

6

**Health and Medical Resources in ELM**

9:00 a.m. - 10:00 a.m., CST

*Online Reference Outreach & Instruction Training session*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=259>

9

**ELM and Local Government**

9:00 a.m. - 9:30 a.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=377>

9

**Connexion Client Module 2: Basic Searching**

1:00 p.m. - 3:00 p.m., CST

*Online OCLC Training*

<http://z.umn.edu/6fz>

10

**Britannica's Interactive Learning Materials**

8:30 a.m. - 9:00 a.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=369>

11

**Connexion Browser Module 2: Basic Searching**

1:00 p.m. - 3:00 p.m., CST

*Online OCLC Training*

<http://z.umn.edu/6g5>

11

**Just Search - InfoTrac**

1:00 p.m. - 1:45 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=381>

12

Accessing Your ELM Usage Statistics

11:00 a.m. – 12:00 Noon, CST

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=278>

13 - 23

RDA for the Seasoned Copy Cataloger – Music Materials

1:30 p.m. – 4 p.m. CST

*Online, 4-part Training Session Series*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=331>

17

Copy Cataloging with CatExpress

10:00 a.m. - 11:00 a.m., CST

*Online OCLC Training*

<http://z.umn.edu/6g9>

18

Just Search - ProQuest

1:00 p.m. – 1:45 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=382>

20

Accessing Your ELM Usage Statistics

12:30 p.m. – 1:30 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=278>

25

Making ELM Access Easy

3:30 p.m. – 4:00 p.m., CST

*Online Reference Outreach & Instruction Training*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=358>

25

RDA for the Seasoned Copy Cataloger – Print Monographs (Part 1 of 4)

9:30 a.m. – 12 Noon CST

*Online, 4-part Training Session Series*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=329>

27

RDA for the Seasoned Copy Cataloger – Print Monographs (Part 2 of 4)

9:30 a.m. – 12 Noon CST

*Online, 4-part Training Session Series*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=329>

## MAY

---

1

The All-New ProQuest

12:30 p.m. – 1:30 p.m., CST

*Online Reference Outreach & Instruction Training*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=318>

1

RDA for the Seasoned Copy Cataloger – Print Monographs (Part 3 of 4)

9:30 a.m. – 12 Noon CST

*Online, 4-part Training Session Series*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=329>

2

Connexion Client Module 2: Basic Searching

1:00 p.m. - 3:00 p.m., CST

*Online OCLC Training*

<http://z.umn.edu/6g0>

3

ELM: More Than Just Articles

12:30 p.m. – 1:30 p.m., CST

*Online Reference Outreach & Instruction Training*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=362>

3

Delivering Curriculum Content through Shareable Online Lesson Plans

2:30 p.m. – 3:30 p.m., CST

*Online Reference Outreach & Instruction Training*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=378>

3

RDA for the Seasoned Copy Cataloger – Print Monographs (Part 4 of 4)

9:30 a.m. – 12 Noon CST

*Online, 4-part Training Session Series*

<https://www.minitex.umn.edu/Training/Details.aspx?SessionID=329>

8

ELM and Local Government

3:30 p.m. – 4:00 p.m., CST

*Online Reference Outreach & Instruction Training*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=377>

9

RDA for the Original Cataloger (Part 1 of 3)

9:30 a.m. – 12 Noon CST

*Online, 3-part Training Session Series*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=376>

10

Accessing Your ELM Usage Statistics

11:30 a.m. – 12:30 p.m., CST

*Online Reference Outreach & Instruction Training*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=278>

13

CONTENTdm Basic Skills 1: Getting Started

1:00 p.m. - 3:00 p.m., CST

*Online OCLC Training*

<http://z.umn.edu/6g1>

15

Connexion Browser Module 2: Basic Searching

1:00 p.m. - 3:00 p.m., CST

*Online OCLC Training*

<http://z.umn.edu/6g6>

15

Copy Cataloging with CatExpress

10:00 a.m. - 11:00 a.m., CST

*Online OCLC Training*

<http://z.umn.edu/6ga>

16

RDA for the Seasoned Copy Cataloger - Continuing Resources (Part 1 of 4)

1:30 p.m. - 4:00 Noon CST

*Online, 4-part Training Session Series*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=352>

18

RDA for the Seasoned Copy Cataloger - Continuing Resources (Part 2 of 4)

1:30 p.m. - 4:00 Noon CST

*Online, 4-part Training Session Series*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=352>

22

RDA for the Seasoned Copy Cataloger - Continuing Resources (Part 3 of 4)

1:30 p.m. - 4:00 Noon CST

*Online, 4-part Training Session Series*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=352>

24

RDA for the Seasoned Copy Cataloger - Continuing Resources (Part 4 of 4)

1:30 p.m. - 4:00 Noon CST

*Online, 4-part Training Session Series*

<http://www.minitex.umn.edu/Training/Details.aspx?SessionID=352>

## JUNE

---

1

CONTENTdm Basic Skills 2: Working with Text

1:00 p.m. - 3:00 p.m., CST

*Online OCLC Training*

<http://z.umn.edu/6g2>

4

CONTENTdm Basic Skills 3: Maintaining Collections

1:00 p.m. - 3:00 p.m., CST

*Online OCLC Training*

<http://z.umn.edu/6g3>

19

Copy Cataloging with CatExpress

1:00 p.m. - 2:00 p.m., CST

*Online OCLC Training*

<http://z.umn.edu/6gc>