



Hennepin County Library

M E M O

TOMORROW'S LIBRARY... TODAY

DATE: March 5, 2002  
TO: Hennepin County Library Staff  
FROM: Charles Brown, Library Director  
SUBJECT: Changes in the Hennepin County Library Catalog

Hennepin County Library (HCL) has long been a leader in using the library catalog to provide exceptional access to library resources. Over the past thirty years, HCL has led the way in cataloging all formats - print as well as AV; applying subject headings to works of fiction; adopting new subject headings for emerging concepts; changing outdated terminology; and providing summary notes and content notes in the bibliographic record.

Thanks, in part, to the dedication of the HCL cataloging staff and their willingness to act as advocates for this approach, these cataloging practices are increasingly becoming the norm throughout the library field.

However, things change. Some of our past practices regarding the selective use of accepted cataloging standards now severely limit our ability to use emerging technology - technology that can expand our users' access to the full range of HCL's rich resources and the ever-increasing range of global information resources as well.

For example, we are unable to import the catalog records that are available for government documents, eBooks, and eJournals into our existing catalog because of the inconsistencies in our subject headings and in our use of identifying record tags. Combining new records with our non-standard data for similar types of information results in a clutter of conflicting terms that can be confusing for the user. As technology advances, the current HCL catalog will become increasingly isolated from web-based resources elsewhere and the new information formats our users want and expect from their library.

Fortunately, we have the opportunity to preserve some of the benefits of the HCL catalog while adopting more standardized cataloging practices. Over the next 6 to 9 months we will replace our current catalog database with records that conform to internationally accepted cataloging standards. We also will maintain some of the value-added features of HCL's existing catalog.

We will adopt the standard use of MARC tags and the AACR2 record format. We also will adopt Library of Congress subject headings. HCL cataloging information that adds value (unique subject heading concepts, notes, etc.) will be placed in keyword searchable fields accessible to library catalog users.

By making this change now, we will be able to adopt new service improvements quickly. We will be able to offer faster and more convenient access to such valuable parts of our collection as government documents and web resources. And we will be able to provide our users with more information about our holdings in reference and periodicals - both onsite and via the Internet.

We also will be in a position to offer users the option to initiate an integrated search of multiple databases from a single access point - the HCL catalog. Users will have greater access to cataloging information via personal digital assistants and similar, growing technologies. And we will be better positioned to adapt to future technological advances.

More detailed information about the project is posted on the Staff Web under "Catalog Project." Over the next weeks and months, there will be opportunities for staff to discuss the technical aspects of these changes in more detail. Staff also will have an opportunity to identify training needs to be addressed.

While I acknowledge that this change may be difficult for some, the potential benefits for our users, both current and future, are immeasurable. Thank you, in advance, for your cooperation and support of this important step in maintaining HCL's leadership role in making library resources available and inviting to the broadest possible audience.

# Preparing for the 21<sup>st</sup> Century Catalog: Data Standardization Project

## GENERAL PROJECT INFORMATION

In today's modern library, collections contain many formats—print, audiovisual, and more recently, remote information accessed via the web. Hennepin County Library's (HCL) goal is to provide users with a web-based catalog that is a single point of entry to the library's resources and allows data to be easily exchanged. In order to make the catalog truly web-based, to have a tool with which the user can connect to multiple databases and sources of information, HCL must replace its current catalog database with records that are in internationally used standardized forms. We will continue to adopt cataloging practices that significantly reduce customized bibliographic records. A new authority file will be created that replaces the existing one. It will use subject headings based primarily on Library of Congress Subject Headings (LCSH) and name and series headings that are structured using Anglo-American Cataloging Rules (AACR2). Current HCL forms will be moved to keyword searchable fields and made available to those users who are accustomed to HCL forms.

The Web is about interoperability of disparate computer systems and standards are what make these systems interoperable. By fully adopting international standards (AACR2, MARC (Machine Readable Cataloging), LCSH, etc.) we can take advantage of the computing power at our disposal and host a catalog that is web-based and meets the needs of our user population. AACR2 determines how information is structured and displayed. MARC is about how data is stored and transferred, and the Library of Congress Subject Headings (LCSH) is a shared body of authorized terms used to describe the content of library materials.

HCL is working with OCLC to upgrade our bibliographic records to meet MARC, AACR2 and LCSH standards. We have also contracted with a MARC database consultant who is assessing our existing database, helping to prepare our records for conversion and participating in the project's success. In September 2002, HCL will have an online catalog, providing good, timely, user-centered access to our traditional collections and will set the stage to incorporate an array of exciting new resources.

Our catalog, services, and programs will be responsive to a global population that increasingly relies on the web to communicate, study, shop and be entertained. HCL's catalog will be a gateway to and member of that web community.

# Preparing for the 21<sup>st</sup> Century Catalog: Data Standardization Project

## WHY

HCL has reached a watershed. The choices made in the past to selectively use national standards have put us in an untenable position. We are stalled in our ability to make expanding library collections accessible to users. Without this action, the Library's strategic directions and service objectives are unattainable.

To move forward and succeed we need to be able to collaborate with others. We need the ability to accept standard MARC, LCSH and AACR2 records from others for services such as the delivery of added copy titles, building Opening Day Collections and adding eBooks, web sites, and government documents to the HCL Catalog. We need to be positioned to add content and value such as cover graphics, reviews and links to author websites. We need to allow users to search across multiple databases (e.g. HCL Catalog and General Reference Center Gold) or across multiple library catalogs (e.g. MnLINK) and we need to provide options for accessing the catalog via new devices such as personal digital assistants.

This project is about standardizing our data. All integrated library systems (Dynix now, Horizon in the near future) rely on national and international standards. Our current catalog records were developed in an isolation that worked for that time. Today we need records that operate in an online, standards-based environment in use by the library world. This project is necessary regardless of current or future hardware or software platforms. It is about organizing our data elements in an internationally recognized structure for the benefit of our users. Arriving at the decision to make this change has been an ongoing discovery process that began in the 1980's.

HCL's data structure problems are multidimensional and interconnected, especially with HCL subject headings. For example, some of HCL's existing subject headings are Library of Congress Subject Headings (LCSH); some headings differ from LCSH in spelling, synonymous terms, word order, structure (especially subdivisions) and specificity; some headings reflect concepts not covered by LCSH. Making the distinction among headings that are LCSH, LCSH variations, and unique concepts is impossible because of the way the database was developed. Since we cannot tell what we have, there is no way to standardize our current records without a complete replacement. Because we do not know what percentage of our headings are variant or unique, we cannot predict the degree of change the user will experience.

## **Preparing for the 21<sup>st</sup> Century Catalog: Data Standardization Project**

These data structure problems present a situation where staff edit incoming records so that the form of the LCSH heading conforms to the HCL form, if one is found. This situation also means that the data cannot be reasonably maintained. We cannot take advantage of many features that would enhance customer service. Currently, it is not possible to batch load MARC records for government documents, eBooks, and eJournals. Combining standard and nonstandard records creates a clutter of conflicting terms and a confused user. Because we cannot load standard records, we cannot incorporate and maintain bibliographic records for websites in our catalog or have accurate links within the bibliographic record.

Today, access is about much more than subject headings. Its a matter of interaction of data, data structure and software to enable searching, retrieval and display. If the data is inaccurate or absent, access is impeded. If the data is not structured (tagged) correctly, access is impeded. If the software is flawed or lacks functionality, access is impeded. If data is not keyword searchable, access is impeded. Access is also about being able to search multiple databases at once, and if multiple databases are searched using different authorized terms, access is again impeded. Subject headings provide more precision in searching and keywords provide greater recall. HCL's web-based catalog must address all these elements of access in order to be successful.

### **CATALOG INFORMATION**

HCL cataloging staff will continue to enhance imported records in ways that add value for users, including summaries, content notes, and subject and genre headings. Staff will continue to provide fuller cataloging of mass-market paperbacks and books in series. Staff will also begin to load government document records, and begin to catalog web sites and other eResources. Staff will spend less time editing the LCSH/ACCR2 forms to make an imported record conform to HCL forms.

Most library systems have local headings to describe materials not covered by LCSH. HCL will too. This may be material on a local or regional topic or a concept not yet adopted by LCSH. When a library chooses to establish a local heading, it is structured (tagged) in such a way that it is identifiable as a local heading. If the Library of Congress later establishes a term for that new concept, libraries retag, and if necessary revise the heading to conform to the established term. In that way, everyone "speaks the same language."

## Preparing for the 21<sup>st</sup> Century Catalog: Data Standardization Project

In September new bibliographic and authority records will be loaded into Dynix and replace existing ones. Associated item information will be attached to the replacement records. The replacement bibliographic records will:

- Contain more contents and summary notes. Contents notes may be formatted in such a way that contents could be indexed and made searchable by title. This is useful especially for songbooks and short story anthologies.
- Add missing information such as ISBN numbers, physical descriptions, and fixed field coding used to refine indexes and provide search limits.
- Have a more consistent look, using common punctuation and abbreviations. All the records will look more like our recently added bib records.
- Contain name and series headings that are structured using AACR2. Author authorities will contain death dates only if such a date is needed to distinguish one author from another.
- Contain subject headings that are based primarily on LCSH. Local headings may be introduced after the replacement records are loaded into Dynix if they are tagged and maintained properly and they do not create conflicts with the LCSH structure.
- Move existing HCL forms that are not valid LCSH headings to a keyword searchable field where they will be available to those users who are accustomed to the HCL forms.
- Contain the original call number, Dynix bib number, copyright date, and pieces notes.

### COMMUNICATION

Charles Brown, Janet Leick, Sharon Charles and the UCS Steering Team will ensure that there is regular and ongoing communication, information and training.

Information and training are critical to the success of this project. In order to ensure that project information is available to staff, questions are answered and training is planned and implemented, each member of the UCS Steering Team has taken a lead role. Margaret Gillespie, Lois Langer, Sandy Louis and Michael McConnell are responsible for project information and updates. Mark Ranum is responsible for the training component. Sharon Charles and Elizabeth Feinberg will lead the technical plan and development of a catalog project communication website.

The purpose of the project website is to provide a neutral place for staff to ask questions to build knowledge and understanding. It will provide a two-way communication tool. It will also be a place

## Preparing for the 21<sup>st</sup> Century Catalog: Data Standardization Project

for staff to find information about the project. The web site is currently under construction and will be available more fully in the next few weeks. Staff are encouraged to participate and communicate feedback, questions and training needs to their direct supervisor, resource group manager/section head or via the project website. Staff participation will be needed in the development of a Training Plan and the Bibliographic Services Plan, which will document cataloging and catalog procedures.

### TRAINING

Staff will be provided with the tools and knowledge they need to provide excellent service and be proficient in their work. This may include building a better understanding of indexing and access points in an online environment, using alternative search methods, identifying and using other helpful reference tools, and understanding basic MARC, AACR2 and LCSH. In the coming weeks, training needs will be identified based on staff input and a plan will be developed. Training plans for cataloging staff have already begun. Several cataloging staff will be attending a variety of cataloging classes over the course of the next 8 months. Courses include Minnesota Opportunities for Technical Service Excellence (MOTSE) sessions on "Access", "LCSH", "eBooks", "Websites", and "eJournals". In addition, all Cataloging Staff are scheduled to attend a four day cataloging course "Book Blitz, or How to Make MARC Records that Really Work." to be conducted by Deborah Fritz, author of *Cataloging with AACR2 and USMARC*.

### SUMMARY

HCL wants the user to have access to a world of information and resources via the catalog.

In a web-based catalog, increased search functionality, improved sortability, more effective contents notes searching, and broadcast searching across multiple databases make an expanding collection more accessible to users. In the world of electronic resources, the catalog descriptions and the resources are linked. Find an eBook from netLibrary in the catalog and two "clicks" later, read the eBook.

Making this change now requires great effort. The long-term benefits will outweigh short-term losses. The underlying philosophy of cataloging endures: make our current and future collections accessible to users as we strike a balance between adding user-centered value to our bibliographic data and the costs and time required for cataloging. HCL will use best practices for cataloging excellence as we weave our collection into the web and weave the web into our catalog.

# Preparing for the 21<sup>st</sup> Century Catalog: Data Standardization Project

## TIMELINE

A general timeline is provided, describing the various stages of the project.

### **March-April 2002: Project Planning**

- Project announcement
- Technical plan finalized.
- Training plans initiated.
- Database assessment and record preparation
- Bibliographic Services Plan (which documents cataloging procedures) is initiated.

### **May 2002**

- Cataloging staff training
- New cataloging practices following AACR2, MARC, LCSH and any local practices that have been defined are implemented.

### **June-August 2002**

- Bibliographic database sent to OCLC for match process and authority control processing.
- Clean up areas identified and addressed
- Prep work with Dynix to load records.
- Staff training

### **September 2002**

- New bibliographic records are imported into Dynix
- New authority records are imported into Dynix
- Search indexes are rebuilt
- Testing for conversion to Horizon begins.



Note: This plan was prepared using the best information available to date. There are some aspects of the plan that may change as work with OCLC, Epixtech, and a technical consultant continue.

## **PART 1 - HCL BIBLIOGRAPHIC RECORD PRE-PROCESSING AND EXPORT**

Our database is exported in meaningful and appropriate batches in order to perform any record pre-processing. The purpose of the pre-processing would be to:

- improve matching with WorldCat records
- make data to be retained more uniform
- make corrections to fixed field elements
- identify problem records

After pre-processing, the outgoing records are sent to OCLC for bibliographic record processing (Part 2) and authority controlled tag processing (Part 3).

## **PART 2 - BIBLIOGRAPHIC RECORD PROCESSING**

One by one, OCLC compares the HCL records to the millions of bibliographic records in WorldCat. If a match is found, the WorldCat record is saved to a new file and will eventually be imported into HCL's Dynix system and will be used to replace the original HCL bibliographic record.

Data from some MARC tags (if present) will be copied from the original HCL record to the matching WorldCat record so that they can be retained in the final replacement record. See **Appendix A** for a list of tags, descriptions and notes pertaining to the retention of data from our original HCL records.

If no match is found, the original HCL record is included in the replacement batch and sent through Authority Controlled Tag Processing (see Part 3, *Authority Controlled Tag Processing*).

## **PART 3 - AUTHORITY CONTROLLED TAG PROCESSING**

Once the new set of replacement bibliographic records has been created, they are processed for authority control.

One by one, each of the replacement bibliographic records are examined for authority controlled data elements. When data in an authority controlled tag is encountered, the data in the tag is processed according to the table in **Appendix B**.

Each bibliographic record is checked for duplicate authority controlled tags. If the same heading is used multiple times in the same bibliographic record, only one is retained in the final replacement record.

The process will also search for authority records from the OCLC database to match all of the authority controlled headings in the final replacement records. If an authority record can be found, it is saved and will be sent to HCL to be imported into Dynix in Part 4, *Introducing the Records into the HCL Dynix System*. If a matching authority record cannot be found, a skeletal authority record will be created and will also be sent to HCL for import.

### **PART 4 - INTRODUCING THE RECORDS INTO THE HCL DYNIX SYSTEM**

With the new bibliographic and corresponding authority records in hand, the task of importing the new database records begins.

- Step 1: The HCL authority files are completely removed from the Dynix system
- Step 2: The new **authority** records are imported into Dynix. *This process will take many hours to complete.*
- Step 3: The new **bibliographic** records are imported into Dynix. *This process will take many hours to complete.* The replacement bibliographic records automatically replace their original counterparts, leaving all associated item records unaffected. Dynix will automatically connect the bibliographic records with the authority records imported in step 2.
- Step 4: All search indexes are rebuilt. *This process will take many hours to complete.*

Appendix A: Bibliographic Data Retention

2/27/02

Original HCL Tag	Repl. Record Tag	Description	Note
035	035	Dynix BIB number (local)	The final replacement record must retain the original's 035 in order to correctly match during record re-introduction. Any 035 data in the base WorldCat record must be removed to ensure that only the original HCL data is present in the final replacement record's 035.
092	092	HCL call number (local)	The final replacement record must retain the original's 092. Any 092 data in the base WorldCat record must be removed to ensure that only original HCL data is possible in the final replacement record's 092.
440	594	Series title	Series title is copied from the original's 440 to the final replacement record's 594. Any data in the 594 tag in the base WorldCat record must be removed to ensure that only original HCL 440 data is possible in the replacement's 594.
505	505	Contents note	<b>Conditional:</b> If the original HCL record has no 505, do nothing. If the original HCL record has a 505 and the base WorldCat record doesn't, copy the original's 505 into the final replacement record. If both records have a 505, preference goes to the formatted 505. If the formatting is the same in both 505s, preference goes to the base WorldCat record's 505.
520	520	Summary note	<b>Conditional:</b> If the original HCL record has no 520, do nothing. If the original HCL record has a 520 and the base WorldCat record doesn't, copy the original's 520 into the final replacement record. If both records have a 520, preference goes to the base WorldCat record's 520.
591	591	Copyright date (local)	The final replacement record must retain the original's 591, if present. Any 591 data in the base WorldCat record must be removed to ensure that only original HCL data is possible in the final replacement record's 591.
593	593	Pieces note (local)	The final replacement record must retain the original's 593, if present. Any 593 data in the base WorldCat record must be removed to ensure that only original HCL data is possible in the final replacement record's 593.
600	600_4	Subject: Personal name	All subject access fields in the original HCL record should have their second indicators changed (if necessary) to "4" and included in the final replacement record. This will flag them for further examination and processing in the <u>Authority Controlled Tag Processing</u> step of the project.
610	610_4	Subject: Corporate name	
611	611_4	Subject: Conference name	
630	630_4	Subject: Uniform title	
650	650_4	Subject: Topical term	
651	651_4	Subject: Geographic name	

## Appendix A: Bibliographic Data Retention

2/27/02

700	700	Added entry: Personal name	All controlled added entry data in the original HCL record should be copied to the final replacement record. Since these tags are authority controlled, they will be examined and processed in the <u>Authority Controlled Tag Processing</u> step of the project.
710	710	Added entry: Corporate name	
711	711	Added entry: Conference name	
730	730	Added entry: Uniform title	
	947	Processing Note	<p>A note should be added to each final replacement record processed to indicate the result of the match process (\$a) and the date of processing (\$b).</p> <p>Subfield \$a should contain one of the following character strings:</p> <p style="padding-left: 40px;"><b>unique-key</b> <b>extended</b> <b>ddr</b> <b>no match</b></p> <p>The date included in subfield \$b should follow the pattern YYYYMMDD. Example: 20020214 for February 14, 2002.</p>

## Appendix B: Authority Controlled Tag Processing

2/27/02

Authority Controlled Tag	Description	Source	Processing Note														
100	Personal name	WorldCat	Main entry headings can only come from the base WorldCat record. The OCLC authority database is checked for available updates. Headings with no matching OCLC authority record are retained as tagged.*														
110	Corp. name	WorldCat															
111	Conf. name	WorldCat															
130	Uniform title	WorldCat															
440	Series title	WorldCat	Series titles can only come from the base WorldCat record. The OCLC authority database is checked for available updates. Headings with no matching OCLC authority record are retained as tagged.*														
800, 810, 811, 830	Series title	WorldCat															
600_0	Personal name	WorldCat	Subject access fields in the record with a second indicator of 0 can only come from the base WorldCat record. The OCLC authority database is checked for available updates. Headings with no matching OCLC authority record are retained as tagged.														
610_0	Corporate name	WorldCat															
611_0	Conference name	WorldCat															
630_0	Uniform title	WorldCat															
650_0	Topical term	WorldCat															
651_0	Geographic name	WorldCat															
600_4	Personal name	HCL	Subject access fields in the record with a second indicator of 4 can only come from the original HCL record. The OCLC authority database is searched for an exact match (from beginning to end). Headings which match an OCLC authority record are retained in the replacement record and the second indicator is changed from a 4 to a 0. Any available updates are applied to the heading. Headings with no matching OCLC authority record are retained in the final bibliographic replacement record, but re-tagged as follows:														
610_4	Corporate name	HCL															
611_4	Conference name	HCL															
630_4	Uniform title	HCL															
650_4	Topical term	HCL															
651_4	Geographic name	HCL															
		<table border="0"> <thead> <tr> <th><u>Original Tag</u></th> <th><u>Moved to...</u></th> </tr> </thead> <tbody> <tr> <td>600_4</td> <td>690</td> </tr> <tr> <td>610_4</td> <td>691</td> </tr> <tr> <td>611_4</td> <td>692</td> </tr> <tr> <td>630_4</td> <td>693</td> </tr> <tr> <td>650_4</td> <td>695</td> </tr> <tr> <td>651_4</td> <td>696</td> </tr> </tbody> </table>		<u>Original Tag</u>	<u>Moved to...</u>	600_4	690	610_4	691	611_4	692	630_4	693	650_4	695	651_4	696
<u>Original Tag</u>	<u>Moved to...</u>																
600_4	690																
610_4	691																
611_4	692																
630_4	693																
650_4	695																
651_4	696																

700	Personal name	WorldCat or HCL	<p>Added entry headings can come from either the base WorldCat record or the original HCL record. The OCLC authority database is searched for an exact match. Headings which match an OCLC authority record are retained in the final replacement record and any available updates are applied. Headings with no matching OCLC authority record are retained in the final replacement bibliographic record, but re-tagged as follows:</p> <table border="0"> <thead> <tr> <th><u>Original Tag</u></th> <th><u>Moves to...</u></th> </tr> </thead> <tbody> <tr> <td>700</td> <td>790</td> </tr> <tr> <td>710</td> <td>791</td> </tr> <tr> <td>711</td> <td>792</td> </tr> <tr> <td>730</td> <td>793</td> </tr> </tbody> </table>	<u>Original Tag</u>	<u>Moves to...</u>	700	790	710	791	711	792	730	793
<u>Original Tag</u>	<u>Moves to...</u>												
700	790												
710	791												
711	792												
730	793												
710	Corp. name	WorldCat or HCL											
711	Conf. name	WorldCat or HCL											
730	Uniform title	WorldCat or HCL											

\* It is no longer true that Main Entry and Series Title headings can only come from the base WorldCat record. Original HCL records that don't match a WorldCat record will be sent through the Authority Controlled Tag Process (this is a recent development) and will contain Main Entry and Series Title headings. The process for these two types of headings will be refined as project development work continues with OCLC. The goal will be to validate and include as much data from the original HCL record as possible, while minimizing real or potential authority record conflicts.