

Suggested Priorities for Bibliographic, Holding, and Item Record Maintenance

Originally authored by the Consortial Cataloging and Authority Control Committee
(Fall, 2005)

Revisions and new projects added by the I-Share Cataloging and Authority Control Team
(December 2009 through August 2011)

Newest projects added November 23, 2011

As of May 2011, an online version of this document is available from:
<http://www.carli.illinois.edu/mem-prod/I-Share/cat/maintpriority.html>

Introduction

Voyager's pre-packaged reports, the I-Share shared SQL and shared macro web sites, and other resources offer a number of tools to help maintain our individual library databases. I-Share library staff have expressed an interest in a prioritized list of database cleanup activities so that they may focus on most urgent needs first. The following list is the I-Share Cataloging and Authority Control Team's consensus on this question. The list is arranged in four categories: general good practice, then projects that should be done frequently, those that should be done occasionally, and those that should only need to be done once. Within each section, projects are listed in order of importance.

The projects listed here are believed to be of general interest to a number of libraries. It is recognized that individual libraries may have unique needs for projects based on their particular history, and are encouraged to adopt their own priorities, as well as develop queries and the corresponding maintenance procedures as needed. Libraries are responsible for correcting records as appropriate in their local Voyager database.

This document is based on queries and macros that are available at the time of writing. Because new queries and macros are regularly added to the shared SQL and Shared Macros pages, the list will expand from time to time.

For the fall 2009 revision of this document, ICAT adopted a set of "conceptual priorities" to help both the team in their assignments of priority order, and the user of this document decide on the relative importance of one project over another. These conceptual priorities have as their core the goal of "reducing patron annoyance," and are described below.

#1 Delivery. These projects correct data problems that prevent efficient charging at the circulation desk, the placing of requests, or the efficient processing of requested items -- the patron has already identified the item she needs, and wants to get out of the door with the materials in hand with minimal difficulty at the circ desk. These projects are in the "Frequently" category, and at the top of that category.

#2 Locating. These projects correct data problems that prevent the patron from finding the item, already identified as desired, within the library or on the web, in the case of e-resources. These projects are in the “Frequently” category, but follow the Delivery-type of projects.

#3 Discovery. These projects correct data problems that will help the patron’s searching activities. These projects are in the “Occasionally” category, in priority order by these subcategories:

(3A) Data problems that prevent overall access in the local database, the I-Share Union Catalog, and/or other public interfaces (such as VuFind).

(3B) Adding/correcting access points to bibs in this order: titles, subjects, authors, control numbers, then “other,” including data elements such as language and format used for sorting/limiting.

(3C) Eliminating duplicate records in local databases or in the I-Share Union Catalog.

(3D) Enhancements to description or projects to bring records up to MARC standard coding, when the MARC coding errors don’t cause problems with a higher priority above.

#4 Legacy system data problems, that in theory once fixed should not recur. These projects are in the “Once” category, and use the same conceptual priorities as above.

Projects added in May 2011 are labeled as “<NEW 05/2011>”; projects added in August 2011 are labeled as “<NEW 08/2011>”; projects added in November 2011 are labeled as “<NEW 11/2011>”.

For the December 2010 revision, an appendix was added to provide estimated run times for each of the SQL queries cited in this document. Note that there are several factors that affect the time it takes an Access query to run, so the timings in this appendix are just estimates. But it is hoped that they will help library staff with their implementation of the various maintenance projects described in this document.

Also for the December 2010 revision, some of the existing Shared SQL queries were revised or split into multiple queries to make them more efficient to run. The descriptions of these queries were revised in this document, and they are labeled “(REV. 12/2010).” Please note that the changes to these queries were made after the CARLI_reports_2009.mdb file was issued, so staff will need to copy the revised (and sometimes new) queries from the Shared SQL page to take advantage of the new queries.

Many of these projects are based on queries that can be found in the Local Catalog Maintenance section of the I-Share “Shared SQL for Reports in Voyager” web page <<http://www.carli.illinois.edu/mem-prod/I-Share/secure/sql.html>>. This webpage is secured, but should be able to be accessed without restrictions from I-Share campus IP addresses. To access this page from off-campus, enter your institutional affiliation and personal Voyager patron record barcode number at the prompts.

To run these queries on your personal workstation, you will need to install MS Access and the appropriate version of ODBC drivers. Instructions for configuring ODBC as well as an installation program are available from <<http://www.carli.illinois.edu/mem-prod/I-Share/secure/execute.html#oracle10g>>.

Your local information technology staff may be able to help with ODBC installation. If catalogers are not able to run these queries personally, other library staff who use MS Access may be able to run them for you; or, as a last resort, you may ask CARLI staff to run the queries for you by having your library’s I-Share Liaison submit a Work Request Online (WRO).

Some of the projects can be greatly facilitated through the use of Macro Express macros to correct the records using the Voyager cataloging client. Like the Shared SQL web page, there is also an I-Share Shared Macro web page at <<http://www.carli.illinois.edu/mem-prod/I-Share/secure/macros.html>>. This page is also secured. It includes instructions for downloading and using the Macro Express software, as well as postings of individual macros for a variety of cataloging tasks. All of the macros described in this document can be linked to from the above URL, and navigating to the “Macros for the Voyager Cataloging Client” section, then navigating to the appropriate type of record that each macro edits.

Some of the projects can be facilitated by using Voyager’s Pick and Scan functionality. More information about using Pick and Scan is available in the *Voyager 7.2 Cataloging User’s Guide*, chapter 7.

Some of the projects suggest that downloading fresh copies of bibs from OCLC may be an effective fix. One technique for doing this using the Connexion client’s batch processing options is described at <http://www.carli.illinois.edu/mem-serv/mem-train/090430cat/090430JW_RepairorTrade.pdf>.

Project Table of Contents

General good practice	6
G1. Use all of Voyager’s validation options in the cataloging client.....	6
G2. Consider implementation of OCLC’s Bibliographic Notification service.....	6
Projects to do frequently	6
F1. Eliminate duplicate item barcodes.....	6
F2. Correct barcodes that do not contain the right number of digits, or add active barcodes to item records that lack them.....	7
F3. Eliminate “error” item type code.....	7
F4. Revise item record variable fields containing double quote character.....	7
F5. Resolve cases where item types are not consistent with an item’s location.....	8
F6. Resolve conflicts when item records have permanent locations different from the location in the MFHD.....	8
F7. Evaluate bibliographic records without MFHDs.....	9
F8. Supply missing call numbers in MFHDs with items attached.....	9
F9. Perform link checking and maintenance on URLs.....	9
F10. Correct MFHDs missing a call number prefix for a specific location.....	10
F11. Evaluate MFHD call number prefixes for typos or miscoded data.....	10
F12. Correct MFHDs missing a call number suffix for a specific location.....	11
F13. Evaluate Notes fields in Holding records.....	11
Projects to do occasionally	12
OC1. Correct bib records that are unsuccessfully indexed by Voyager.....	12
OC2. Evaluate suppressed bibliographic records with items attached.....	12
OC3. Evaluate suppressed MFHDs with items attached.....	13
OC4. Correct errors in bib record field tags for control numbers. <NEW 08/2011>.....	13
OC5. Correct invalid indicator values in bibliographic records.....	16
OC6. Eliminate multiple OCLC numbers in a single bib.....	17
OC7. Delete 035 \$a in format (XXXdb)nnnnnn from bib records.....	17
OC8. Resolve name, title (series), and subject heading problems identified by “Voyager Cat 6: See Refs with linked bib records report”.....	17
OC9. Correct typographical errors.....	18
OC10. Add an alternate title field (generally 246) to bib records as needed.....	18
OC11. Correct bibliographic records with more than one 245 field.....	18
OC12. Add uniform titles to bib records that lack them. <NEW 11/2011>.....	19
OC13. Correct subject headings in bibliographic records as instructed in the Cataloging Service Bulletin List of Revised LC Subject Headings.....	21
OC14. Supply subject access to autobiographies that lack it.....	21
OC15. Identify bib records that have MESH (and other) subject headings but lack LC subject headings.....	21
OC16. Correct geographic subject headings coded as topical headings <NEW 08/2011>.....	22
OC17. Update name headings using the Closed Dates in Authority Records lists.....	23
OC18. Correct MFHDs with unindexed call numbers.....	24
OC19. Evaluate serial bib records without an ISSN.....	24
OC20. Correct multiple ISBNs found in the same 020 field.....	24
OC21. Evaluate bib records with an ISSN but the format is not for serials.....	25

OC22. Correct errors in bib record format codes.....	25
OC23. Correct errors in bib record fixed field Language codes.	27
OC24. Correct bib records without 260 subfield c but date information is present elsewhere in the 260.	28
OC25. Correct bib records with 260 subfield c date information using letter EL instead of digit 1.....	28
OC26. Correct errors in bib record fixed field Date codes.	29
OC27. Correct errors in bib record fixed field Government Publication codes.....	31
OC28. Evaluate item records with copy number zero.	38
OC29. Evaluate MFHD records without a specific copy number.....	39
OC30. Eliminate duplicate OCLC numbers.	39
OC31. Reconcile superseded OCLC numbers.	39
OC32. Supply full description for Cataloging in Publication level bibliographic records.	40
OC33. Evaluate MFHDs with 'OK to Export' box checked.	40
OC34. Correct MFHD record type codes.	41
OC35. Correct bib record fixed field Literary Form codes.....	41
OC36. Correct invalid or obsolete bib record fixed field Publication status codes. <NEW 05/2011>	43
OC37. Correct bib record Leader Record Status codes.....	49
OC38. Correct bibliographic records with invalid use of code m in 008 Nature of Contents. <NEW 05/2011>	49
Projects to do once.....	51
ON1. Correct item barcodes that do not belong to your library.	51
ON2. Evaluate DRA- assigned barcodes beginning 38888.	51
ON3. Correct the "653 problem."	51
ON4. Replace "MARCettes" with full cataloging.	52
ON5. Correct ISBN numbers with fewer than 10 digits.	52
ON6. Evaluate 028 field data for miscoded corporate authors.....	52
Appendix: Estimated Run Times for the SQL Queries.....	53

General good practice

G1. Use all of Voyager's validation options in the cataloging client.

Turning on (i.e., not bypassing) the validation functions in Options → Preferences → Validation tab in the cataloging client will help assure that headings, MARC coding, and ISBN/ISSNs are valid for records as they are added to or updated within your database.

G2. Consider implementation of OCLC's Bibliographic Notification service.

This service allows libraries to replace bibs in their catalogs with new versions when certain improvements have been made to the master record in OCLC. For further details, see <http://www.carli.illinois.edu/mem-prod/I-Share/cat/OCLC_bibnotif.html>.

Projects to do frequently

F1. Eliminate duplicate item barcodes.

Resolving these problems is very important for circulation and interlibrary loan service. Duplicated item barcodes force staff to select from a list of items when a barcode is scanned. This is particularly problematic for items sent to fill UB requests, because the item information available to the borrowing library is limited. Duplicated item barcodes can also result in the wrong item being used to fill a UB request.

To help prevent future instances of this problem, it is strongly recommended that all cataloging and circulation operators enable the client preference to “Check for Duplicate Item Barcodes.”

(F1.A) Pre-Packaged report: “Duplicate Item Barcodes” finds cases where the same barcode has been used on more than one item.

(F1.B) Shared SQL (Item Records: Barcode category): “Find holding and item data for a given barcode” is a query that helps with the resolution of problems found by the “Duplicate Item Barcodes” pre-packaged report. This query prompts for an individual barcode number, and outputs the data needed to identify the item records with identical barcodes. This query may be most helpful when there are only a few duplicated barcodes found in a database.

<NEW 08/2011> (F1.C) Shared SQL (Item Records: Barcode category): “Find holding and item data for all duplicated barcodes” is another query designed to help with the resolution of problems found by the “Duplicate Item Barcodes” pre-packaged report. This query uses the Duplicate Item Barcodes query as a table, and then outputs the data needed to identify all of the item records with identical barcodes. This query may be most helpful when there are many duplicated barcodes found in a database.

F2. Correct barcodes that do not contain the right number of digits, or add active barcodes to item records that lack them.

From time to time, barcode readers (or typists) skip or repeat digits. And from time to time, staff forget to add the barcode number to the item record during processing. Finding and fixing these cases quickly will save trouble at the circulation desk and ILL office.

(F2.A) Shared SQL (Item Records: Barcode category): “Item barcodes with invalid length.” The query prompts for the correct number of digits for your library’s item barcodes.

(F2.B) Shared SQL (Item Records: Barcode category): “Items with no Active barcodes.” The query finds item records with no barcode at all, or only barcodes with a status of inactive. Of course, some items without barcodes are legitimate (e.g., rare book room materials or other non-circulating collections), but this query may find item records where an active barcode should be added.

F3. Eliminate “error” item type code.

“Error” item types could have been created during the library’s data conversion to Voyager, or from ongoing bulk import jobs where the incoming data does not match a “real” entry in the mapping table. Items with the error item type code often do not generate the proper due date when charged, because many libraries profile them to have restricted loan policies. Finding and fixing these cases will save trouble at the circulation desk and ILL office.

Shared SQL (Item Records: General category): “Find “error” item type code”

F4. Revise item record variable fields containing double quote character.

There is a newly discovered bug in WebVoyage Classic, whereby the presence of a double quote character (i.e., “) in an item record variable field will prevent the UB request form from populating correctly in the I-Share Union Catalog. This means that

patrons from any library cannot place requests on any titles in the UC that have a double quote character in the Enumeration, Chronology, Year, Caption or Freetext fields of the item record, no matter which library has that data in their item record.

Shared SQL (Item Records: General category): “Item records containing double quote.” This query finds item records linked to bibs that are not OPAC suppressed that contain the double quote character in any of the item record variable fields.

Shared macro (Item Record Maintenance section): “Changeitem_doublequote.mex.” If you have too many of the item records to fix manually, this macro can help with cleanup.

F5. Resolve cases where item types are not consistent with an item’s location.

Items with this problem may be circulated inappropriately or for incorrect periods of time.

(F5.A) Pre-Packaged report: “Item Count by Location and Type.” Running this report will show the number of occurrences of each combination of permanent location and item type. From this, you can quickly see cases where an item type and permanent location combination may be incorrect, such as a circulating reference book.

(F5.B) Shared SQL (Item Records: General category): “Items with specific combinations of perm loc and item type” is a query designed to be used in conjunction with the “Item Count by Location and Type” pre-packaged report, and will show you the specific items with questionable combinations, which can then be investigated and corrected.

F6. Resolve conflicts when item records have permanent locations different from the location in the MFHD.

Because location information is independently coded in the MFHD and in the item, sometimes errors are committed that result in conflicting location information. A common underlying cause of this discrepancy is the use of Voyager Pick and Scan to change the item record’s permanent location without also setting the option to change the holding (MFHD) location. It is strongly recommended to change both locations when using this tool.

WebVoyage version 7 makes this data discrepancy much more noticeable to patrons than previous versions of that public interface.

Shared SQL (Item Records: General category): “Item records with perm loc different from the MFHD loc.” Revised summer 2009 to include the item barcode, so that Pick and Scan can be used as part of the fix.

Pick and Scan has functionality to edit both the item and MFHD location. Pick and Scan requires the item record to contain a barcode, so not all appropriate records may be fixable with this tool.

F7. Evaluate bibliographic records without MFHDs.

These may be “orphaned” bib records left over from incomplete withdrawal or bibs created for ordering that have not had orders placed for them. Bibs without holdings can be confusing and misleading to patrons and staff. Note: Some bib records in Voyager cannot be deleted if they are associated with Acquisitions POs. In these cases, you may need to suppress the bib record instead of delete it.

Shared SQL (Bibliographic Records: General category): “Bibliographic records without MFHDs”

Shared macro (Bibliographic Record Maintenance section): “Del_bibs.mex.” If it is determined that bib records identified by this query can and should be deleted, this macro can help with cleanup.

F8. Supply missing call numbers in MFHDs with items attached.

If catalogers forget to create a call number, or miscode it in various ways, users will see no call number in the public catalog.

Shared SQL (MFHD Records: Call Numbers category): “Missing call numbers in MFHDs with items attached”

F9. Perform link checking and maintenance on URLs.

For electronic resources, URLs function similarly to a physical item’s call number. Invalid URLs make the electronic resource unavailable to users via the OPAC.

Each I-Share library is encouraged to evaluate their local requirements for more extensive link checking than is described in this project, and to initiate the best possible solution for their institution.

(F9.A) Shared SQL (Bibliographic Records: General category): “Bibliographic or MFHD records with typos at beginning of URL”. This query finds 856 fields in bibs

and/or MFHDs that have typographical errors at the beginning of the field, which makes the URL inaccessible.

(F9.B) Requestable via WRO: a re-run of “MFHD_866withURL” which will identify specific holdings records that have at least one 866 field that contains the text string “http”. URLs in 866 fields are treated as a note, and are not hyperlinked to provide access to the online resource for patrons. The URL must be stored in the MFHD record’s 856 field in order to be hotlinked. In May 2010, this server-side report was run for all I-Share libraries; at that time, only 16 of the I-Share databases had any of these “errors,” so it is possible that a re-run of this report will not find any records in your database.

F10. Correct MFHDs missing a call number prefix for a specific location.

Some libraries use a prefix for all items in a specific location. If the prefix is missing, users may have difficulty knowing where to find an item.

Shared SQL (MFHD Records: Call Numbers category): “MFHDs missing call number prefix for a specific location”

Shared macro (MFHD (Holdings Record) Maintenance section): “Addmfhd_852k.mex.” In cases where the prefix needs to be added, this macro is available to help.

F11. Evaluate MFHD call number prefixes for typos or miscoded data.

Some libraries use a prefix for all items in a specific location. Sometimes the prefix contains a typographical error or is added to the call number proper (852 \$h) instead of the subfield designated for prefixes (852 \$k). There is a suite of queries designed to help identify the MFHDs with these types of errors. Be sure to read the description of each query on the Shared SQL page, for more information about the intention of each query.

(F11.A) Shared SQL (MFHD Records: Call Numbers category): “Odd Prefixes Query 1 – List All Prefixes”. This make-table query is designed for use by the other queries in the suite. It creates a table of all MFHDs that have an 852 \$k. This query may take up to several hours to run.

(F11.B) Shared SQL (MFHD Records: Call Numbers category): “Odd Prefixes Query 2 – Rare Prefixes”. This make-table query is designed for use by the other queries in the suite. It creates a table of all MFHDs from Odd Prefixes Query 1 where the prefix is found in 20 or fewer MFHDs. This query is designed to identify potential typographical errors in prefixes, and is used by the other queries in the suite.

(F11.C) Shared SQL (MFHD Records: Call Numbers category): “Odd Prefixes Query 3 – MFHDs with Rare 852\$k”. This query compares the tables created by the previous two queries and produces a list of the MFHD IDs that contain the “rare” prefixes found by query 2. These query results can be used to manually correct typographical errors found in 852 \$k prefixes.

Shared macro (MFHD (Holdings Record) Maintenance section): “Changemfhd_852k.mex.” In cases where the existing prefix needs to be revised, and there are too many records to process manually, this macro is available to help.

(F11.D) Shared SQL (MFHD Records: Call Numbers category): “Odd Prefixes Query 4 – Prefixes not in 852\$k”. This query takes the list of all prefixes found by query 1 and looks for “normalized” call numbers that begin with the prefix. A common error is to put the prefix in 852 \$h instead of 852 \$k.

F12. Correct MFHDs missing a call number suffix for a specific location.

Some libraries use a suffix for all items in a specific location. If the suffix is missing, users may have difficulty knowing where to find an item.

Shared SQL (MFHD Records: Call Numbers category): “MFHDs missing call number suffix for a specific location”

Shared macro (MFHD (Holdings Record) Maintenance section): “Addmfhd_852m.mex.” In cases where the suffix needs to be added, this macro is available to help.

F13. Evaluate Notes fields in Holding records.

Many libraries add public notes to their holdings records, and consistency in the content of these notes can be important to help patrons locate materials in the library. Non-public notes may be useful to staff for processing purposes.

(F13.A) Shared SQL (MFHD Records: General category): “Finding Public and Non-Public Notes in Holdings”.

This is a series of three queries. Query 1 prompts for a location code, and query 2 looks up the 852 fields from the MFHDs in the specified location. The output from query 3 contains holdings records in that location that have any data in 852 \$x and/or \$z, so the contents of the existing notes can be reviewed.

(F13.B) Shared SQL (MFHD Records: General category): “MFHDs lacking a specific public note”.

This is a series of three queries. Query 1 prompts for a location code, and query 2 looks up the 852 fields from the MFHDs in the specified location. Query 3 prompts for the beginning text of what should be the valid public note for that location. The query results identify holding records in the desired location that have either no 852\$z or that lack the valid note text. The 852 \$x is also included in the results, in case some public notes were miscoded as \$x.

Shared macro (MFHD (Holdings Record) Maintenance section):
“Addmfhd_852z.mex.” If the fix is determined to be to add a new public note (852 \$z) to the holding record, this macro can add a note with the same text to all records in the macro’s input file.

Projects to do occasionally

OC1. Correct bib records that are unsuccessfully indexed by Voyager.

If there is an error in the MARC coding of the bib record, or a subfield with no data or only punctuation in it, Voyager may create an index entry with no data in it, and users will not be able to search the entry.

(OC1.A) Shared SQL (Bibliographic Records: Titles category): “Bibs without indexed titles”. This query looks at the bib record’s Title index. Some bibs are incomplete and lack a 245 field. Some bibs have errors in subfielding or other MARC coding. Some bibs have a diacritic or special character among the characters that would be unindexed because of the non-filing indicator, which seems to confuse Voyager.

(OC1.B) Shared SQL (Bibliographic Records: General category): “Identifying Bibs with empty index entries.” This query looks at all indexed fields in the bib record and finds cases where the index is present but contains no data.

(OC1.C) Requestable via WRO: a re-run of “Bad tag list” which will identify specific bib records that contain invalid field tags. Most of the fields with invalid tags will not be indexed properly by Voyager. In December 2010, this server-side report was run for all I-Share libraries; at that time, approximately one-third of the databases had any of these errors, so it is possible that a re-run of this report will not find any errors in your database.

OC2. Evaluate suppressed bibliographic records with items attached.

In cases where a library suppresses its bib records during the order process, occasionally the cataloger may forget to unsuppress the bib at the time of cataloging.

If suppressed bibs with items attached are not intended, it is important to correct this problem quickly so users can find items in the catalog.

Shared SQL (Bibliographic Records: General category): “Suppressed bibliographic records with items attached”. Revised query output (fall 2009) includes item status, item barcode, and location/call number data, to help make processing the query results more efficient.

Shared macro (Bibliographic Record Maintenance section):
“Changebib_OPACsuppress.mex” is available to unsuppress bibs in batches based on results of the query above.

Pick and Scan (Voyager cataloging client, version 7.1.0 and higher) has functionality to suppress/unsuppress bibs. Pick and Scan requires the item record to contain a barcode, so not all appropriate records may be fixable with this tool.

OC3. Evaluate suppressed MFHDs with items attached.

In some embedded order data workflows, holding records are suppressed as well as bib records. If holdings are not un-suppressed, they will not be visible to patrons.

Shared SQL (MFHD Records: General category): “MFHDs that are suppressed from OPAC with Items attached”. Revised query output (fall 2009) includes item status, item barcode, and location/call number data, to help make processing the query results more efficient.

Shared macro (MFHD (Holdings Record) Maintenance section):
“Changemfhd_OPACsuppress.mex” is available to unsuppress MFHDs in batches based on results of the query above.

Pick and Scan (Voyager cataloging client, version 7.1.0 and higher) has functionality to suppress/unsuppress MFHDs. Pick and Scan requires the item record to contain a barcode, so not all appropriate records may be fixable with this tool.

OC4. Correct errors in bib record field tags for control numbers. <NEW 08/2011>

Some I-Share databases have errors in the MARC field tags for various control numbers. For example, a bib may contain an 035 \$a that contains the text “(pbk.)” which implies the content of the field is actually an ISBN rather than an OCLC number. Records with these types of incorrect field tags are usually not retrievable with a control number search for the valid tagging (e.g., the bib described above would not be found with an ISBN number search). In addition, these incorrect field tags may have negative consequences on de-duplication both in the local database and

the I-Share union catalog, and some errors can result in the bib being discarded from the union catalog loads, meaning that the title is not represented in that database.

There are ten individual queries for this project that look for various data elements that imply incorrectly tagged control number fields. Because these control number fields are used the I-Share union catalog's duplicate detection profiles, staff should follow the "suppress/replace routine" when correcting the bibs identified by these queries. More details about this routine are available from <http://www.carli.illinois.edu/mem-prod/I-Share/cat/safebibrep.pdf>.

Some of the queries can retrieve the same bib record, so it is assumed the queries will be run in the order listed below, and that the appropriate records will be fixed before the next query in the series is run.

(OC4.A) Shared SQL (Bibliographic Records: Control Numbers category):
"Bibliographic record 035 \$a contains pbk may be ISBN." This query looks for bib records where the 035a field contains descriptive text that is commonly found in ISBNs. It is expected that most, but not all, of the records retrieved by this query may be mis-tagged ISBNs and should have their field tag changed from 035 to 020.

(OC4.B) Shared SQL (Bibliographic Records: Control Numbers category):
"Bibliographic record 035 \$a contains double slash may be LCCN." This query looks for bib records where the 035a field contains a double slash, which is often found in LCCNs. It is expected that most, but not all, of the records retrieved by this query may be mis-tagged LCCNs and should have their field tag changed from 035 to 010.

(OC4.C) Shared SQL (Bibliographic Records: Control Numbers category):
"Bibliographic record 035 \$a contains hyphen may be LCCN." This query looks for bib records where the 035a field contains a hyphen, which is often found in LCCNs. It is expected that many, but not all, of the records retrieved by this query may be mis-tagged LCCNs and should have their field tag changed from 035 to 010.

(OC4.D) Shared SQL (Bibliographic Records: Control Numbers category):
"Bibliographic record 035 \$a contains DRA bib ID." This query looks for bib records where the 035a field contains text that matches the format of DRA bib IDs (e.g., ABC-1234). This former system control number should be coded as 035 \$9 rather than 035 \$a. Libraries that were not members of the consortium when the ILS was DRA could still have these errors in their database, if the library copies bibs from the I-Share union catalog into the local database. It is expected that most, but not all, of the records retrieved by this query may be mis-tagged DRA control numbers and should have their subfield code changed from 035 \$a to \$9. Another option is to delete the 035 field that contains the DRA control number, rather than changing the subfield coding. The former system control number was very valuable during the conversion to Voyager, but is now considered an obsolete number.

Shared macro (Bibliographic Record Maintenance section):
“Changebib_035aDRANumber.mex” is available to change the bib 035 subfield coding in batches based on results of the query above, if too many records are retrieved to edit manually.

(OC4.E) Shared SQL (Bibliographic Records: Control Numbers category):
“Bibliographic record 035 \$a contains punctuation may be barcode.” This query looks for bib records where the 035a field contains the exclamation point character. This coding was used in previous systems to indicate the number following the character was a barcode number. Libraries that were not members of the consortium prior to our implementation of Voyager could still have these errors in their database, if the library copies bibs from the I-Share union catalog into the local database. It is expected that most, but not all, of the records retrieved by this query should have the 035 field that represents a barcode deleted from the bib.

(OC4.F) Shared SQL (Bibliographic Records: Control Numbers category):
“Bibliographic record 010 \$a contains ocm may be OCLC number.” This query looks for bib records where the 010a field contains the text “ocm” or “ocn” which is the prefix found in OCLC numbers. It is expected that most, but not all, of the records retrieved by this query may be mis-tagged OCLC numbers and should have their field tag changed from 010 to 035.

(OC4.G) Shared SQL (Bibliographic Records: Control Numbers category):
“Bibliographic record 010 \$a contains pbk may be ISBN.” This query looks for bib records where the 010a field contains descriptive text that is commonly found in ISBNs. It is expected that most, but not all, of the records retrieved by this query may be mis-tagged ISBNs and should have their field tag changed from 010 to 020.

(OC4.H) Shared SQL (Bibliographic Records: Control Numbers category):
“Bibliographic record 020 \$a contains ocm may be OCLC number.” This query looks for bib records where the 020a field contains the text “ocm” or “ocn” which is the prefix found in OCLC numbers. It is expected that most, but not all, of the records retrieved by this query may be mis-tagged OCLC numbers and should have their field tag changed from 020 to 035.

(OC4.I) Shared SQL (Bibliographic Records: Control Numbers category):
“Bibliographic record 020 \$a contains double slash may be LCCN.” This query looks for bib records where the 020a field contains a double slash, which is often found in LCCNs. It is expected that most, but not all, of the records retrieved by this query may be mis-tagged LCCNs and should have their field tag changed from 020 to 010.

(OC4.J) Shared SQL (Bibliographic Records: Control Numbers category):
“Bibliographic record 022 \$a not 8 digits.” This query looks for bib records where the 022a field is not in the format of NNNN-NNNN or NNNN-NNNX (where N represents a digit). It is expected that most, but not all, of the records retrieved by this

query may be malformed ISSNs or other control numbers that are mis-tagged as an ISSN, that should be corrected as appropriate.

OC5. Correct invalid indicator values in bibliographic records.

By the MARC bibliographic standard, the first or second indicator values (when applicable per field) should only be a digit between 0 and 9. For various reasons, some bib records contain invalid indicators in various fields, including the pipe character, other punctuation, or letters. Currently in Voyager, when representing a non-filing indicator, the invalid character equates to a zero, which may or may not be the correct number of non-filing characters. Supplying the correct number in cases where zero is not correct improves title access. Replacing these bibs with fresh copies from OCLC may fix this problem as well as resolve other obsolete MARC usages and possibly provide other enhancements to access and description.

In addition, bib records with this invalid indicator are rejected by default from processes that extract and convert MARC bibs to MARCXML, such as the eXtensible Catalog project. Also, future implementations may be more sensitive to this non-standard data than is Voyager.

(OC5.A) Shared SQL (Bibliographic Records: Titles category): “Find bib records where 245 indicator 2 contains the | (pipe) character” queries (REV. 12/2010). This is a sequence of three queries that should be run in order. The output from the third query is the one that will report on the problematic records.

(OC5.B) Shared SQL (Bibliographic Records: Titles category): “Find bib records where 245 indicator 2 is blank” queries (REV. 12/2010). This is a sequence of two queries that should be run in order. The output from the second query is the one that will report on the problematic records.

(OC5.C) Requestable via WRO: a re-run of “Bad indicator count” which will provide an overview to show which bib record fields contain invalid indicators, and how many per field. This is part of an alternative to running the first query above.

(OC5.D) Requestable via WRO: a re-run of “Bad indicator list” which will identify the specific bib records that contain invalid indicators. This report will find the same bibs as the first shared SQL above, plus any pipe characters and other invalid indicator values found in any bib record fields, not just the 245 field. In December 2010, this server-side report was run for all I-Share libraries; at that time, most databases had at least some records with these errors.

Shared macro (Bibliographic Record Maintenance section):

“Changebib_0xx5xx_INDS.mex” can reset the first and/or second indicator(s) to new user-specified value(s) in bib record fields from 010 through 599.

Shared macro (Bibliographic Record Maintenance section):
“Changebib_6xx9xx_INDS.mex” can reset the first and/or second indicator(s) to new user-specified value(s) in bib record fields from 600 through 999.

OC6. Eliminate multiple OCLC numbers in a single bib.

When an individual Voyager bib contains more than one OCLC control number, the record will almost certainly be discarded from the I-Share Union Catalog. A bib record should only contain one OCLC control number.

Shared SQL (Bibliographic Records: Control Numbers category): “Individual bibs with more than one OCLC number”

OC7. Delete 035 \$a in format (XXXdb)nnnnnn from bib records.

When bibliographic records are copied from the I-Share Union Catalog to a local library database, by default the Voyager bib record ID and database code (e.g., (EIUdb)123456) from the contributing library is added as an 035 \$a to the newly copied bib. If this unwanted 035 \$a data is not manually deleted when the bib is saved to the local database, its presence when that bib is fed back into the UC can contribute to discards from the UC. Deleting all existing 035 \$a in the format (XXXdb)nnnn will reduce the number of discards from the I-Share Union Catalog.

Shared SQL (Bibliographic Records: Control Numbers category): “Bib record 035 \$a contains (XXXdb)nnnn”

Shared macro (Bibliographic Record Maintenance section):
“Delbib_XXXdb_049u.mex.”

OC8. Resolve name, title (series), and subject heading problems identified by “Voyager Cat 6: See Refs with linked bib records report”.

This report must be requested through WRO (Work Request Online). It lists headings used in bibliographic records that exist as see references in authority records. Working through this report continuously and a little at a time will result in many name, title, and subject heading corrections to bibliographic records. Headings that exist as references can result in split files for authors, subjects, or series as well as confusing OPAC displays that will make retrieval more difficult for users. More information on using Cat 6 can be found at <http://www.carli.illinois.edu/mem-prod/I-Share/secure/cat/voy_cat_reports.pdf>.

OC9. Correct typographical errors.

A list of most commonly misspelled words is available at <http://terryballard.org/typos/typoscomplete.html>. The list is organized by probability of appearance of words. Working through the list will improve retrieval of information. Catalogers may wish to initially limit their searches for the misspellings to specific search types (such as Title or Subject searches), because keyword searching for these errors may retrieve large numbers of records in some databases.

OC10. Add an alternate title field (generally 246) to bib records as needed.

If a record with an ampersand in the 245 is missing a 246, 247, or 740 with the string “and”, users may have more difficulty retrieving the record by title. Similarly, if titles have a symbol in the title without a corresponding alternate title field, the record will be difficult to retrieve.

(OC10.A) Shared SQL (Bibliographic Records: Titles category): “Ampersands in Titles”

Shared macro (Bibliographic Record Maintenance section): “Addbib_246.mex.” This macro can use the bib IDs from the “Ampersands in Titles” query above, copy the contents of 245 \$a into a new 246 \$a, and change the 246 field’s ampersand to “and”. After the macro has processed the ampersands found in 245 \$a, the “Ampersands in Titles” query should be re-run, and a determination made about the appropriateness of manually changing ampersands found in other 245 subfields to “and.”

(OC10.B) Shared SQL (Bibliographic Records: Titles category): “Symbols in Titles”

OC11. Correct bibliographic records with more than one 245 field.

According to the MARC standard, the bibliographic 245 field is non-repeatable. In the VuFind public catalog, records with multiple 245 fields display no title information at all, which can result in confusing displays and the inability to retrieve additional views of the record. Voyager will index multiple 245 fields for title searching, but it is unknown how other systems (such as the eXtensible Catalog) will index this non-standard coding.

Shared SQL (Bibliographic Records: Titles category): “Bibliographic records with multiple 245 fields”. This is a sequence of three queries that should be run in order. The output from the third query is the one that will report on the problematic records.

In sampling the results of this query in several databases, often these errors are cases where a 246 field was incorrectly coded as a 245 field.

OC12. Add uniform titles to bib records that lack them. <NEW 11/2011>

In both the VuFind and WebVoyage public catalogs, the uniform title is an access point. It is also expected that the eXtensible Catalog (XC) project will use the data in the 240 field as an important data element in its FRBRization functionality. Therefore, libraries may wish to add a uniform title to bibliographic records that should have one, per national cataloging standards.

There are five individual parts/queries for this project that look for bib records that do not currently have a uniform title but that may be candidates to add one. Some of the queries can retrieve the same bib record, so it is assumed the queries will be run in the order listed below, and that the appropriate records will be fixed before the next query in the series is run.

(OC12.A) Request a “Voyager Cat 6: See Refs with Linked Bib Records Report” and process the Name-Title entries to add a uniform title where appropriate.

This is the same report as referenced in project OC8. The Name-Title entries included in this report are often cases where the 1xx and 245 field in a bib record match a cross-reference in an authority record. As a result, the information needed to determine what the correct uniform title should be discernable from a comparison of the authority record cited in the report with the bib record(s) in question.

This report must be requested through WRO (Work Request Online) by either the library director or I-Share Liaison. When filing the WRO, if you choose the record format option for “place in my library’s directory on the reports server for local manipulation,” then the output will be saved as a pipe-delimited text file, which can be easily imported into Excel and sorted as needed locally. In this format, you can sort the output so all Name-Title entries are collected together, and then you could do a secondary sort by number of bib records, so you could focus on entries that affect multiple records first.

More information on using Cat 6 can be found at <http://www.carli.illinois.edu/mem-prod/I-Share/secure/cat/voy_cat_reports.pdf> and in particular, see the section on processing Name-Title entries, which are examples 7 through 9.

(OC12.B) Shared SQL (Bibliographic Records: Titles category): “Bib record without UT but title contains transl”. This query looks for non-serial bib records that lack a uniform title, but the full title (including 245 \$c) contains text inferring a translation, including variations on “transl*” in other major languages. It is expected that most, but not all, of the records retrieved by this query may be candidates to have a uniform title added to the bib.

The other queries in this project described below will omit bibs that include “transl*” (etc.) in the title field, so it is expected that the library will process all appropriate records found by this query before running the subsequent queries below.

In some databases, these queries may retrieve a large number of records, so this query and those that follow prompt for a range of bib IDs. Testing was successful using a range of 50,000 bib IDs at a time. For all queries, the output includes the call number from the MFHD, so staff can concentrate cleanup efforts on targeted subject areas, if desired.

(OC12.C) Shared SQL (Bibliographic Records: Titles category): “Bib record without UT but 700 field contains tr.” This query looks for non-serial bib records that lack a uniform title, and the 245 field does not include “transl*” or variations of “transl*” in other major languages, but the 700H index includes the text “tr.” which implies that a translator is an added author. It is expected that most, but not all, of the records retrieved by this query may be candidates to have a uniform title added to the bib.

(OC12.D) Shared SQL (Bibliographic Records: Titles category): “Bib record without UT but 6xx field contains Translations into.” This query looks for non-serial bib records that lack a uniform title, and the 245 field does not include “transl*” or variations of “transl*” in other major languages, but the 60x or 65x field includes the text “Translations into.” It is expected that many, but not all, of the records retrieved by this query with an index code of 600H may be candidates to have a uniform title added to the bib. It is also expected that many, but not all, of the records retrieved by this query with an index code of 6500 may represent anthologies, which may not need a uniform title added.

(OC12.E) Shared SQL (Bibliographic Records: Titles category): “Bib record without UT but 041 indicator 1 is 1.” This is a series of three queries that looks for non-serial bib records that lack a uniform title, and the 245 field does not include “transl*” or variations of “transl*” in other major languages, and the 041 field has its first indicator set to 1, implying a translation. It is expected that many, but not all, of the records retrieved by this query may be candidates to have a uniform title added to the bib.

OC13. Correct subject headings in bibliographic records as instructed in the Cataloging Service Bulletin List of Revised LC Subject Headings.

Working through this list when each issue of CSB is published will help keep subject headings current. While similar to working through the “Voyager Cat 6” report described above, working with the CSB list will also identify headings with subdivisions, which are not included in the “Voyager Cat 6” report. Paper copies of CSB are available by subscription from the Library of Congress; electronic copies are available from <<http://www.loc.gov/cds/PDFdownloads/csb/index.html>>.

In addition, Library of Congress publishes an electronic weekly list of new and changed subject headings, which has the advantage of currency. Libraries may wish to consult the weekly list <<http://www.loc.gov/aba/cataloging/subject/weeklylists/>> in addition to, or in place of, the CSB list of revised headings.

OC14. Supply subject access to autobiographies that lack it.

Correcting these problems improves subject access to autobiographical materials.

Shared SQL (Bibliographic Records: Subject Headings category): “Find autobiographies without a 600 field.” The query uses fixed field coding to determine that a bib record is an autobiography, and occasionally this coding is incorrect. Catalogers will need to examine records carefully to determine the correct course of action for cleanup (i.e., is the fixed field coding wrong, or does the bib need subject heading(s) to be added).

OC15. Identify bib records that have MESH (and other) subject headings but lack LC subject headings.

Adding LC topical and geographical subject headings to those records that lack them will improve subject access/cross-linking between records for those libraries that rely on LCSH. The queries below default to working with MeSH headings, but query 2 may be modified to find bibs with other types of subject headings. They are designed to be run in the order listed below, with query 4 providing the output to be used by cataloging staff.

(OC15.A) Shared SQL (Bibliographic Records: Subject Headings category): “Missing LC subject query 1.” This make-table query is designed for use by the other queries in the suite. It lists all bibs with LC subject headings.

(OC15.B) Shared SQL (Bibliographic Records: Subject Headings category): “Missing LC subject query 2.” This make-table query is designed for use by the other queries in the suite. It lists all bibs with MeSH subjects not found in query 1.

(OC15.C) Shared SQL (Bibliographic Records: Subject Headings category):
“Missing LC subject query 3.” This make-table query is designed for use by the other queries in the suite. It lists all subject headings from the bibs found in query 2. Note: because this query involves “the BLOB,” it may take a long time to complete.

(OC15.D) Shared SQL (Bibliographic Records: Subject Headings category):
“Missing LC subject query 4.” This query is designed to present the results to the cataloging staff member. It lists all the bibs found in query 3 that do not have any LC or LC Children’s topical or geographical subject headings.

OC16. Correct geographic subject headings coded as topical headings
<NEW 08/2011>

Prior to the implementation of VuFind, none of the former public catalogs supported for I-Share libraries made distinctions between subject headings coded as 650 versus 651 for searching or display. However, VuFind displays subject headings coded as 650 under the label of “Topics” and subject headings coded as 651 under the label of “Regions.” It is also expected that the public catalog for the eXtensible Catalog (XC) project will distinguish between geographic and topical headings for some faceting/limiting functionality.

There are two queries that can be used to identify bib records with geographic subject headings that are coded as topical subject headings (650). One is limited to headings that begin with the names of U.S. states. The other is limited to headings that begin with the names of various countries.

The queries do not make distinctions based on 65x field indicator values; by default, all 650 headings are inspected for these coding errors. If the library prefers to work only with subject heading systems they support (for example, LCSH but not MeSH), the INDEX_CODE value in each query can be edited to limit to the desired subject system. Using this example, editing the INDEX_CODE from “650*” to “6500” would limit the query to only LCSH headings. Similarly, editing the INDEX_CODE to “6502” would limit results to only MeSH headings. The final digit in the query’s 650 INDEX_CODE value matches the second indicator value in bib 650 fields.

The queries also include additional text to omit false matches, such as valid 650 headings that begin with a country name (e.g., Vietnam War). Due to MS Access field length limitations, not all applicable terms can be used as criteria, so there will likely be some false hits retrieved by both queries. Catalogers will need to use their judgment as to which records retrieved by each query should be edited to change the 650 coding to 651. Checking the authority file for questionable headings retrieved by the queries may be useful in this decision-making.

In some databases, these queries may retrieve a large number of records, so both of the queries prompt for a range of bib IDs; testing was done using a range of 50,000 bib IDs at a time.

(OC16.A) Shared SQL (Bibliographic Records: Subject Headings category):
“Bibliographic record 650 index begins with State Name”

This query will find bibliographic records where a subject heading is coded as 650 and the heading begins with the name of a U.S. state. Additional limiting criteria are included to help reduce the number of false hits. It is expected that that the fix for most, but not all, of the records retrieved with this query should be to change the field tag from 650 to 651.

IMPORTANT NOTE: This query relies on a table called StateNames that must be manually imported before the queries can be run. Instructions for creating the StateNames table are available as Example 1 from http://www.carli.illinois.edu/mem-prod/I-Share/secure/sql/importing_table.pdf. This step only needs to be completed once per Access mdb file. If, for example, you previously create the StateNames table for the maintenance project regarding fixed field government publication codes, you don't have to repeat that step for this query.

(OC16.B) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic record 650 index begins with Country Name”

This query will find bibliographic records where a subject heading is coded as 650 and the heading begins with the name of a country. Additional limiting criteria are included to help reduce the number of false hits. It is expected that that the fix for most, but not all, of the records retrieved with this query should be to change the field tag from 650 to 651.

IMPORTANT NOTE: This query relies on a table called CountryNames that must be manually imported before the queries can be run. Instructions for creating the CountryNames table are available as Example 2 from http://www.carli.illinois.edu/mem-prod/I-Share/secure/sql/importing_table.pdf. This step only needs to be completed once per Access mdb file.

OC17. Update name headings using the Closed Dates in Authority Records lists.

The Library of Congress changed their policy as of Feb. 1, 2006, about adding death dates to personal name headings. Lists for each week since the new policy went into effect are available at <http://www.oclc.org/us/en/rss/feeds/authorityrecords/default.htm> and as an RSS feed. Many libraries had received patron comments about the catalog being out of date due to the previous practice of not adding death dates, and therefore actively

asked for this policy change. Often, headings are re-evaluated during the process of reconciling versions of the name, and other elements of the heading may be changed as well.

One reason to use the LC lists is that the revised authority records do not always contain a reference from the name form with the open date, and therefore would not appear on the Voyager Cat 6 report. For those libraries using Gary Strawn's Cataloger's Toolkit, the software creates a "virtual" cross reference, thereby significantly reducing the effort to create batch changes, which will correct both name and subject usage of a particular name heading in the catalog.

OC18. Correct MFHDs with unindexed call numbers.

A MFHD without a call number is not necessarily an error (e.g., for e-resources). But records that have a call number in the 852 \$h but that do not have an 852 first indicator value are not indexed for MFHD call number searches. Cleaning these up will provide better call number searching to both users and library staff.

Shared SQL (MFHD Records: Call Numbers category): "Unindexed Call Numbers"

Shared macro (MFHD (Holdings Record) Maintenance section): "Addmfhd_852ind1.mex." If you have many unindexed call numbers, this macro can help with cleanup. Be sure to create a separate input file for each different new 852 indicator 1 value (e.g., one input file for all MFHD IDs that need an Ind1 value of 8 (for Other), and a separate input file for all MFHDs IDs that need an Ind1 value of 0 (for LC classification).

OC19. Evaluate serial bib records without an ISSN.

The ISSN is an important searching and matching mechanism. Supplying ISSNs whenever possible will enhance the functionality of link resolver software (e.g., SFX).

Shared SQL (Bibliographic Records: Control Numbers category): "Serial bib records without an ISSN"

OC20. Correct multiple ISBNs found in the same 020 field.

Current MARC standards dictate one \$a per 020 field, and to repeat 020 when more than one ISBN is present. Voyager only indexes the first \$a when there are multiple \$a's in a single 020 field. Correcting this problem enhances retrieval and matching functionality when using the ISBN. Records with this problem are generally NOT yet corrected in OCLC, so manual editing must be done. If libraries catalog new items or

update existing bibs with this problem, MARC validation will identify 020 \$a as unrepeatable. However, if MARC validation is turned off, new cases may be introduced.

Shared SQL (Bibliographic Records: Control Numbers category): “Find 020’s with multiple \$a’s.”

OC21. Evaluate bib records with an ISSN but the format is not for serials.

A miscoded format field may adversely impact retrieval in both the OPAC and staff clients (e.g., Journal title searching relies on bib format coding).

Shared SQL (Bibliographic Records: Control Numbers category): “Bib records with an ISSN but format is not for serials”

Shared macro (Bibliographic Record Maintenance section):
“Changebib_Format.mex.” This macro will change the bib record’s format coding as specified by the user. However, this macro should be used ONLY if the number of bibs to be edited is too many to process manually. Changing the format code in the cataloging client can invalidate other fixed field data, and therefore is best done manually.

OC22. Correct errors in bib record format codes.

Fixed field record type and bib level codes are combined into “format” in Voyager tables. The format is used for display purposes and limiting of searches in both the WebVoyage and VuFind public interfaces. However, the display of format is more pronounced in VuFind than WebVoyage, so this project may have higher importance for libraries that use VuFind as their public catalog interface.

There are several queries that can be used to identify bib records with various types of format errors. It is assumed the queries will be run in the order listed below.

(OC22.A) Shared SQL (Bibliographic Records: Fixed Fields category):
“Bibliographic format counts”. This query counts all combinations of bib level and record type values found in all bib records in the database. This is useful for statistics and can suggest formats to spot check for possible cleanup projects.

(OC22.B) Shared SQL (Bibliographic Records: Fixed Fields category):
“Bibliographic records with specific format codes.” This query returns a list of bib IDs for a specific bib level/record type combination entered at the prompt, designed as a follow-up to the query above. It may be useful for identifying specific bib records that may have unexpected format data, for possible cleanup projects.

(OC22.C) Shared SQL (Bibliographic Records: Fixed Fields category): “Bibliographic Records with invalid or obsolete format coding.” This query looks for records where the combination of the Leader’s Type of Record and Bibliographic Level contains blank, invalid, or obsolete values. Depending on which Leader byte contains the error, records with invalid format codes may display an incorrect material type icon in VuFind. They will often also lack a format facet in VuFind and may not be retrieved in WebVoyage searches that limit by format. In addition, records with this problem may not have a valid fixed field language code for the language facet in VuFind or language limits in WebVoyage.

In the cataloging client, records with invalid format coding often do not display any data in the 008 field. Invalid Type of Record and/or Bibliographic Level codes must be corrected before the 008 can be edited, else an error will occur in the cataloging client when these records are edited.

(OC22.D) Shared SQL (Bibliographic Records: Fixed Fields category): “Bibs with video GMD but record type is not ‘g’.” This query looks for a video or motion picture GMD in the 245 field, but the format code does not begin with “g”. Usually, the records retrieved have problematic format codes.

(OC22.E) Shared SQL (Bibliographic Records: Fixed Fields category): “Bibs with sound recording GMD but record type is not ‘i’ or ‘j’.” This query looks for a sound recording GMD in the 245 field, but the format code does not begin with “i” or “j”. Usually, the records retrieved have problematic format codes.

Shared macro (Bibliographic Record Maintenance section): “Changebib_Format.mex.” This macro will change the bib format to new values specified by the user. However, this macro should ONLY be used if there are too many records to be fixed manually, because changing the format code can reset other fixed field values to an invalid code.

(OC22.F) Shared SQL (Bibliographic Records: Fixed Fields category): “Bibs with format beginning ‘g’ but no video GMD.” This query looks for bib records with missing, obsolete, or incorrect GMDs, and potential incorrect format codes may also be retrieved. Bibs with very long 245 fields may have valid GMDs retrieved, due to field indexing limitations in Voyager.

(OC22.G) Shared SQL (Bibliographic Records: Fixed Fields category): “Bibs with format beginning ‘i’ or ‘j’ but no sound recording GMD.” This query looks for bib records with missing, obsolete, or incorrect GMDs, and potential incorrect format codes may also be retrieved. Bibs with very long 245 fields may have valid GMDs retrieved, due to field indexing limitations in Voyager.

OC23. Correct errors in bib record fixed field Language codes.

Fixed field language (LANG) codes are used for display purposes and limiting of searches in both the WebVoyage and VuFind public interfaces. However, the display of language is more pronounced in VuFind than WebVoyage, so this project may have higher importance for libraries that use VuFind as their public catalog interface.

There are several queries that can be used to identify bib records with various types of language errors. It is assumed the queries will be run in the order listed below.

(OC23.A) Shared SQL (Bibliographic Records: Fixed Fields category):

“Bibliographic records with Lang code set to ‘und’.” “Und” is a valid language code, but often the query retrieves records that could be edited to contain a “real” LANG code. This is especially true for “MARCettes” from the consortium’s LCS days.

(OC23.B) Shared SQL (Bibliographic Records: Fixed Fields category):

“Bibliographic records with Lang code set to ‘zxx’.” “Zxx” is a new valid language code, meaning “no linguistic content.” The results of this query should be spot-checked for bibs that could be edited to contain a more appropriate LANG code. It is expected that a high number of records retrieved with this query will not require correction.

(OC23.C) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib records with Lang code set to obsolete or invalid value.” This query looks for bib records with the LANG code set to “n/a”, null, pipe character, or two letter codes.

(OC23.D) Shared SQL (Bibliographic Records: Fixed Fields category):

“Bibliographic records with Lang code not common.” This query looks for bib records with the LANG code set to values OTHER than “eng”, “fre”, “ger”, “ita”, “lat”, “rus”, “spa”, and “zxx.” The query will retrieve many valid less common codes, and also will retrieve typos in common codes. Spot-check the retrieval set for bibs that represent obvious errors in LANG code. It is expected that a high number of records retrieved with this query will not require correction.

Shared macro (Bibliographic Record Maintenance section):

“Changebib_LANG.mex.” This macro can be used with the results of the queries above and will change the bib LANG code to a new value specified by the user. There is some labor involved in creating/sorting the input files for the macro run(s) because each individual input file must be for the same bib format code and desired new language code.

OC24. Correct bib records without 260 subfield c but date information is present elsewhere in the 260.

If the date is not found in 260 \$c, the publication date may not display to users. In addition, correcting this error will give better results when running queries that rely on the publication date field, such as those described in the next maintenance project.

(OC24.A) Shared SQL (Bibliographic Records: General category): “Bibliographic records with null 260c, but 260b contains a date”

(OC24.B) Shared SQL (Bibliographic Records: General category): “Bibliographic records with null 260c, but 260a contains a date”

Shared macro (Bibliographic Record Maintenance section):

“Changebib_260parseC.mex.” This macro can be used with the output from both queries above. The macro will parse date information found in 260 \$a or 260 \$b into a new 260 \$c.

(OC24.C) Shared SQL (Bibliographic Records: General category): “Bib records with null 260c, for scores and sound recordings only”

Shared macro (Bibliographic Record Maintenance section):

“Changebib_260dtoC.mex.” This macro can be used with the output from the query above. The macro will change the subfield code from obsolete 260 \$d to \$c.

(OC24.D) Shared SQL (Bibliographic Records: Fixed Fields category):

“Bibliographic records with null 260c but 008 DATE 1 is not null.” This query looks for bib records where the 260 \$c is null, but the fixed field DATE1 is not null, not “uuuu”, not “||||” and the bib format is not for serials. The bibs found by this query will need to be fixed manually, often by adding a new 260 \$c based on DATE1, but sometimes the error is in 260 subfield coding. This query should be run after the bibs found by the previous queries dealing with 260 \$c have been corrected.

OC25. Correct bib records with 260 subfield c date information using letter EL instead of digit 1.

Date information in 260 \$c containing the letter EL instead of the digit 1 may be confusing to users. In addition, if future public catalogs use the 260 \$c for sorting, bib records with this typographical error will not sort properly by date. Also, fixing the bibs with this error may improve results from queries that include publication date information.

(OC25.A) Shared SQL (Bibliographic Records: General category): “Bibliographic records with 260c beginning with letter EL”

Shared macro (Bibliographic Record Maintenance section):

“Changebib_260cELdate.mex.” This macro can be used with the output from the query above. The macro will change the letter EL to digit 1 when found at the beginning of the 260 \$c.

(OC25.B) Shared SQL (Bibliographic Records: General category): “Bibliographic records with 260c containing the letter EL”

This query should be run after bibs found by the query above have been corrected. This query will look for dates containing the letter EL instead of digit 1 anywhere in the 260 \$c. Some perfectly valid 260 \$c dates may be retrieved by this query (e.g., April 1994). But some records may be found that have the publisher information incorrectly coded in 260 \$c instead of 260 \$b.

OC26. Correct errors in bib record fixed field Date codes.

Fixed field Date1 values are used for limiting of searches and sorting search results in both the WebVoyage and VuFind public interfaces. However, the use of date for sorting is more pronounced in VuFind than WebVoyage, so this project may have higher importance for libraries that use VuFind as their public catalog interface.

There are several queries that can be used to identify bib records with various types of fixed field Date errors. It is assumed the queries will be run in the order listed below. The queries that compare the Date1 with 260 \$c will also give better results if the maintenance project described above is completed before this set of queries is run.

(OC26.A) Shared SQL (Bibliographic Records: Fixed Fields category):

“Bibliographic records with null 008 DATE1, but 260c is not null.” This query looks for bib records where the 260 \$c is not null, is not [n.d.] and variations, but the fixed field DATE1 is null and the bib format is not for serials.

(OC26.B) Shared SQL (Bibliographic Records: Fixed Fields category): “Bibs with 008 DATE1 like ‘uuuu’ or ‘||||’, but 260c is not null.” This query looks for bib records where the 260 \$c is not null, is not [n.d.] and is not “uuuu”, but the fixed field DATE1 is “uuuu” or “||||.”

(OC26.C) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib records with 008 DATE1 may contain invalid “u” data.” This query looks for bib records where the 260 \$c is not null, does not contain “u” or “?” or “-“, and the fixed field DATE1 contains at least one “u”. Identifies bib records that could be edited to include

more specific DATE1 data and for which 260\$c contains more specific information. This query may retrieve some valid DATE1 data.

<NEW 05/2011> (OC26.D) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib records with 008 DATE1 less than 4 chars.” This query looks for bib records where the fixed field DATE1 contains less than four characters.

<NEW 05/2011> (OC26.E) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib records with 008 DATE2 less than 4 chars.” This query looks for bib records where the fixed field DATE2 contains less than four characters, omitting records with the publication status code of “e”. While WebVoyage and VuFind do not currently use the DATE2 data for any functionality, it is expected that the eXtensible Catalog will use this data element for some faceting/limiting, so some catalogers may wish to prepare now.

This query may retrieve a larger than expected number of records, due to changes made to the definition of some codes in the fixed field’s Publication status (a.k.a., Date type status). Prior to MARC record changes adopted in 1995, the 008 Publication status code of “d” was used to represent a “detailed date” for books and visual materials, in combination with a DATE2 value of two or four characters. This coding was frequently used for technical reports. The currently valid Publication status value for this type of detailed date is the code “e”. Currently, the Publication status code of “d” means “continuing resource ceased publication.” Many, but not all, records retrieved with this query should be corrected by changing the Publication status code from “d” to “e”, when the DATE2 represents a Month and optionally also a day value, in the format MMDD.

<NEW> Shared macro (Bibliographic Record Maintenance section):
“Changebib_008pubstatus.mex.” This macro can be used with the output from the query above, when the proper fix is to change the publication status value to “e” (for Detailed date), which allows the DATE2 value to consist of less than 4 characters. The macro will edit the bib record’s Publication Status code to a new value specified by the user.

(OC26.F) Shared SQL (Bibliographic Records: Fixed Fields category):
“Bibliographic records where 008 DATE1 does not equal 260c.” This query looks for bib records where the fixed field DATE1 is not null and does not contain a “u” character, and the DATE1 is not found in some format in the 260 \$c. Identifies bib records with inconsistencies between 260\$c and DATE1 data. This query will likely retrieve a large number of records. It should be run after all other date scenarios (008 DATE1 and 260 \$c) have been fixed. Sometimes the error found by this query will be in DATE1 and sometimes the error will be in 260 \$c.

Shared macro (Bibliographic Record Maintenance section):
“Changebib_008Date1.mex.” This macro can be used with the output from the

queries above, when the fix is to change the fixed field DATE1 value. The macro will edit the bib record's DATE1 value to a new date specified by the user.

OC27. Correct errors in bib record fixed field Government Publication codes.

In March 2011, a government publication limit was added to the VuFind advanced search screen. The limit options are to: Include gov pubs (selected by default), Exclude gov pubs, or Gov Pubs only.

The VuFind limit looks at the bibliographic record's fixed field (MARC 008/28) gov't pub codes of a, c, f, i, l, m, o, s, u, or z. Bibs that have any one of these codes will be considered to be a government publication. None of the former or current public catalogs supported for I-Share libraries have previously used this fixed field value for searching or display, so it is expected that there will be many records found by some of the queries for this project, especially for libraries that are currently or have been government depositories. This project may have higher importance for libraries that use VuFind as their public catalog interface.

This project consists of two series of queries. Queries in the "1" series below (one sub-letter for each valid 008 gov't pub code, projects OC27.A through K) are designed to look for bibs that have a specific gov't pub code, but that could represent records that are mis-coded as a gov't pub. Queries in the "2" series below (with sub-letters for most of the valid 008 gov't pub codes, projects OC27.L through X) are designed to look for bibs that with a blank gov't pub code, but that could represent a gov't pub. It is expected that more records will be retrieved by queries in the "2" series than by the "1" series.

The VuFind interface doesn't distinguish between the different types of documents, so for example, a federal document with code f and a state document with code s are both considered a gov't pub for VuFind limiting purposes. It is unknown at this time if future public catalogs such as the eXtensible Catalog will make distinctions based on the specific type of gov't pub. Therefore, libraries that undertake this maintenance project are encouraged to use the specific coding as appropriate for the work as much as possible, as opposed to assigning a generic gov't pub code such as "o."

Due to MS Access field length limitations, not all appropriate terms can be used as criteria, so there will likely be some false hits retrieved by the queries in both series. Catalogers will need to use their judgment as to which records retrieved by each query should be edited to fix the 008 gov't pub code.

Some of the queries in both series, but especially the "2" series, can retrieve the same bib record, so it is assumed the queries will be run in the order listed below, and that the appropriate records will be fixed before the next query in the series is run. A

shared macro is available to edit the targeted records to a user-specified new gov't pub code (see details at the end of this project description).

Some of the queries may retrieve a large number of records, so all of the queries prompt for a range of bib IDs; testing was done using a range of 50,000 bib IDs at a time. In addition, all of the queries use criteria that limit to the bib formats for which the gov't pub code is applicable, per the MARC standard.

IMPORTANT NOTE 1: several of the queries in both series that deal with State government publications rely on a table called StateNames that must be manually imported before the queries can be run. These queries also consist of a subquery and a main query, to be able to pull from the table properly. Instructions for creating the StateNames table are available from <http://www.carli.illinois.edu/mem-prod/I-Share/secure/sql/importing_table.pdf>. This step only needs to be completed once per Access mdb file. If, for example, you create the StateNames table for query 1B, you don't have to repeat that step for any subsequent queries that use that table.

IMPORTANT NOTE 2: I-Share libraries will need to make an informed decision about the fixed field coding of state university press publications before they begin editing their bib records for this maintenance project. Please see the document entitled "ICAT Statement on State University Press Publications and Maintenance Projects" <www.carli.illinois.edu/mem-prod/I-Share/cat/state_UP_pubs.html> for details about this local decision.

Libraries that choose to follow current national standards should run the query described in project OC27.T and NOT run the query described in project OC27.C.

Libraries that elect to not treat state university press publications as state documents should run the query described in project OC27.C and NOT run the query described in project OC27.T.

(OC27.A) Shared SQL (Bibliographic Records: Fixed Fields category): "Bib 008 govpub query 1A code is f." This query looks for bib records where the 074 field (GPO item number) is null, but the fixed field gov't pub code is f. Additional limiting criteria are included in the Author and Publisher fields to help reduce the number of false hits. It is expected that some, but not all, of the records retrieved by this query may not be federal/national documents, and so the fixed field element may need to be changed to a more appropriate code or changed to blank.

(OC27.B) Shared SQL (Bibliographic Records: Fixed Fields category): "Bib 008 govpub query 1B code is s." This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov't pub code is s, and the publisher field does not contain an entry from the StateNames table. Additional limiting criteria are included in the Author and Publisher fields to help reduce the number of false hits. It is expected that some, but not all, of the records retrieved by this query

may not be state/provincial/territorial documents, and so the fixed field element may need to be changed to a more appropriate code or changed to blank.

(OC27.C) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 1C code is s PUB maybe UP.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is s, and the Publisher field contains text that could imply a university press, with additional limiting text to reduce false hits and to also omit some known private university presses. It is expected that some, but not all, of the records retrieved by this query may be university press publications that are coded as state documents.

NOTE: Due to MS Access field length limitations, not all appropriate terms can be used as limiting criteria, so there will likely be some false hits retrieved by the query.

NOTE 2: By the MARC standard, university press publications from public institutions are technically considered state documents. But there is concern among some catalogers that coding state university press materials as state documents could have negative consequences on VuFind advanced searches that utilize the government publication limit. I-Share libraries need to make a decision about how they wish to handle this before they undertake this maintenance project. ICAT’s statement about this decision is available at <http://www.carli.illinois.edu/mem-prod/I-Share/cat/state_UP_pubs.html>.

Libraries that opt to follow national standards should NOT run this query, but instead should run the query described in project OC27.T.

But those libraries that decide to code state university press publications as “blank” can use this query to identify university press publications that are currently coded as “s” so that the 008/28 can be edited to “blank.” These libraries should NOT also run the query described in project OC27.T.

(OC27.D) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 1D code is L.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is l (the letter EL). Additional limiting criteria are included in the Author and Publisher fields to help reduce the number of false hits. It is expected that some, but not all, of the records retrieved by this query may not be local government documents, and so the fixed field element may need to be changed to a more appropriate code or changed to blank.

(OC27.E) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 1E code is c.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is c. Additional limiting criteria are included in the Author and Publisher fields to help reduce the number of false hits. It is expected that some, but not all, of the records retrieved by this query may not be multilocal government documents, and so the fixed field element may need to be changed to a more appropriate code or changed to blank.

(OC27.F) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 1F code is i.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is i. Additional limiting criteria are included in the Author and Publisher fields to help reduce the number of false hits. It is expected that some, but not all, of the records retrieved by this query may not be international intergovernmental documents, and so the fixed field element may need to be changed to a more appropriate code or changed to blank.

(OC27.G) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 1G code is m.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is m. Additional limiting criteria are included in the Publisher field to help reduce the number of false hits. It is expected that some, but not all, of the records retrieved by this query may not be multistate documents, and so the fixed field element may need to be changed to a more appropriate code or changed to blank.

(OC27.H) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 1H code is a.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is a. Additional limiting criteria are included in the Publisher field to help reduce the number of false hits. It is expected that some, but not all, of the records retrieved by this query may not be autonomous or semi-autonomous component documents, and so the fixed field element may need to be changed to a more appropriate code or changed to blank.

(OC27.I) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 1I code is o.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is o. Additional limiting criteria are included in the Publisher field to help reduce the number of false hits. It is expected that some, but not all, of the records retrieved by this query may not be government publication level-undetermined documents, and so the fixed field element may need to be changed to a more appropriate code or changed to blank.

(OC27.J) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 1J code is z.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is z. It is expected that some, but not all, of the records retrieved by this query may not be Other government publications, and so the fixed field element may need to be changed to a more appropriate code or changed to blank.

(OC27.K) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 1K code is u.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is u. It is expected that some, but not all, of the records retrieved by this query may not be unknown if item is

a government publication, and so the fixed field element may need to be changed to a more appropriate code or changed to blank.

(OC27.L) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 1L code is invalid.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is not a valid code per the MARC standard. It is expected that all of the records retrieved by this query should be edited so that the fixed field gov’t pub element is updated to a valid code or changed to blank.

(OC27.M) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 2A code not f and GPOnum not null.” This query looks for bib records where the 074 field (GPO item number) is not null, and the fixed field gov’t pub code is not f. It is expected that most, but not all, of the records retrieved by this query may be federal/national documents, and so the fixed field element may need to be changed to code f, or to a more appropriate gov’t pub code.

(OC27.N) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 2B code is blank by loc code.” Libraries that do not have a specific location code for their documents will not find this query useful. This query prompts for a location code, and then looks for linked bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is either blank or the pipe character. It is expected that most, but not all, of the records retrieved by this query will represent government documents, and so the fixed field element should be changed to the appropriate gov’t pub code.

(OC27.O) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 2C code is blank and call no SuDoc.” Libraries that do not use SuDoc call numbers for their documents will not find this query useful. This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is either blank or the pipe character, and the linked MFHD contains a SuDoc call number type. It is expected that most, but not all, of the records retrieved by this query may be federal/national documents, and so the fixed field element may need to be changed to code f, or to a more appropriate gov’t pub code.

(OC27.P) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 2D code is blank could be f PUB.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is either blank or the pipe character, and the Publisher field contains text that could imply a federal document, with additional limiting text to reduce false hits. It is expected that many, but not all, of the records retrieved by this query may be federal/national documents, and so the fixed field element may need to be changed to code f, or to a more appropriate gov’t pub code.

(OC27.Q) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 2E code is blank could be f AUTH.” This query looks for bib records

where the 074 field (GPO item number) is null, and the fixed field gov't pub code is either blank or the pipe character, and the Author field contains text that could imply a federal document, with additional limiting text to reduce false hits. It is expected that some, but not all, of the records retrieved by this query may be federal/national documents, and so the fixed field element may need to be changed to code f, or to a more appropriate gov't pub code.

(OC27.R) Shared SQL (Bibliographic Records: Fixed Fields category): "Bib 008 govpub query 2F code is blank could be s PUB not UP." This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov't pub code is either blank or the pipe character, and the Publisher field contains text that could imply a state document, with additional limiting text to reduce false hits and to also omit University Press publications. It is expected that many, but not all, of the records retrieved by this query may be state documents, and so the fixed field element may need to be changed to code s, or to a more appropriate gov't pub code.

NOTE: By the MARC standard university presses from public institutions are technically considered state documents. But some libraries may not want to spend the time distinguishing between public and private university presses, so this project is designed to remove most of that ambiguity and omit all university press publications from the query results.

(OC27.S) Shared SQL (Bibliographic Records: Fixed Fields category): "Bib 008 govpub query 2G code is blank could be s AUTH not UP." This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov't pub code is either blank or the pipe character, and the Author field contains text that could imply a state document, with additional limiting text to reduce false hits and to also omit University Press publications. Because this query looks at the Author field in conjunction with the StateNames table, there will be some false hits for authors with first or last names that are also the name of a state (e.g., first name of Virginia). It is expected that some, but not all, of the records retrieved by this query may be state documents, and so the fixed field element may need to be changed to code s, or to a more appropriate gov't pub code.

NOTE: By the MARC standard university presses from public institutions are technically considered state documents. But some libraries may not want to spend the time distinguishing between public and private university presses, so this project is designed to remove most of that ambiguity and omit all university press publications from the query results.

(OC27.T) Shared SQL (Bibliographic Records: Fixed Fields category): "Bib 008 govpub query 2H code blank could be s PUB maybe UP." This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov't pub code is either blank or the pipe character, and the Publisher field contains text that could imply a university press, with additional limiting text to reduce false hits and to also omit some known private university presses. It is expected that some, but not all,

of the records retrieved by this query may be state documents, and so the fixed field element may need to be changed to code s, or to a more appropriate gov't pub code.

NOTE: Due to MS Access field length limitations, not all appropriate terms can be used as limiting criteria, so there will likely be some false hits retrieved by the query.

NOTE 2: By the MARC standard, university press publications from public institutions are technically considered state documents. But there is concern among some catalogers that coding state university press materials as state documents could have negative consequences on VuFind advanced searches that utilize the government publication limit. I-Share libraries need to make a decision about how they wish to handle this before they undertake this maintenance project. ICAT's statement about this decision is available at <http://www.carli.illinois.edu/mem-prod/I-Share/cat/state_UP_pubs.html>.

Libraries that opt to follow national standards should run this query, and should NOT run the query described in project OC27.C.

(OC27.U) Shared SQL (Bibliographic Records: Fixed Fields category): "Bib 008 govpub query 2I code is blank could be L." This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov't pub code is either blank or the pipe character, and the Publisher field contains text that could imply a local document, with additional limiting text to reduce false hits. It is expected that some, but not all, of the records retrieved by this query may be local documents, and so the fixed field element may need to be changed to code l, or to a more appropriate gov't pub code.

(OC27.V) Shared SQL (Bibliographic Records: Fixed Fields category): "Bib 008 govpub query 2J code is blank could be c." This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov't pub code is either blank or the pipe character, and the Publisher field contains text that could imply a multilocal document, with additional limiting text to reduce false hits. It is expected that some, but not all, of the records retrieved by this query may be multilocal documents, and so the fixed field element may need to be changed to code c, or to a more appropriate gov't pub code.

(OC27.W) Shared SQL (Bibliographic Records: Fixed Fields category): "Bib 008 govpub query 2K code is blank could be i." This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov't pub code is either blank or the pipe character, and the Publisher field contains text that could imply an international intergovernmental document, with additional limiting text to reduce false hits. It is expected that some, but not all, of the records retrieved by this query may be international intergovernmental documents, and so the fixed field element may need to be changed to code i, or to a more appropriate gov't pub code.

(OC27.X) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 2L code is blank could be m.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is either blank or the pipe character, and the Publisher field contains text that could imply a multistate document, with additional limiting text to reduce false hits. It is expected that some, but not all, of the records retrieved by this query may be multistate documents, and so the fixed field element may need to be changed to code m, or to a more appropriate gov’t pub code.

(OC27.Y) Shared SQL (Bibliographic Records: Fixed Fields category): “Bib 008 govpub query 2M code is blank could be any code.” This query looks for bib records where the 074 field (GPO item number) is null, and the fixed field gov’t pub code is either blank or the pipe character, and the Publisher field contains generic text that could imply any type of government document. It is expected that some, but not all, of the records retrieved by this query may be government documents, and so the fixed field element may need to be changed to a more appropriate gov’t pub code.

Shared macro (Bibliographic Record Maintenance section):

“Changebib_008GovtPub.mex.” This macro will change the bib record’s 008 gov’t publication code to a new value, as specified by the user.

OC28. Evaluate item records with copy number zero.

Item records with copy number zero result in no copy number displaying in the WebVoyage or VuFind public catalogs when item information is presented to the user, and no copy number included in overdue and other patron notices. Depending on your library’s policies, these may represent errors. Correcting them may also suggest examination of preference settings on staff computers and staff awareness training.

Shared SQL (Item Records: General category): “Find item records with copy number zero”

Shared macro (Item Record Maintenance section): “Changeitem_copy1.mex.” If you have many of these, this macro can help with cleanup.

In addition, if your library’s cataloging practice is to include the copy number in the associated MFHD (recommended), there is a macro available that will add an 852 \$t to the MFHD (Addmfhd_852t.mex).

OC29. Evaluate MFHD records without a specific copy number.

The MARC standard for holding records states that the copy number is to be stored in 852 \$t, but this subfield is not considered to be required. In WebVoyage, when the MFHD does not have an 852 \$t, a default of Copy 1 is displayed to users along with the other holding information. However, VuFind does not currently display a default of Copy 1 at the holdings level display when the MFHD is lacking an 852 \$t. This project may have higher importance for libraries that have a local policy for adding the 852 \$t and/or that use VuFind as their public catalog interface.

Shared SQL (MFHD Records: General category): “MFHDs lacking a copy number” queries (REV. 12/2010). This is a sequence of four queries that should be run in order. The output from the fourth query is the one that will report on the problematic records.

Shared macro (MFHD (Holdings Record) Maintenance section): “Addmfhd_852t.mex.” If your library’s cataloging practice is to include the copy number in the associated MFHD (recommended), this macro will insert a new 852 \$t with default value of 1.

OC30. Eliminate duplicate OCLC numbers.

When more than one Voyager bib contains the same OCLC control number, all but one should be deleted (preferably) or at least OPAC-suppressed. Multiple unsuppressed records with the same OCLC number will cause confusion to users and staff and result in problems in the I-Share Union Catalog.

Shared SQL (Bibliographic Records: Control Numbers category): “Duplicate OCLC#s Main Query”

OC31. Reconcile superseded OCLC numbers.

When OCLC merges bib records in WorldCat, the superseded control number is placed in the 019 field of the retained bib. Over time, both versions of the records can make their way into the local library’s database, resulting in a type of “duplicate” OCLC number. The presence of these bibs in the I-Share Union Catalog can also result in duplicates and/or discards for incoming records into the UC.

Libraries that participate in OCLC’s Bibliographic Notification Service are likely to have more of these “019 duplicates” in their local database.

Shared SQL (Bibliographic Records: Control Numbers category): “Bibs whose OCLC number is in another bib’s 019 field”

OC32. Supply full description for Cataloging in Publication level bibliographic records.

Shared SQL (Bibliographic Records: General category): “Identifying Cataloging in Publication level bibliographic records” queries (REV. 12/2010). This is a sequence of three queries that should be run in order. The output from the third query is the one that will report on the problematic records.

Specifically, these queries look for a 300 field that contains no data other than the string “p cm.” Catalogers can then determine whether downloading a fresh OCLC record or manually completing the 300 field is the better strategy for correcting the problem. Often, a fresh, upgraded OCLC record will contain a completed 300 field as well as additional access points.

A technique for using the output from a shared SQL query and downloading fresh OCLC records in batch using Connexion client can be found at http://www.carli.illinois.edu/mem-serv/mem-train/090430cat/090430JW_RepairorTrade.pdf.

OC33. Evaluate MFHDs with ‘OK to Export’ box checked.

In itself, this is not a problem, as we do not export MFHDs anywhere. However, it may indicate that an operator intended to set the ‘OK to Export’ option in the Bib record (to enable backloading of the bib record to OCLC) and did not. Libraries with DRA experience would be most likely to do this, as export readiness was controlled in the MFHD in that system.

NOTE: a few I-Share libraries have multiple OCLC symbols for different administrative entities (e.g., the four campuses of Illinois Eastern Community Colleges), and their bib record backloading routines are legitimately based on OK to export settings in the MFHD. Libraries with this type of OCLC setup should ignore this project.

Shared SQL (MFHD Records: General category): “Find MFHDs with OK to export set”. This query will find cases where the MFHD has the OK to export flag set to Yes, but the linked bib has the OK to export flag set to No.

Shared macro (Bibliographic Record Maintenance section): “Changebib_OKexport.mex.” In cases where the bib record needs the OK to Export option set, this macro is available to help with cleanup.

OC34. Correct MFHD record type codes.

Voyager currently does not base functionality on MFHD record type. However, it is possible that future implementations may do so, and some catalogers may want to prepare now.

Also, libraries that participate in OCLC's Local Holding Record (LHR) backloading projects might want to place more emphasis on these values in their Voyager holdings records. When MFHDs are submitted to OCLC for the LHR project, OCLC will reset the MFHD record type value to "y" in the WorldCat LHR record (no change is made to the Voyager MFHD). Having the MFHD record type set correctly in Voyager may alleviate the chance that this code would be reset to an incorrect default in WorldCat.

(OC34.A) Shared SQL (MFHD Records: General category): "MFHD rec type not 'y' for serial bib"

(OC34.B) Shared SQL (MFHD Records: General category): "MFHD rec type 'x' for mono bib, but has item enum"

(OC34.C) Shared SQL (MFHD Records: General category): "MFHD rec type 'v' for mono bib, but no item enum"

(OC34.D) Shared SQL (MFHD Records: General category): "MFHD rec type 'y' for mono bib"

(OC34.E) Shared SQL (MFHD Records: General category): "MFHD rec type not v, x, or y"

Shared macro (MFHD (Holdings Record) Maintenance section):
"Changemfhd_rectype.mex" can help with this task. This macro will change the MFHD record type to a value of "y" (by default, but can be edited as needed) for all MFHD IDs included in the input file.

OC35. Correct bib record fixed field Literary Form codes.

None of the former or current public catalogs supported for I-Share libraries use the coding in the bib record fixed field's Literary Format (MARC 008/33 for books) for searching or display. However, there is an enhancement request to add this functionality to VuFind, and it is expected that the public catalog for the eXtensible Catalog project will use the data in this fixed field for faceting/limiting. Some catalogers may want to prepare now.

There are several queries that can be used to identify bib records with various types of Literary Form coding errors. It is assumed the queries will be run in the order listed below.

(OC35.A) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic record 008 fiction filter query 1”

This query will find bibliographic records for books where the fixed field’s Literary Format value is other than 1, f, or j (generally representing non-fiction), but a subject heading in the record contains “fiction” and variations. It is expected that many, but not all, of the records retrieved with this query will be coded incorrectly.

(OC35.B) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic record 008 fiction filter query 2”

This query will find bibliographic records for books where the fixed field’s Literary Format value is other than 1, f, or j (generally representing non-fiction), but the title in the record contains “*novel*” or “*stor*” (but not “*histor*”), and the call number begins with 8 or P (but not PN). It is expected that many, but not all, of the records retrieved with this query will be coded correctly. For example, criticism about a work of fiction could be retrieved with this query, and such a book would be properly coded as 0. Library staff will want to review the query results to decide which records need to be edited and which records are coded correctly.

(OC35.C) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic record 008 fiction filter query 3A”

This query will find bibliographic records for books where the fixed field’s Literary Format value is 1, f, or j (generally representing fiction), but the call number in the linked MFHD begins with a value other than 8 or P. It is expected that many, but not all, of the records retrieved with this query will be coded incorrectly.

(OC35.D) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic record 008 fiction filter query 3B”

This is a follow-up query to Bibliographic record 008 fiction filter query 3A. It looks for the same criteria as query 3A (bib records for books where the fixed field’s Literary Format value is 1, f, or j), but no call numbers are omitted from the results. . It is expected that most, but not all, of the records retrieved with this query will be coded correctly. For example, most of the results of this query are expected to be works of fiction, but the query can also retrieve criticism about a work of fiction. The criticism’s Literary Form value should be coded as 0. Library staff will want to review the query results to decide which records need to be edited and which records are coded correctly.

(OC35.E) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic record 008 fiction filter query 4A”

This query will find bibliographic records for books where the fixed field’s Literary Format value is blank, but the call number in the linked MFHD begins with a value other than 8 or P. While blank is technically a valid value for Literary Form, future OPAC functionality may consider a blank value to represent non-fiction. It is expected that most, but not all, of the records found by this query would qualify to have their LitForm value to be set to 0 (for non-fiction), to remove the ambiguity of a blank value.

(OC35.F) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic record 008 fiction filter query 4B”

This is a follow-up query to Bibliographic record 008 fiction filter query 4A. It looks for the same criteria as query 4A (bib records for books where the fixed field’s Literary Format value is blank), but only call numbers that begin with 8 or P are included in the results. It is expected that many, but not all, of the records found by this query would qualify to have their LitForm value to be set to 0 (for non-fiction), to remove the ambiguity of a blank value. However, this query will retrieve some records that are works of fiction. Library staff will need to review the results before determining what the correct LitForm code should be.

Shared macro (Bibliographic Record Maintenance section):
“Changebib_008LitForm.mex” can help with this task. This macro will change the bib record status to a value of “0” (by default, but can be edited as needed) for all bib IDs included in the input file.

OC36. Correct invalid or obsolete bib record fixed field Publication status codes. <NEW 05/2011>

None of the former or current public catalogs supported for I-Share libraries use the coding in the bib record fixed field’s Publication status (MARC 008/06) for searching or display. However, it is expected that the public catalog for the eXtensible Catalog (XC) project will use the data in this fixed field for some faceting/limiting functionality. Some catalogers may want to prepare now.

There are several queries that can be used to identify bib records with various types of Publication status (a.k.a., Date type status) coding errors. It is assumed the queries will be run in the order listed below.

NOTE: It is expected that records with invalid Publication status coding retrieved by queries 1 through 7 (sub-projects A through G) below will result in negative consequences in the XC public catalog. It is unknown at this time if the records retrieved by queries 8 through 17 below will prove to be problematic in the XC as

well. Therefore, while queries 8 - 17 do represent invalid MARC coding for this data element, they can be considered optional at this time for maintenance projects.

Some of the queries below may retrieve a larger than expected number of records, often due to changes made to the definition of some codes in the fixed field's Publication status. For example, prior to MARC record changes adopted in 1995, the 008 Publication status code of "c" for books was essentially the equivalent of today's publication status code of "t." The Publication status code of "c" now means "continuing resource currently published" and thus should only be used for serial records.

If needed, more information about the history of Publication status coding practices can be found on the Library of Congress description of the MARC 21 format for bib records, in the "Content Designator History" section:

<http://www.loc.gov/marc/bibliographic/bd008a.html>

(OC36.A) Shared SQL (Bibliographic Records: Fixed fields category):
"Bibliographic records with invalid 008 pubstatus query 1"

This query will find bibliographic records where the fixed field's Publication status code is c, and the bib format is for a serial, and the 008 date2 is not 9999. It is expected that that the fix for most, but not all, of the records retrieved with this query should be to change the Publication status code to d.

(OC36.B) Shared SQL (Bibliographic Records: Fixed fields category):
"Bibliographic records with invalid 008 pubstatus query 2"

This query will find bibliographic records where the fixed field's Publication status code is d, and the bib format is for a serial, and the 008 date2 is null. It is expected that the fix for most, but not all, of the records retrieved with this query should be to add an appropriate 008 date2.

(OC36.C) Shared SQL (Bibliographic Records: Fixed fields category):
"Bibliographic records with invalid 008 pubstatus query 3"

This query will find bibliographic records where the fixed field's Publication status code is s, and the bib format is for a serial. It is expected that that the fix for records retrieved with this query should be to either correct the bib format code, or to change the Publication status code to c or d, and add an 008 date2, as applicable.

NOTE: Each of the three queries above are limited to serial records, which are likely to have updated content over time. In testing these queries and potential fixes to the invalid coding, it was noticed that many of the records retrieved have already been corrected in OCLC. Therefore, a recommended strategy for fixing the records retrieved by the three queries above is to get fresh copies of the serial bibs from

OCLC, replace the existing bibs in the local Voyager database, and then re-run the queries and manually correct any remaining problematic records. This strategy may often result in other updates to the bib, such as a completed 362 field or the addition of a 785 field.

A technique for using the output from a shared SQL query and downloading fresh OCLC records in batch using Connexion client can be found at http://www.carli.illinois.edu/mem-serv/mem-train/090430cat/090430JW_RepairorTrade.pdf.

(OC36.D) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 4”

This query will find bibliographic records where the fixed field’s Publication status code is c, and the bib format is not for a serial, and the 260 \$c contains a hyphen, and the 008 date2 is not null. It is expected that that the fix for most records retrieved with this query should be to change the Publication status code to m.

(OC36.E) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 5”

This query will find bibliographic records where the fixed field’s Publication status code is c, and the bib format is not for a serial, and the 260 \$c does not contain a hyphen, and the 008 date2 is not null. It is expected that that the fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of less than 8 should be to change the Publication status code to s and to delete the 008 date2. The fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of greater than 7 should be to change the Publication status code to t, or if the format is for sound recordings, to code p.

(OC36.F) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 6”

This query will find bibliographic records where the fixed field’s Publication status code is d, and the bib format is not for a serial, and the 008 date2 is not null. It is expected that that the fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of less than 8 should be to change the Publication status code to e. The fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of greater than 7 should be to change the Publication status code to t, or if the work is a multivolume set, to code m.

(OC36.G) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 7”

This query will find bibliographic records where the fixed field’s Publication status code is c, d, or k, with no limits on bib format, and the 008 date2 is null. It is

expected that that the fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of less than 8 should be to change the Publication status code to s. The fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of greater than 7 should be to change the Publication status code to t and to add the appropriate 008 date2.

(OC36.H) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 8”

This query will find bibliographic records where the fixed field’s Publication status code is m, with no limits on bib format, and the 260c does not contain a hyphen, and the 008 date2 is null. It is expected that that the fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of less than 8 should be to change the Publication status code to s. The fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of greater than 7 should be to change the Publication status code to t add the appropriate 008 date2.

(OC36.I) Shared SQL (Bibliographic Records: Fixed fields category): “Bibliographic records with invalid 008 pubstatus query 9”

This query will find bibliographic records where the fixed field’s Publication status code is s, with no limits on bib format, and the 260c contains a hyphen, and the 008 date2 is not null. It is expected that that the fix for most, but not all, of the records retrieved with this query should be to change the Publication status code to m, or to code q in the cases of questionable dates.

(OC36.J) Shared SQL (Bibliographic Records: Fixed fields category): “Bibliographic records with invalid 008 pubstatus query 10”

This query will find bibliographic records where the fixed field’s Publication status code is s, with no limits on bib format, and the 008 date2 is not null. It is expected that that the fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of less than 8 should be to change the Publication status code to r, or code p for videos and sound recordings. However, if there is no data in the bib’s Notes or other fields indicating a reprint/reissue, the fix may be to delete 008 date2. The fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of greater than 7 should be to change the Publication status code to t.

(OC36.K) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 11”

This query will find bibliographic records where the fixed field’s Publication status code is n, with no limits on bib format, and the 008 date1 is not uuuu. It is expected that that the fix for most, but not all, of the records retrieved with this query that

contain a value in Length260c of less than 8 should be to change the Publication status code to s. The fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of greater than 7 should be to change the Publication status code to t or r, and to add the appropriate 008 date2, if currently blank.

(OC36.L) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 12”

This query will find bibliographic records where the fixed field’s Publication status code is code i, and the bib format is for sound recordings, and the 008 date2 is not null. It is expected that that the fix for most, but not all, of the records retrieved with this query should be to change the Publication status code to p.

(OC36.M) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 13”

This query will find bibliographic records where the fixed field’s Publication status code is e, i, m, q, r, or t, with no limits on bib format, and the 008 date2 is null. It is expected that that the fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of less than 8 should be to change the Publication status code to s. The fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of greater than 7 should be to add the appropriate 008 date2.

(OC36.N) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 14”

This query will find bibliographic records where the fixed field’s Publication status code is p, with no limits on bib format, and the 008 date2 is null. It is expected that that the fix for most, but not all, of the records retrieved with this query should be to add the appropriate 008 date2.

(OC36.O) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 15”

This query will find bibliographic records where the fixed field’s Publication status code is u and the bib format is not for serials. It is expected that that the fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of less than 8 should be to change the Publication status code to s, and to edit the 008 date1 and date2 as appropriate. The fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of greater than 7 should be to change the Publication status code to t or r, and to add the appropriate 008 date2, if currently blank.

(OC36.P) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 16”

This query will find bibliographic records where the fixed field’s Publication status code is u, the bib format is for serials, and the 008 date2 is not uuuu. It is expected that that the fix for most, but not all, of the records retrieved with this query that contain a 008 date2 value other than 9999 should be to change the Publication status code to d. The fix for most, but not all, of the records retrieved with this query that contain a 008 date2 value of 9999 should be to change the Publication status code to c, or keep as u, and to adjust the 008 date1 and date2 as appropriate for an uncertain publication date.

NOTE: because this query targets serial records, replacing the existing bibs with fresh copies from OCLC and then re-running the query again may prove to be worthwhile. A technique for using the output from a shared SQL query and downloading fresh OCLC records in batch using Connexion client can be found at http://www.carli.illinois.edu/mem-serv/mem-train/090430cat/090430JW_RepairorTrade.pdf.

(OC36.Q) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 17”

This query will find bibliographic records where the fixed field’s Publication status code is b or | (pipe character), with no limits on bib format, and the 008 date1 and date2 are not null. It is expected that that the fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of less than 8 should be to change the Publication status code to s and to adjust 008 date2. The fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of greater than 7 should be to change the Publication status code to t.

(OC36.R) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid 008 pubstatus query 18”

This query will find bibliographic records where the fixed field’s Publication status code is not any valid code per the MARC standard, with no bib format or other limits. It is expected that that the fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of less than 8 should be to change the Publication status code to s. The fix for most, but not all, of the records retrieved with this query that contain a value in Length260c of greater than 7 should be to change the Publication status code to t.

<NEW> Shared macro (Bibliographic Record Maintenance section):
“Changebib_008pubstatus.mex.” This macro can be used with the output from the any of the queries described above, when the proper fix is to change the publication status code. The macro will edit the bib record’s Publication Status code to a new value specified by the user, for all bib IDs included in the input file.

<NEW> Shared macro (Bibliographic Record Maintenance section):
“Changebib_008date2.mex.” This macro can be used with the output from the any of the queries described above, when the proper fix is to edit the fixed field date2 and there are too many records to fix manually. The macro will edit the bib record’s Date2 code to a new value specified by the user, for all bib IDs included in the input file.

OC37. Correct bib record Leader Record Status codes.

Voyager currently does not base any functionality the record status code in the bib record’s Leader. However, it is possible that future implementations may do so, and some catalogers may want to prepare now.

(OC37.A) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with invalid Rec Status”

This query will find bibliographic records with a record status code other than a, c, d, n, or p. These records should be edited to contain a valid record status code; a suggestion is to use “c”.

(OC37.B) Shared SQL (Bibliographic Records: Fixed fields category):
“Bibliographic records with Rec Status code set to ‘d’”

A record status code of “d” (for deleted) is valid per the MARC standard, and one can manually set this value in the Voyager cataloging client. But some staff may have set the Rec status value of “d” under the false assumption that changing this value would result in the bib record being deleted from the database. In addition, although the cataloging client bases no functionality on this code, it is unknown if future implementations may use it to delete records. It is recommended that if any records retrieved by this query represent works still held in the library’s collection, that the record status code be edited to a value other than “d”; a suggestion is to use “c.”

Shared macro (Bibliographic Record Maintenance section):
“Changebib_RecStatus.mex” can help with this task. This macro will change the bib record status to a value of “c” (by default, but can be edited as needed) for all bib IDs included in the input file.

OC38. Correct bibliographic records with invalid use of code m in 008 Nature of Contents. <NEW 05/2011>

None of the former or current public catalogs supported for I-Share libraries use the coding in the bib record fixed field’s Nature of Contents (MARC 008/24-27 for books) for searching or display. However, it is expected that the public catalog for

the eXtensible Catalog (XC) project will use the code of “m” in this fixed field under certain circumstances. Some catalogers may want to prepare now.

At this time, it is unclear exactly how the XC public catalog will use this data. But the XC documentation includes a specification for the use of, or creation of, a field/label of “Thesis”, to be used in some way in the FRBRization of the catalog at the work level. Specifically, a field/label of “Thesis” will be added to XC records as follows:

- * when the bib contains a 502 field, the content of \$a will populate the XC Thesis field;
- * if the bib has no 502 field, the generic text “Thesis” will be added to the XC record if any of the fixed field Nature of contents bytes contains the code “m”.

This maintenance project is designed to find records that do not contain a 502 field but do contain Nature of contents code m. Any records for which an XC field/label of Thesis would be inappropriate should be edited to delete the code m from Nature of contents. Note that in this context, “Thesis” refers to dissertations as well as master’s theses.

Shared SQL (Bibliographic Records: Fixed fields category): “Bib records with 008 Nature of contents code for thesis”. This is a sequence of three queries that must be run in order. The output from the third query is the one that will report on the potentially problematic records.

Special note: According to OCLC’s Bibliographic Formats and Standards document, field 502 is to be used only for original theses and their copies. A 500 field is to be used for notes about the work originating as a thesis but now published. The third query in this series may retrieve many records that represent commercially published titles that contain a 500 note with text similar to “Originally presented as the author’s thesis...” These records should be edited to delete the code “m” from the Nature of contents fixed field.

Some records retrieved by query 3 may represent original theses/dissertations where the thesis note is found in field 500. For these records, the field tag for the thesis note should be changed to 502.

Some records retrieved by query 3 are cases where the Nature of contents code m was applied in error, when the item is not and was never based on a thesis. These records should be edited to delete the code “m” from the Nature of contents fixed field.

Projects to do once

ON1. Correct item barcodes that do not belong to your library.

During the consortium's DRA era, it was possible for one library to edit another library's item records. As a result, some items received barcode numbers belonging to a different library. Correcting these problems is important for circulation and ILL service.

Shared SQL (Item Records: Barcode category): "Item barcodes that do not belong to your library"

ON2. Evaluate DRA- assigned barcodes beginning 38888.

Libraries that were not ILCSO members during the DRA era will not have this problem. Barcode numbers of this type were machine generated when items in our system previous to DRA contained no barcode, and thus they should now identify items that still do not have real barcodes. Whether these represent a problem or not depends on the circulation status of the item and local policies. If the item's type is one that allows requesting, replacing these with real barcodes will help ILL staff, because all items sent out of the library to fill Universal Borrowing requests require a physical barcode on the piece that matches the item record's barcode number.

Shared SQL (Item Records: Barcode category): "Identifying DRA assigned barcodes beginning 38888"

ON3. Correct the "653 problem."

Only libraries that were consortium members during the FBR era (pre-1998) should have this problem. Because of legacy system indexing limitations, alternate title information wound up in field 653. Correcting this problem improves title access to affected records. More information on the "653 problem" is available from <<http://www.carli.illinois.edu/mem-prod/I-Share/cat/corr653prob.html>>.

Shared SQL (Bibliographic Records: Titles category): "Correcting the 653 problem" queries (REV. 12/2010). This is a sequence of three queries that should be run in order. The output from the third query is the one that will report on the problematic records.

ON4. Replace “MARCettes” with full cataloging.

Only libraries that were consortium members during the FBR era (pre-1998) should have these records. Short “MARCette” records that remain in the database from the consortium’s LCS era had only title entries and sometimes a single author field as access points, as well as very limited description. In particular, MARCettes have no subject headings. Replacing these records with full cataloging improves subject access as well as other access points such as added authors, series, standard control numbers, etc. Some records found by this query may require original cataloging, if they represent uniquely held items.

Shared SQL (Bibliographic Records: Control Numbers category): “Bib record 035 \$a begins with LCS Scope Code”

ON5. Correct ISBN numbers with fewer than 10 digits.

In the early years, there were nine-digit standard book numbers, which were soon expanded to 10 digits. Correcting this problem enhances retrieval and matching functionality when using the ISBN. Reloading fresh bibs from OCLC usually corrects this problem, with the added advantage of other possible enhancements to access, description, and MARC coding.

Shared SQL (Bibliographic Records: Control Numbers category): “Find ISBN numbers less than 10 digits”

ON6. Evaluate 028 field data for miscoded corporate authors.

Libraries that were not ILCSO members during the FBR era should not have this problem. A number of older bibliographic records for print monographs and serials have corporate added entries or other invalid data in the 028 field. The query below searches for 028 subfields a or b in records with bib type "am" (print monographs) or "as" (print serials) so the information in the 028 can be evaluated and placed in the correct field if appropriate, or (sometimes) deleted. If the 028 field data is a valid publisher number, the problem could be that the format coding for the bib record is in error.

Shared SQL (Bibliographic records: General category): “Bibliographic records with 028 in non-media records”

Appendix: Estimated Run Times for the SQL Queries

It is difficult, but not impossible, to estimate how long an Access query will run. The timing is affected by the size of your Voyager database and the speed of your workstation. The capacity of your internet connection is also an important factor and this may vary depending on how busy your campus connection is at any particular time.

The chart below gives an estimated run time for the queries cited in this document. The database on which the timings were done has 120,000 bib records, which is near the median size for I-Share libraries. Without doing timings of your own, this chart will at least tell you which queries should run in under a minute and which ones may take substantially longer. However, if you make a note of your own timings, you may find that your timings are typically, say, double what you see in this chart. This will allow you to make more accurate estimates for your technical environment.

*** IMPORTANT:** For queries marked with an asterisk, be sure that you use a version dated 12/1/2010 or later. Earlier versions of these queries will run significantly slower. To see the date in Access, view your list of queries in Details view and check the Modified date. If you need a newer copy of the query, you can get it from the Shared SQL page on the CARLI web site <<http://www.carli.illinois.edu/mem-prod/I-Share/secure/sql.html>>.

Query Name	Estimated Timing
Ampersands in Titles	< 1 minute
Bib 008 govpub query 1A code is f	< 1 minute, for 50,000 bibs
Bib 008 govpub query 1B code is s	1 to 15 minutes, for 50,000 bibs
Bib 008 govpub query 1C code is s PUB maybe UP	1 to 15 minutes, for 50,000 bibs
Bib 008 govpub query 1D code is L	< 1 minute, for 50,000 bibs
Bib 008 govpub query 1E code is c	< 1 minute, for 50,000 bibs
Bib 008 govpub query 1F code is i	< 1 minute, for 50,000 bibs
Bib 008 govpub query 1G code is m	< 1 minute, for 50,000 bibs
Bib 008 govpub query 1H code is a	< 1 minute, for 50,000 bibs
Bib 008 govpub query 1I code is o	< 1 minute, for 50,000 bibs
Bib 008 govpub query 1J code is z	< 1 minute, for 50,000 bibs
Bib 008 govpub query 1K code is u	< 1 minute, for

	50,000 bibs
1Bib 008 govpub query 1L code is invalid	< 1 minute, for 50,000 bibs
Bib 008 govpub query 2A code not f and GPOnum not null	< 1 minute, for 50,000 bibs
Bib 008 govpub query 2B code is blank by loc code	< 1 minute, for 50,000 bibs
Bib 008 govpub query 2C code is blank and call no SuDoc	< 1 minute, for 50,000 bibs
Bib 008 govpub query 2D code is blank could be f PUB	< 1 minute, for 50,000 bibs
Bib 008 govpub query 2E code is blank could be f AUTH	< 1 minute, for 50,000 bibs
Bib 008 govpub query 2F code is blank could be s PUB not UP	1 to 15 minutes, for 50,000 bibs
Bib 008 govpub query 2G code is blank could be s AUTH not UP	1 to 15 minutes, for 50,000 bibs
Bib 008 govpub query 2H code is blank could be s PUB may not UP	1 to 15 minutes, for 50,000 bibs
Bib 008 govpub query 2I code is blank could be L	< 1 minute, for 50,000 bibs
Bib 008 govpub query 2J code is blank could be c	< 1 minute, for 50,000 bibs
Bib 008 govpub query 2K code is blank could be i	< 1 minute, for 50,000 bibs
Bib 008 govpub query 2L code is blank could be m	< 1 minute, for 50,000 bibs
Bib 008 govpub query 2M code is blank could be any code	< 1 minute, for 50,000 bibs
Bib record 035 \$a begins with LCS Scope Code	< 1 minute
Bib record 035 \$a contains (XXXdb)nnnn	< 1 minute
Bib record without UT but title contains transl	< 1 minute, for 50,000 bibs
Bib record without UT but 700 field contains tr	< 1 minute, for 50,000 bibs
Bib record without UT but 6xx field contains Translations into	< 1 minute, for 50,000 bibs
Bib record without UT but 041 indicator 1 is 1	depends on query 1; 1 to 15 minutes for 50,000 bibs
Bib records with 008 DATE1 less than 4 chars	< 1 minute
Bib records with 008 DATE1 may contain invalid "u" data	< 1 minute
Bib records with 008 DATE2 less than 4 chars	< 1 minute
Bib records with 008 Nature of contents code for thesis queries	< 1 minute for each query

Bib records with an ISSN but format is not for serials	< 1 minute
Bib records with Lang code set to obsolete or invalid value	< 1 minute
Bib records with null 260c, for scores and sound recordings only	< 1 minute
Bibliographic format counts	< 1 minute
Bibliographic or MFHD records with typos at beginning of URL	< 1 minute
Bibliographic record 008 fiction filter queries	< 1 minute, for 25,000 bibs, for each query
Bibliographic record 010 \$a contains ocm may be OCLC number	1 to 15 minutes
Bibliographic record 010 \$a contains pbk may be ISBN	1 to 15 minutes
Bibliographic record 020 \$a contains double slash may be LCCN	1 to 15 minutes
Bibliographic record 020 \$a contains ocm may be OCLC number	1 to 15 minutes
Bibliographic record 022 \$a not 8 digits	1 to 15 minutes
Bibliographic record 035 \$a contains double slash may be LCCN	1 to 15 minutes
Bibliographic record 035 \$a contains DRA bib ID	1 to 15 minutes
Bibliographic record 035 \$a contains hyphen may be LCCN	1 to 15 minutes
Bibliographic record 035 \$a contains pbk may be ISBN	1 to 15 minutes
Bibliographic record 035 \$a contains punctuation may be barcode	1 to 15 minutes
Bibliographic record 650 index begins with Country Name	1 to 15 minutes
Bibliographic record 650 index begins with State Name	1 to 15 minutes
Bibliographic records where 008 DATE1 does not equal 260c	< 1 minute
Bibliographic records with 028 in non-media records	< 1 minute
Bibliographic records with 260c beginning with letter EL	< 1 minute
Bibliographic records with 260c containing the letter EL	< 1 minute
Bibliographic records with invalid 008 pubstatus queries	< 1 minute, for each query
Bibliographic Records with invalid or obsolete format coding	< 1 minute
Bibliographic records with invalid Rec Status	< 1 minute
Bibliographic records with Lang code not common	< 1 minute
Bibliographic records with Lang code set to 'und'	< 1 minute
Bibliographic records with Lang code set to 'zxx'	< 1 minute
Bibliographic records with multiple 245 fields queries	< 1 minute, for each query
Bibliographic records with null 008 DATE1, but 260c is not null	< 1 minute
Bibliographic records with null 260c but 008 DATE 1 is not null	< 1 minute

Bibliographic records with null 260c, but 260a contains a date	< 1 minute
Bibliographic records with null 260c, but 260b contains a date	< 1 minute
Bibliographic records with Rec Status code set to 'd'	< 1 minute
Bibliographic records with specific format codes	< 1 minute
Bibliographic records without MFHDs	16 to 60 minutes
Bibs whose OCLC number is in another bib's 019 field	< 1 minute
Bibs with 008 DATE1 like 'uuuu' or ' ', but 260c is not null	< 1 minute
Bibs with format beginning 'g' but no video GMD	< 1 minute
Bibs with format beginning 'i' or 'j' but no sound recording GMD	< 1 minute
Bibs with sound recording GMD but record type is not 'i' or 'j'	< 1 minute
Bibs with video GMD but record type is not 'g'	< 1 minute
Bibs without indexed titles	< 1 minute
* Correcting the "653 problem" Query 1	1 to 15 minutes
* Correcting the "653 problem" Query 2	depends on Query 1; 1 to 15 minutes for 10,000 bibs
* Correcting the "653 problem" Query 3	< 1 minute
Duplicate Item Barcodes	< 1 minute
Duplicate OCLC#s Main Query	< 1 minute
Find 020's with multiple \$a's	< 1 minute
Find autobiographies without a 600 field.	< 1 minute
* Find bib records where 245 Ind 2 is a pipe Query 1	< 1 minute
* Find bib records where 245 Ind 2 is a pipe Query 2	depends on Query 1; 1 to 15 minutes for 10,000 bibs
* Find bib records where 245 Ind 2 is a pipe Query 3	< 1 minute
* Find bib records where 245 indicator 2 is blank Query 1	16 to 60 minutes for 10,000 bibs
* Find bib records where 245 indicator 2 is blank Query 2	< 1 minute
Find "error" item type code	< 1 minute
Find holding and item data for a given barcode	< 1 minute
Find holding and item data for all duplicated barcodes	1 to 15 minutes
Find ISBN numbers less than 10 digits	< 1 minute
Find item records with copy number zero	< 1 minute
Find MFHDs with OK to export set	< 1 minute
Finding Public and Non- Public Notes in Holdings Query 1	< 1 minute
Finding Public and Non- Public Notes in Holdings Query 2	depends on Query 1; 1 to 15 minutes for 10,000 MFHDs
Finding Public and Non- Public Notes in Holdings Query 3	< 1 minute
Identifying Bibs with empty index entries	< 1 minute

* Identifying Cataloging in Publication level bibliographic records Query 1	< 1 minute
* Identifying Cataloging in Publication level bibliographic records Query 2	depends on Query 1; 1 to 15 minutes for 10,000 bibs
* Identifying Cataloging in Publication level bibliographic records Query 3	< 1 minute
Identifying DRA assigned barcodes beginning 38888	< 1 minute
Individual Bibs with more than one OCLC number	16 to 60 minutes
Item barcodes that do not belong to your library	< 1 minute
Item barcodes with invalid length	< 1 minute
Item Count by Location and Type	< 1 minute
Item records containing double quote	1 to 15 minutes
Item records with perm loc different from the MFHD loc	< 1 minute
Items with no Active barcodes	< 1 minute
Items with specific combinations of perm loc and item type	< 1 minute
MFHD rec type not v, x, or y	< 1 minute
MFHD rec type not 'y' for serial bib	< 1 minute
MFHD rec type "v" for mono bib, but no item enum	< 1 minute
MFHD rec type "x" for mono bib, but has item enum	< 1 minute
MFHD rec type "y" for mono bib	< 1 minute
* MFHDs lacking a copy number Query 1	< 1 minute
* MFHDs lacking a copy number Query 2	< 1 minute
* MFHDs lacking a copy number Query 3	depends on Query 2; 1 to 15 minutes for 3,000 MFHDs
* MFHDs lacking a copy number Query 4	< 1 minute
MFHDs lacking a specific public note Query 1	< 1 minute
MFHDs lacking a specific public note Query 2	depends on Query 2; 1 to 15 minutes for 10,000 MFHDs
MFHDs lacking a specific public note Query 3	< 1 minute
MFHDs missing call number prefix for a specific location	< 1 minute for loc with 5,000 MFHDs
MFHDs missing call number suffix for a specific location	< 1 minute for loc with 5,000 MFHDs
MFHDs that are suppressed from OPAC with Items attached	< 1 minute
Missing call numbers in MFHDs with items attached	< 1 minute
Missing LC subject query 1	< 1 minute
Missing LC subject query 2	< 1 minute

	depends on Query 2; 16 to 60 minutes for 10,000 bibs
Missing LC subject query 3	
Missing LC subject query 4	< 1 minute
Odd Prefixes Query 1 – List All Prefixes	16 to 60 minutes
Odd Prefixes Query 2 – Rare Prefixes	< 1 minute
Odd Prefixes Query 3 – MFHDs with Rare 852\$k	< 1 minute
	depends on Query 1; < 1 minute for 10,000 prefixes
Odd Prefixes Query 4 – Prefixes not in 852\$k	
Serial bib records without an ISSN	< 1 minute
Suppressed bibliographic records with items attached	< 1 minute
Symbols in Titles	1 to 15 minutes
Unindexed Call Numbers	< 1 minute