

All the Cool Kids are Doing it: Website Cataloging in Connexion



Tessa L. H. Minchew, Systems/Catalog Librarian
Georgia Perimeter College
tminchew@gpc.edu
<http://www.gpc.edu/~tminchew>

GaCOMO 2007
October 19, 2007



Extracting Metadata In Connexion

- Client or browser will extract metadata from an electronic file and create a preliminary record for further editing
- Supported file types
 - HTML (htm or *.html files)
 - PDF (*.pdf files)
 - Word (*.doc files)
 - MP3 (*.mp3 files)

2



Extracting Metadata In Connexion

- Connexion will also extract metadata in non-Latin scripts
 - Supported non-Latin scripts: Arabic, Bengali, Chinese, Cyrillic, Devanagari, Greek, Hebrew, Japanese, Korean, Tamil, and Thai

3



Extracting in the Connexion Browser

- On **Cataloging** tab, click the **Create** button
- On **Single Record** menu, select **Extract Metadata**
- Select the appropriate section depending upon whether you want to create single or multiple records

4



Extracting in the Connexion Browser

- Enter the URL for the resource in the field next to **Extract data from URL**
- Click appropriate radio box if you wish to generate Dewey numbers or apply constant data
- Click the **Create** button at the bottom right

5



Extracting in the Connexion Client

- You can extract metadata while logged on or offline, just define a default authorization and password in Tools > Options > Authorizations
- Click Cataloging > Create > Extract Metadata
- Enter the URL for the resource in the field under **Resource (fully qualified URL or path)**
- Examine all settings, adjust accordingly, and click OK

6

006 - Additional Material Characteristics

- Include field 006 for computer file, if Type is not "m"
- Create 006 option at bottom right of template
- Click on arrow to open menu of 006 types and select "Computer Files"

7

006 - Additional Material Characteristics

- Audn - Target Audience
 - Blank - Unknown or unspecified
- File - Type of Computer File
 - j - Online system or service
- GPub - Government Publication
 - Blank - Not a government publication
 - f - Federal/National
 - s - State, provincial, territorial, dependent, etc.

8

008 – Fixed Fields

- For websites, use the Continuing Resources (formerly Serials) workform. (When extracting metadata from a website, Connexion defaults to this)
- Type - Type of Record
 - a - Language material
 - m - Computer file (Software, numeric data, computer-oriented multimedia, online systems or services)

9

008 – Fixed Fields

- Type - Type of Record
 - Code for the most significant aspect of the resource. With websites this will usually be language material. When in doubt consider the item a computer file.
- BLvl - Bibliographic Level
 - i - Integrating resource (for all updating websites)
- S/L Entry Convention
 - 2 - Integrated entry (Record cataloged under latest title)

10

008 – Fixed Fields

- Desc - Descriptive Cataloging Form
 - a - AACR2
- ELvl - Encoding Level
 - K - Less-than-full input by OCLC participants
- Form - Form of Item
 - s - Electronic
- Orig - Form of Original Item
 - blank - None of the following

11

008 – Fixed Fields

- SrTp - Type of Continuing Resource
 - w - Updating Web site (For resources that do not fit into one of the other codes, such as database, newspaper or periodical)
- Srce - Cataloging Source
 - d - Other (organization other than a national bibliographic agency or a participant in a cooperative cataloging program)

12

008 – Fixed Fields

- Conf - Conference Publication
 - 0 - Not a conference publication
 - 1 - Conference publication
- EntW - Nature of Entire Work
 - Blank - Not specified
- Cont - Nature of Contents
 - Blank - Not specified

13

008 – Fixed Fields

- GPub - Government Publication
 - Blank - Not a government publication
 - f - Federal/National
 - s - State, provincial, territorial, dependent, etc.
- Freq - Frequency
 - u - Unknown
- Regl - Regularity
 - u - Unknown

14

008 – Fixed Fields

- DtSt - Type of Date/Publication Status
 - c - Continuing or integrating resource currently published
- Ctrl - Type of Control
 - Blank - No specific type of control
- MRec - Modified Record
 - Blank - Not modified

15

008 – Fixed Fields

- Alph - Original Alphabet or Script of Title
 - Select the appropriate code for the alphabet or script of field 222 or 245
- Dates - Date 1 and Date 2
 - Date 1 - beginning date of publication
 - Date 2 - ending date
- Lang - Language Code
- Ctry - Country of Publication, etc.

16

008 – Fixed Fields

- Lang - Language Code
 - Language of the principal work.
 - Use MARC Code Lists for Languages
- Ctry - Country of Publication, etc.
 - State or country of publication, production or execution
 - MARC Code List for Countries

17

007 – Physical Description Fixed Field (Electronic Resource)

- Consult *OCLC Bibliographic Formats and Standards*
 - <http://www.oclc.org/bibformats/en/>
- Add 007 if missing from copy (a macro makes this a lot easier)
 - 007 __ c \$b r \$d m \$e n \$f _ \$g --- \$h m \$i u \$j u \$k u \$l u

18

AACR2 Chapter 9 (Electronic Resources)

- AACR2 9.0B1
 - Chief source of information
 - The resource itself (e.g. title screen(s), main menus, initial displays of information, home pages, and encoded metadata (e.g. TEI headers, HTML/XML meta tags))
 - If the information in these sources varies, prefer the most complete

19

AACR2 Chapter 9 (Electronic Resources)

- AACR2 9.0B2
 - Prescribed sources of information (Enclose information taken from outside the prescribed source(s) in square brackets.)
 - Title and statement of responsibility (Chief source of information, information issued by the publisher, creator, etc.)
 - Edition (Chief source of information, information issued by the publisher, creator, etc.)

20

AACR2 Chapter 9 (Electronic Resources)

- AACR2 9.0B2
 - Prescribed sources of information (Enclose information taken from outside the prescribed source(s) in square brackets.)
 - Type and extent of resource (Any source)
 - Publication, distribution, etc. (Chief source of information, information issued by the publisher, creator, etc.)
 - Physical description (Any source)

21

AACR2 Chapter 9 (Electronic Resources)

- AACR2 9.0B2
 - Prescribed sources of information (Enclose information taken from outside the prescribed source(s) in square brackets.)
 - Series (Chief source of information, information issued by the publisher, creator, etc.)
 - Note (Any source)
 - Terms of availability (Any source)

22

To Classify or Not to Classify

- Pros
 - Aids in collection assessment
 - Provides access through a call number browse
- Cons
 - Added time and effort for an item that doesn't require a shelving device
 - May be confusing to patrons who might look for the resource on the shelf (An Internet prefix can alleviate this confusion)

23

Variable Fields

- 1XX / 7XX
 - Due to diffuse or corporate responsibility, assigning a 1XX can be tricky
 - 100 should only be assigned for listed authors/webmasters if the site being cataloged is personal in nature
 - 110 can be assigned for the official site for a corporate body or government agency

24

Variable Fields

- 1XX / 7XX – Name Entries
 - Personal names and corporate bodies found in the item, and not receiving a 1XX can be traced in a 7XX
- 245/246 - Titles
 - 245 usually taken from home page screen
 - \$h [electronic resource]
 - Variant titles can be taken from any page on the website or supplied (in square brackets)

25

Variable Fields

- 260 - Publication, Distribution, Etc.
 - Taken from any available contact and date information
- 270 - Address (Optional)
 - \$m - Electronic mail address
- 300 - Physical Description
 - Not used for websites, as they are considered non-physical resources

26

Variable Fields

- 500 – General Note
 - 500 __ Title from caption (viewed on Sep. 14, 2006).
 - 500 __ Title from home page (viewed on Apr. 23, 2007).
- 520 Summary, Etc.
- 538 - System Details Note
 - 538 __ Mode of access: World Wide Web.

27

Variable Fields

- 6XX Subject Headings
 - There is not an authorized form subdivision for websites
 - "Computer network resources" is only authorized as a topical subdivision
 - "Databases" is only appropriate for actual databases, not websites
 - Connexion will sometimes supply 653s from the website metadata. Examine them and convert them to authorized headings if desired.

28

Variable Fields

- 856 - Electronic Location and Access
 - Connexion creates this field for you during extraction
 - \$u Uniform Resource Identifier
 - \$q Electronic format type
- 887 - Non-MARC Information Field
 - Optional field
 - Connexion pulls keywords from the resource and places them here, indicating the source in \$2

29

Completing the Record

- Follow the same steps as any other format
- Connexion Browser
 - Cataloging Tab
 - Edit > Validate Record
 - Action > Update Holdings
 - Action > Export Record in MARC

30

Completing the Record

- Connexion Client
 - Offline
 - Edit > Validate
 - Action > Update Holdings
 - Batch > Process Batch
 - Online
 - Edit > Validate
 - Action > Update Holdings

31

Resources

- *Anglo-American Cataloguing Rules* (2nd ed.). Chicago: American Library Association, 2002.
- *Cataloging Electronic Resources: OCLC-MARC Coding Guidelines*
<http://www.oclc.org/support/documentation/worldcat/cataloging/electronicresources/>

32

Resources

- *OCLC Bibliographic Formats and Standards* (3rd ed.)
<http://www.oclc.org/bibformats/en/default.shtm>
- *OCLC Connexion Browser Tutorials*
<http://www.oclc.org/support/training/connexion/browser/tutorial/>
- *OCLC Connexion Client Tutorials*
<http://www.oclc.org/support/training/connexion/client/tutorial/default.htm>

33