Minitex

A Publication of the Minitex Digitization, Cataloging & Metadata Education Unit

DIGITIZATION, CATALOGING & METADATA MAILING

July

Announcements

MDL Staff Join Minitex

Sara Ring, Minitex/DCME

The Digitization, Cataloging & Metadata Unit recently expanded! Greta Bahnemann, Metadata Coordinator for the Minnesota Digital Library (MDL), joined us this spring. Previously, her office was within the University of Minnesota Libraries' Digital Library Services Department (just down the hall from Minitex). Structural changes that resulted from MDL strategic planning led to MDL's and Greta's transfer to Minitex.

Greta has an undergraduate degree in History and Art History from the University of Minnesota, Morris. She has a Master's Degree in Early American Culture from the Winterthur Program and the University of Delaware, worked in the museum field for 10 years, then received her library degree from the University of British Columbia in Vancouver. After that, she worked at Cornell University as a fine arts cataloger and joined MDL in 2010.

You may see Greta out and about at various library conferences, or work directly with her if your library submits a project for digitization and inclusion in *Minnesota Reflections*. ◼

Cataloging & Metadata

RDA Q&A

Mark K. Ehlert, Minitex/DCME

Topic #1: Reporting RDA Errors

Question
*It appears that **editor of compilation** has been removed from Appendix I. However, I.2.1 makes a reference to it under **compiler:***
*"**compiler:** A person, family, or corporate body responsible for creating a new work (e.g., a bibliography, a directory) by selecting, arranging, aggregating, and editing data, information, etc. For a compiler as a contributor, see **editor of compilation** at I.3.1."*

But the term is not in I.3.1 nor anywhere in RDA that I can find. Is this an error? If so, to whom do we report these? Or do you?

Answer
Yes, this is indeed an error. The relationship designator *editor of compilation* was removed from RDA with the April 2014 update. That term's particular role is now part of *editor*, also found under I.3.1.

You can notify ALA Publishing of errors in the RDA text via the RDA Toolkit website:

1. Go here: www.rdatoolkit.org/support
2. Click the "Open New Ticket" button
3. Fill in your name, email address, and (optionally) phone number
4. From the Help Topic drop-down menu choose "Content Related"
5. Fill in the subject line and post your message
6. Click the "Submit Ticket" button
7. You'll get a confirmation message via e-mail that includes a service ticket number and a link to your message

It may take a few days, but someone from ALA Publishing will respond. They pass on these content-related messages to the Joint Steering Committee for Development of RDA (JSC). For technical and display issues, however, responsibility for corrections lies with ALA Publishing as administrator of the system infrastructure that drives the Toolkit site.

Below is a sample of errors I recently submitted through the aforementioned ticketing system:

- RDA 2.3.4.5 and 2.3.4.6: the last two sentences under each instruction are duplicates.
- RDA 3.4.1.7.1: the last sentence points to the wrong note (3.21.2.10); should it instead point to 3.21.2.11?
- RDA 6.9.1.3: in the last sentence under the Alternative, the word *in* appears twice.
- RDA 6.16.1.3.1: in the 1st block of examples, 1st example, the medium of performance is given as *strings* (the old form) rather than *violins, viola, cello* [or *violoncello*] (the new form).
- RDA 6.23.2.9.3: *TABLE 6.3* should be renumbered to "6.2" due to the April 2014 demise of the *Standard Combinations of Instruments* table under 6.15.1.5, which had been numbered "6.2."
- RDA 6.25.1.4: regarding the if-then list about halfway down, it's unclear whether this is an *or* list, an *and* list, or an *and/or* list.
- RDA 6.28.1.11: the last sentence is missing the term *collective*; it should likely begin "Record the conventional **collective** title…"
- RDA 10.2.2.4: the underlined *NAME* header should appear above 10.2.2, matching the position of the same header over 9.2.2 and 11.2.2.

As you can see, many errors are quite minor, others are uncertain. But reporting even small ones helps make RDA's instructions more meaningful to catalogers.

Topic #2: Relationship Designators for Joint Enterprises

Question
What is the accepted procedure for adding relator terms for people who jointly create a work? My guess is that they all get the same relator term so that they will appear in the same index.

Answer
You're correct. Whether there's one person or five persons responsible for writing the words to a particular song, they each get the $e designator *lyricist* posted with their name access points.

```
245 00 Auf wiederseh'n sweetheart / ‡c
        Storch, Sexton, Turner. ...
700 1_  Parsons, Geoffrey ‡c (Lyricist), ‡e
        lyricist.
700 1_  Phillips, James John Turner, ‡e
        lyricist.
```

If you have more than one composer responsible for a single musical work, all names get the *composer* stamp, even if one name appears in the 100 field and the remaining names in the 700(s).

```
100 1_  Smith, Patti, ‡e composer.
245 10 Because the night : ‡b LP version / ‡c
        Smith, Springsteen.
700 1_  Springsteen, Bruce, ‡e composer.
```

The same holds true for other joint ventures: multiple writers of a single story each get *author*; multiple editors of a single dictionary each get *editor*; multiple narrators in a single film each get *narrator*. Any application of, say, *co-author* or *joint composer* that you find in RDA records—terms that are variations on those in RDA's Appendix I—is a local practice.

Topic #3: Note Order

Question
When adding notes to bibliographic records, are they supposed to be in numerical MARC tag order or do they go in another prescribed order? For DVDs, for example, the fields could appear in this manner: 538, 546, 500, 511, 508...

Answer
RDA prescribes no order to notes. Any direction on this point would be made by specialized cataloging communities or through locally established practices. (There's no easy automated way in MARC-based catalogs to arrange notes in a particular display pattern

Minitex

short of forcing tag order. Other courses of action involve intervention via special programming or manual shifting of MARC fields.)

An example of a community practice is found in serials cataloging, where it's common to give notes in MARC tag order. It seems this continues under RDA, though that section of the *CONSER Cataloging Manual*, now in the middle of an extensive RDA update, hasn't yet been revised:

www.loc.gov/aba/pcc/conser/more-documentation.html

Where the question of practice is met with silence, catalogers may put notes in any order they choose or continue following AACR2's 1.7B and related rules in later chapters. This retains a consistency with older records across the catalog—that is, if this is even noticed by the user. But with the automated ingestion of hundreds or thousands of bib records of varying quality from sundry sources into catalogs these days, note order may be, regrettably, a thing of the past.

Topic #4: Large Print

Question
Is there a set way to enter "large print" in the 3xx fields? Most of our records are coming through with "pages (large print)" entered in the 300 $a. Maybe 10% have the large print in the 340 $n subfield.

Which way is correct? Or perhaps, which way is "more" correct??

Answer
There is no standard way of describing large-print font sizes in RDA-MARC records. Which is to say that no best practices have yet come to pass from vetted sources like LC or OLAC or the Music Library Association.

RDA is a how-to-fill-in-a-blank-line guide rather than a how-it-should-display guide. In that sense, using the 340 field is more in line with its expectations. That field's $n is a discrete slot for holding and indexing font size terms from RDA 3.13, like the prescribed *large print* and *giant print*, or cataloger-supplied wording—preferably from a controlled list—such as *larger print*. (The same RDA instructions describe font size for tactile text, which I put aside here.)

The 340 field also serves as the source from which programming can customize various displays and indexing. This MARC coding provides and example:

```
300 __  ix, 123 pages : ‡b maps ; ‡c 29 cm
336 __  text ‡2 rdacontent
336 __  cartographic image ‡2 rdacontent
337 __  unmediated ‡2 rdamedia
```

```
338 __  volume ‡2 rdacarrier
340 __  ‡n large print (20 point)
```

In the public catalog, this might yield:

> **DESCRIPTION (ILLUSTRATED BOOK):** ix, 123 pages (large print (20 point)) : maps ; 29 cm

or this, highlighting the font size by breaking it out separately:

> **EXTENT (ILLUSTRATED BOOK):** ix, 123 pages
> **ILLUSTRATIONS:** maps
> **SIZE:** 29 cm high
> **PRINT SIZE:** large print (20 point font size)

If such programming is not available, should the 340 field (or any 34x or 33x field, for that matter) be projected on the screen as-is and/or indexed behind the scenes? What would such a display look like to the user? How would the indexing function as part of a user's search experience?

And if *not* leveraging the 340, whether by choice or by system limitation, what about the 300 field?

AACR2 2.5B23 tells us to add a parenthetical *large print* after the page/volume information in the Physical Description Area—in MARC, the 300 field. The familiar display certainties of this field plus the fact that few catalogs today do anything imaginative with the 340 have led some catalogers to opt for either a) duplicating the content of the 340 $n in the 300 field (thus covering present and future needs), or b) posting the font size term in the 300 field alone (thus treading wholly down AACR2's how-it-should-display path). Either method might supplement or replace more eye-catching tokens, like an icon, generated from the long-standing fixed field Form 008/23 code **d**.

Speaking of Form codes, **d** is only defined as *large print* in the generic sense, thus lacking the nuance provided by RDA's terminology. [1] Some library users may want to easily distinguish between large print and giant print texts when using the catalog. ■

Cataloging & Metadata

Linked Data and the (Public) Library Catalog

Mark K. Ehlert, Minitex/DCME

An article in *SCATNews: Newsletter of the Standing Committee of the IFLA Cataloguing Section* caught my eye

---

1 www.oclc.org/bibformats/en/fixedfield/form.html

recently. It was written by Asgeir Rekkavik, a consultant with the Oslo Public Library. He describes how the new Public Library, opening in 2018, will be "a digital discovery center where the presentation of the physical collection will merge with digital content and user generated content." This is not a unique goal; libraries here in the States have announced similar objectives. But the author goes on to explain that:

> This calls for new ways of describing both physical and digital content, and for new ways of working with cataloguing in the library. To get there, the library has decided to drop their integrated library system and to drop MARC as [the] cataloguing format. Instead we will use RDF linked data as the primary cataloguing format, starting already in 2015.

Rekkavik recounts how Oslo Public Library staff have been exploiting cataloging data as linked data since 2010, thus gaining experience for future, (inter)national linked data endeavors while furnishing new services to the public in ways that the author argues could not have been possible through "the use of ordinary MARC records."

The root of this sea change, according to Rekkavik, is that "it is the public, the patrons[,] the data should be made to facilitate," whereas up until now, catalogers "have produced catalogue data as a tool for librarians, so that they in turn could use them to assist the patrons"—an inefficient model. The transition to linked data is a means to empowering library users.

> Cataloguing for the public should avoid making assumptions about what patrons will be interested in or in what their motivation for being interested in it would be. It should simply focus on making the data as rich and expressive as possible, so that it can be applied and combined in as many ways as possible.

Read the rest of the article in the June 2014 issue of *SCATNews*, pages 13–16. It's available in PDF format here: www.ifla.org/cataloguing/newsletter ■

SERIALST Transition to a New Home
Post to SERIALST (July 22, 2014); edited

We are writing today to let you know what's going to happen shortly with this email list.

SERIALST (Serials in Libraries Discussion Forum) has always been hosted by the University of Vermont, where our founder, Birdie MacLennan, worked. With her death

earlier this year, we realized the university would probably not want to continue hosting the list permanently. We began searching for a new host.

We quickly settled on NASIG, the North American Serials Interest Group, as a potential new home: www.nasig.org

NASIG is well-known within the serials community, and has experience hosting and managing their own web and email services. All the people involved with running SERIALST have been active NASIG members. NASIG's philosophy of being a common forum for all parties in the serials community (aggregators, libraries, publishers, subscription agents, etc.) fits very well with the way SERIALST has always worked.

When we approached them, NASIG was immediately interested in the idea, and negotiations have moved quickly to fruition. So, on a date in the near future, SERIALST will transition from the University of Vermont's servers to NASIG's. The list archives will also move to a NASIG server.

The immediate effects will be fairly minimal:

- Subscribers will *not* have to re-subscribe to the list; your subscription and settings should be ported to NASIG with little or no difficulty.
- NASIG has contracted with L-Soft to provide list-hosting services.
- Since L-Soft currently provides those same services for the University of Vermont, that should make the transition especially easy.
- The email address for sending messages to the list will change. We'll send the new address to the list once the transition timetable is settled.
- The web address of the archives will change, but the full archive of past messages will continue to be available.

Though there may be minor changes, NASIG plans to generally continue the editorial policies currently in place, which can be found at: www.uvm.edu/~bmaclenn/serialst.html

SERIALST will still be open to everyone; there will be no expectation that people be NASIG members to participate in the list. NASIG will continue to run their own listserv, NASIG-L, for communicating with their membership.

The list will be moderated by members of NASIG's Communications and Marketing Committee, with Beth Ashmore of Samford University as the lead moderator.

More details will follow shortly, including a timeline for the migration. In the meantime, please feel free to contact us if

Minitex

you have questions.

Stephen Clark (sdclar@wm.edu)
Ann Ercelawn (ann.ercelawn@vanderbilt.edu)
Bob Persing (persing@pobox.upenn.edu)
SERIALST associate moderators

NASIG issued a parallel announcement on their blog at:

nasig.wordpress.com/2014/07/21/nasig-and-serialst-breaking-news ■

Cataloging & Metadata

BIBFRAME Notes

Mark K. Ehlert, Minitex/DCME

Considering BIBFRAME and Audiovisual Materials

BIBFRAME (BF) developers at the Library of Congress (LC) commissioned a report from Audiovisual Preservation Solutions,[1] asking how BF can service the description of and access to audiovisual materials (moving image and sound recordings) in a linked data world. These objects have unique characteristics that fall outside the scope of common text-based resources; at the same time, the latter have been the substance around which many library cataloging standards have been built. AACR1's descriptive rules for "phonorecords" read like an afterthought, for instance.

BIBFRAME AV Modeling Study, at a little over 50 pages with appendices, analyzes elements from forward-looking bibliographic instruments both well-known (RDA, FRBR) and less well-known (<indecs>,[2] the draft of the International Federation of Film Archives' moving image cataloging manual,[3] metadata sourced from the Variations3 project out of Indiana University[4]). How can these inform the development of BF for the benefit of library users, researchers, archivists, and catalogers interested in searching, retrieving, describing, and preserving audiovisual resources? The authors cast their response as a series of recommendations for the framers of BF.

This study is likely the first of several that will affect BF development as the new standard makes steady progress toward implementation. Even at 50 pages, I found *BIBFRAME AV Modeling Study* a fairly quick read: a mostly straightforward and at times illuminating account of current audiovisual metadata work (whether in libraries, archives, broadcasting), its relation to user expectations, and future possibilities for serving the public. There are a few technical thickets the reader must to cut through, but I still recommend this work to those catalogers who are interested in the future of audiovisual material description.

Sally MacCullum, Chief of the Network Development and MARC Standards Office at LC, announced the publication of the report on the BIBFRAME discussion list on July 25, 2014. Her message, reprinted below, has been lightly edited for this article.

> Just posted on the BIBFRAME (BF) web site[5] is a study carried out by Audiovisual Preservation Solutions for the Library of Congress concerning an appropriate model for AV material and relating that to the BF model: *BIBFRAME AV Modeling Study: Defining a Flexible Model for Description of Audiovisual Resources*.[6] It discusses the special characteristics of AV material that make it different from textual and other media and makes some general recommendations. With the increasing impact of AV as information resources and the need to preserve material in those media, special attention to it is appropriate as we develop BIBFRAME.
>
> This report was carried out in close consultation with the Library of Congress's National Audiovisual Conservation Center (NAVCC) staff in Culpeper, Virginia, a vast archive that preserves and serves AV resources. The report examines a number of proposed community models such as FRBR/RDA, FIAF, OLAC, indecs, Variations, PBCore, EBCore and others that inform the analysis. The recommendations will become considerations for the BF model development as there are various ways the special attributes of these media could be accommodated.
>
> The report is composed of a base document (the analysis, AV model, and recommendations) with appendices that give more detail about situations encountered with AV material.
>
> As usual, comments, concerns, and discussion are encouraged via the BIBFRAME listserv (see the BIBFRAME home page/contacts)[7] or direct to bfcomment@loc.gov. ■

---

1 www.avpreserve.com

2 www.doi.org/factsheets/indecs_factsheet.html

3 www.filmstandards.org/fiaf/wiki/doku.php?id=start

4 www.dlib.indiana.edu/projects/variations3/metadata/guide

5 www.loc.gov/bibframe

6 *Report and appendices available here (PDF format):* z.umn.edu/npm

7 www.loc.gov/bibframe/contact/index.html

RDA Original Cataloging Training: Outstate Edition

Mark K. Ehlert, Minitex/DCME

Registration is now open for two *RDA Original Cataloging* workshops in outstate Minnesota: Bemidji and St. Cloud. These two-day, on-site training sessions will focus on practical cataloging matters via lecture and exercises. We will revisit FRBR, then step our way through the RDA text to build bibliographic records and access points for a variety of monographic materials. We will also take a look at RDA's depictions of relationships in the catalog.

Visit the links below for full workshop descriptions, prerequisites (including assigned readings), and registration details. **Please note registration deadlines.**

Low enrollment will result in the rescheduling or cancellation of workshops.

A.C. Clark Library, Bemidji State University
Day 1: Monday, September 22, 9:00 a.m. – 4:30 p.m.
Day 2: Tuesday, September 23, 8:30 a.m. – 4:00 p.m.

z.umn.edu/nk2

James W. Miller Learning Resources Center, St. Cloud State University
Day 1: Thursday, September 25, 9:00 a.m. – 4:30 p.m.
Day 2: Friday, September 26, 8:30 a.m. – 4:00 p.m.

z.umn.edu/nk4 ■

ALA Born-Digital Preconference Summary

Sara Ring, Minitex/DCME

In June I attended a preconference at the ALA Annual Conference on managing local and community-produced born-digital media. Within an archive, home movies might come in from a donor as part of a larger collection. Within a university library or archive, one might collect university videos of performances, sporting events, campus lectures, or even media as part of a student thesis. Several speakers presented strategies for collecting and managing this type of content. What follows is an overview of the content covered and a few takeaways, from my perspective. Handouts and even a few audio recordings of the speakers

are available here:

z.umn.edu/npj ■

Overview

How is locally produced content different from other types of collections in libraries and archives? As the first presenter (Howard Besser from NYU's Moving Image Archiving and Preservation Program) explained in the introduction and overview section of the day, these types of materials often do not include consistent file-naming conventions, folder arrangements, or metadata, and may come in various file formats that you as the library or archives need to normalize (migrate to a file format that can be read or that is more stable for long term access). The quality of the content a library or archive receives will also vary. For example, if the digital content is coming from various individuals or departments within a university, no two people or departments will capture and organize their digital content in the same way.

Besser outlined general problems of handling born-digital personal content. In the analog world, when we (ibraries and archives) work with physical content, we can see the various drafts of a document, where text has been crossed out and re-written, etc. In the digital world, where are these traces of thought processes found? When a person creates a document on a computer, a record of the changes disappears as one writes over and re-saves drafts. In libraries and archives, we are interested in collecting the drafts as they can reveal an author or a scientist's changing thoughts about their work. Finding the digital equivalent of drafts/changes to a work will require new tools and new approaches such as digital archeology and digital forensics.

Besser separated the challenge into two stages. In Stage 1 "people write on computers instead of paper." One of the biggest issues here is obsolescense. Libraries and archives may receive collections that are on outdated media. Finding a device that can still play or read the media (e.g. zip discs, 8" floppy), finding an interface plug to hook it to a current computer, and making sure that a recent operating system can even recognize the device, can be challenging.

Once you can read it, getting content off of donor computers and media will also require new approaches. The initial step to getting content off of storage media is to create a disk image. If you simply copied and pasted files from one device to another, a lot of information can be lost in the process. By doing a disc image, you can maintain file edit histories, hidden files, user profiles and other settings. After you have the disk image, you can begin to explore the directory structure of the disk image. You may need to emulate or migrate the file format just to be able to read the file or group of files. You may need to do a bit of

Minitex

digital forensics to decide whether you wish to save the edit histories of the files, and you will probably have to make some decisions about what to redact due to privacy issues (credit card numbers, etc).

In Stage 2, people no longer store their digital works in places over which they have absolute control (email services, cloud storage for documents, social networks). Many terms of service for 3rd-party hosting services (email, social media) don't allow anyone else to use your account. As a library or archive receiving a collection from a donor, it can be difficult or even impossible to get data from third parties.

One of the most intriguing points Besser made was that, as archives, if we hope to preserve digital content, we need to be involved earlier in the life-cycle of the digital objects that we're collecting, long before the content even reaches the archive. We can educate and help our donors at the point of creation so that the content coming to us is not on obsolete media or in file formats we can no longer read.

We can support content creators so that when the content comes to us, we know what it is. Besser's example was of a photograph, with people sitting on a porch where a lake had flooded their residence. Without any context (descriptive information) from the content creator, it would have just looked like a photograph of a flood. In actuality, the flood was caused by a land developer who wanted to get the people in the town to leave their homes. The importance of getting involved early in the digital life cycle was repeated throughout the day.

Examples of Organizations Engaging Digital Content Creators

The Activist Archivist
The Activist Archivists organization was mentioned as a case study. This group of media archivists and academics provides assistance on archiving/preservation to improve access to media content that is being produced by individuals and community activists. They have created the following resources:

- 7 Tips to Ensure Your Video is Usable in the Long Term
- Best Practices for Video Activists
- Tips for Catalogers
- A "Why Archive?" postcard
- Creative Commons Guidance

These guides and handouts are intended for this group's particular audience, but one could consider doing something similar within their own library or archive, for your audience. You can check out the resources they have created here:

activist-archivists.org/wp

WITNESS
An archivist from the organization WITNESS (Yvonne Ng) spoke about the resources they developed for their community. WITNESS trains and supports people using video to document human rights abuse. They offer training on how to use video, host their curriculum on their website for anyone to access, and post their materials on GitHub for re-use. Examples of materials on GitHub include a tip sheet for filming on your mobile phone, obtaining informed consent, and a tip sheet on uploading videos to YouTube. The raw files are provided on the GitHub site, and would be easy to re-use and customize for various audiences. Their impressive online Archive Guide features best practices for organizing, storing, preserving, and sharing video footage. Best practices are grouped in easy-to-digest categories that follow the digital object life cycle: create, transfer, acquire, organize, store, catalog, preserve, share. View the WITNESS resources webpage:

witness.org/resources

Personal Digital Archiving
The Library of Congress personal digital archiving website, although not discussed during this presentation, is another resource that promotes the philosophy of getting involved earlier in the digital life cycle process with content creators. Their website hosts overview guides (introduction to scanning, why digital preservation is important), guides on how to preserve specific types of digital content (digital audio and video included), and a personal digital archiving day kit. If you are a library or archive and wish to hold a personal digital archiving event, take a look at the kit:

digitalpreservation.gov/personalarchiving

Queens History Days
The Queens Library (Queens Memory Project) in Queens, New York has been getting involved in the digital life cycle at the content creation stage by holding "Queens History Days." They invite individuals from the community to come in to the library with their photographs and documents. They provide scanners, digital recorders (for oral histories), and consent forms. Citizens bring their content and give information about each file to library staff. The library staff then contextualize these citizen photos and objects with photos/objects from their collection in Queensmemory.org.

How did they do it? The library had to accommodate a new workflow and build new accession procedures. They trained library staff on accessioning (working with citizens to extract information from them about their photos and documents). Only one person at a time does the scanning so the file naming is all consistent. They also trained one of their staff to conduct oral histories interviews. The consent form they created incorporates Creative Commons licensing of the content.

At end of the day, each donor has a folder that contains:

- digital objects
- accession form (information is immediately mapped to the standards that they use in Queens Memory and on the web: VRA Core and EAD)
- a photo of the donor
- signed release form
- notes (metadata)

For full archival processing, donations of eight items or more are considered a collection. Each collection gets an EAD finding aid created for the Queens Library digital archives website.

Item level records are created by metadata services staff and ingested into their Digital Object Management System to be displayed on the digital archives website.

They see the following benefits to implementing a program like this:

- They now have a pipeline for high quality crowd-sourced donations
- Public programs provide unique opportunities for teaching digital literacy, cross-generational and cross-cultural learning
- The program helps to diversify staff skills and create opportunities for collaboration by cultivating a deeper understanding of expertise across the organization

Tools for Managing Born-Digital Content

Other speakers at the preconference gave presentations ranging from getting information technology staff support to help manage born-digital content to an overview of digital video properties, formats, and streaming options. The last segment of the preconference focused on using various tools to manage born digital content. After a brief tutorial on how to use the command line interface on a Mac, various presenters demonstrated (or discussed) the following tools.

Fixity (Checksum) Tools

Rsync
rsync.samba.org
A command-line open source tool used to transfer or copy files from one location to another. It is also a common tool used for making backups. Rsync creates checksums to ensure proper file transfer (Checksums are like digital fingerprints, generating number strings using a mathematical algorithm). It provides confirmation logs, can preserve file attributes (like date modification times), and can be used in combination with other programs or scripts.

Md5
www.fourmilab.ch/md5
Uses the MD5 algorithm to produce a hash (checksum) on a file. It can be used to generate checksums, compare files and check for corruption. There is also a program called Md5deep where you can compute other types of checksums, beyond just MD5 (SHA-1, SHA-256, for example)

Fixity (From AvPreserve)
www.avpreserve.com/avpsresources/tools
Fixity is a tool you can use to regularly scan files or directories to make sure files have not become corrupt. Fixity monitors files by creating and validating checksums, and emails reports to the user.

Tools for Tranferring Files

BagIt
www.digitalpreservation.gov/documents/bagitspec.pdf
BagIt is a hierarchical file packaging format for storing or transferring digital content.

Bagger
sourceforge.net/projects/loc-xferutils/files/loc-bagger
Bagger is a tool used to produce a package of files according to the BagIt specification. Yvonne Ng from WITNESS uses BagIt/Bagger to package her master files and send them to another organization, that stores and preserves them.

Metadata Tools

Mediainfo
mediaarea.net/en/MediaInfo
A convenient display of the most relevant technical metadata from video and audio files.

MDQC (Metadata Quality Control)
www.avpreserve.com/blog/avpreserve-releases-fixity-and-mdqc-digital-preservation-tools
A free open-source tool for comparing embedded technical metadata against a profile determined by the user. For example, you could use it to compare expected technical metadata against actual technical metadata on a group of files.

File Formats

ffmpeg
www.ffmpeg.org
An open source command-line tool that transcodes and encodes video and audio.

Sustainability of Digital Formats (Library of Congress)
www.digitalpreservation.gov/formats

Minitex

A guide to digital content formats to help determine file sustainability. ■

Digitization & Preservation

Minnesota Reflections Hits 200,000

Sara Ring, Minitex/DCME

Minnesota Reflections hit the 200,000 digital object mark this summer! Two new contributors brought us over 200,000: The Saint Paul Almanac and Quatrefoil Library. Many thanks to them and to each of the more than 160 cultural heritage organizations that helped us reach this milestone.

About Saint Paul Almanac

The Saint Paul Almanac is a literary organization that creates opportunities for understanding, learning, and building relationships through sharing people's stories. Every year the organization publishes its *Almanac* which includes stories, poems, and articles about Saint Paul written by local authors.

z.umn.edu/nlm

About Quatrefoil Library

Quatrefoil Library is a member-supported lending library, located in Minneapolis, that collects, maintains, documents and circulates gay, lesbian, bisexual, transgender and queer materials and information in a safe and accessible space. Established in 1986, Quatrefoil is managed by an all-volunteer board and staff. The library collections include books, periodicals, DVDs and CDs. Quatrefoil's contributions to *Minnesota Reflections* are digitized versions of two Twin Cities GLBT community newspapers (available in the Minnesota Newspapers Collection) and *Focus Point*, a LGBT periodical which ceased publication in 2000.

z.umn.edu/nll ■

Minitex Digitization, Cataloging & Metadata Mailing
Contact Information

The *Digitization, Cataloging & Metadata Mailing* is an informational bulletin sent monthly to libraries in the Minitex region. 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.

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.

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

Minitex
DEDICATION. EXPLORATION. INNOVATION.

JULY 2014

MDL Staff Join Minitex

RDA Q&A

ALA Born-Digital Preconference Summary

DIGITIZATION, CATALOGING & METADATA MAILING

A Publication of the Minitex Digitization, Cataloging & Metadata Education Unit

Minitex

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.

AUGUST

13

NISO Webinar: "Streamlining and Simplifying: Advances in Consortial Licensing"
12:00 p.m. – 1:30 p.m., CST
University of Minnesota - Twin Cities
www.minitex.umn.edu/Events/Niso/#licensing

18-19

Upper Midwest Digital Collections Conference
Conference
**University of St. Thomas, Minneapolis Campus
Minneapolis, MN**
www.wils.org/news-events/wilsevents/umdcc

20

Connexion Client Module 4: Save Files, File Management, and Batch Processing
1:00 p.m. – 3:00 p.m., CST
DCME/OCLC Online Training Session
www.minitex.umn.edu/Training/Details.aspx?SessionID=409

21

Minitex Annual Public Library Node Meeting
Conference
Brookdale Library's Conference Center, Shingle Creek Parkway, Brooklyn Center, MN
www.minitex.umn.edu/Events/Conferences/PublicNode2014.aspx

25

CONTENTdm Basic Skills 1: Getting Started
1:00 p.m. – 3:00 p.m., CST
DCME/OCLC Online Training Session
www.minitex.umn.edu/Training/Details.aspx?SessionID=389

27

CONTENTdm Basic Skills 2: Working with Text
1:00 p.m. – 3:00 p.m., CST
DCME/OCLC Online Training Session
www.minitex.umn.edu/Training/Details.aspx?SessionID=390

29

CONTENTdm Basic Skills 3: Maintaining Collections
1:00 p.m. – 3:00 p.m., CST
DCME/OCLC Online Training Session
www.minitex.umn.edu/Training/Details.aspx?SessionID=391

SEPTEMBER

1

Labor Day
Minitex Offices closed

9

Connexion Client Module 5: Automation and Customization
1:00 p.m. – 3:00 p.m., CST
DCME/OCLC Online Training Session
www.minitex.umn.edu/Training/Details.aspx?SessionID=443

10

NISO, 2-Part Webinar: "E-books for Education, Part 1: Electronic Textbooks - Plug in and Learn"
12:00 p.m. – 1:30 p.m., CST
University of Minnesota - Twin Cities
www.minitex.umn.edu/Events/Niso/#ebooks1

11-12

RDA Original Cataloging
9:00 a.m. – 4:30 p.m. CST
Two-part, In-person Training Session
310 Walter Library, University of Minnesota, Twin Cities Campus, Minneapolis, MN
www.minitex.umn.edu/Training/Details.aspx?SessionID=515

17-19

NDLA Annual Conference: "Library Services for the Next 125 Years"
Conference
Ramada Inn, Bismarck, ND
ndla.info/conference/2014

17

NISO, 2-Part Webinar: "E-books for Education, Part 2: Open Textbook Initiatives"
12:00 p.m. – 1:30 p.m., CST
University of Minnesota - Twin Cities
www.minitex.umn.edu/Events/Niso/#ebooks2

22-23

RDA Original Cataloging for Bemidji
9:00 a.m. – 4:30 p.m. CST
Two-part, In-Person Training Session
A.C. Clark Library, Bemidji State University, Bemidji, MN
www.minitex.umn.edu/Training/Details.aspx?SessionID=543

23

File Not Found: An Introduction to Digital Preservation
1:00 p.m. – 2:00 p.m., CST
DCME Online Training Session
minitex.umn.edu/Training/Details.aspx?SessionID=395

24

NISO Virtual Conference: "Library Data in the Cloud"
10:00 a.m. – 4:00 p.m., CST
University of Minnesota - Twin Cities
www.minitex.umn.edu/Events/Niso/#cloud

25-26

RDA Original Cataloging for St. Cloud
9:00 a.m. – 4:30 p.m. CST
Two-part, In-Person Training Session
Miller Center, St. Cloud State University, St. Cloud, MN
www.minitex.umn.edu/Training/Details.aspx?SessionID=542

OCTOBER

1-3

SDLA Annual Conference: "Libraries - The Center of It All"
Conference
Best Western Ramkota, Pierre, SD
www.sdlibraryassociation.org/?page=32

2-4

Annual MEMO (MN Educational Media Organization) Fall Conference: "MADE IN MN - Information and Technology Educators of Minnesota"
Conference
River's Edge Convention Center / Kelly Inn, St Cloud, MN
memotech.ning.com/page/2014-fall-conference

8-9

Minnesota Library Association (MLA) Annual Conference: "Better Together"
Conference
Verizon Wireless Center, 1 Civic Center Plaza, Mankato, MN
www.mnlibraryassociation.org/events/event_details.asp?id=423359

15

NISO Webinar: "21st Century Resource Sharing: Which Inter-Library Loan Standard Should I Use?"
12:00 p.m. – 1:30 p.m., CST
University of Minnesota - Twin Cities
www.minitex.umn.edu/Events/Niso/#sharing

23

PALS User Group Meeting
Conference
www.mnpals.org/content/user-group-meeting-fall-2014