

Authorities Committee Meetings November 20-21, 1996

*Respectfully submitted by: Mary Ann O'Daniel and Nancy Lynne Williams
January 11, 1997*

The Authorities Committee held two meetings covering 1 1/2 days November 20-21, 1996. The following SUS institutions were represented either by the Authorities Committee member and/or by staff handling authorities work:

University of Florida, University of Florida Legal Information Center, Florida State University, Florida State University Law Library, University of South Florida, University of Central Florida, and University of North Florida.

The first day meeting was devoted to training in the use of Northwestern's CLARR software. Gary Strawn had given UF permission to distribute to the SUS libraries the June version of CLARR. All the above attendees had this software distributed to them on floppy disks. The Medical Library at USF also was given a copy of the software. The meeting started with Phek Su (UF) providing highlights from the presentations she had made on CLARR for the spring FLA Conference and the fall NOTIS Users Group Meeting. Dan Cromwell (UF), who prepared the copies of the software, provided some guidance and instruction on installation. It was agreed that questions could be posted to TS-PLAN or sent to Nancy Williams (nanwill@nervm.nerdc.ufl.edu)

After break Jimmie Lundgren (UF) and Betsy Simpson (UF) provided some information on UF's testing of the more recent versions of CLARR. These versions are being tested and UF has begun doing much of its original cataloging directly on LUIS and using CLARR to capture and post the record directly into OCLC. The CLARR software translates the various LUIS codes into compatible OCLC codes. This technique is being used for both bibliographic records as well as authority records. Comments were made that CLARR is more than an authority tool--it is a cataloging tool and in the future possibly even more than a cataloging tool.

Nancy Williams talked through a 16-page handout everyone was able to take with them covering the most frequently used CLARR buttons and samples of print reports for two titles. Interpreting the reports and pointing out what CLARR can and cannot do was delineated.

After lunch Nancy took the group to her computer and covered a number of examples of CLARR in operation and she discussed results of the searches. Using various CLARR buttons and converting a 9xx record to a firm field record was demonstrated. As always everyone was amazed that one click could take the information on a bibliographic record and create an authority record for any designated heading. Creating a series authority was shown.

For over an hour the participants broke into three groups with three in each group to get hands on experience. Nancy tried to answer questions that arose and USF even identified a bug in the post-June version of CLARR they were working with. Everyone seemed very pleased with the

CLARR software and it was hoped they would all be able to install their copies soon on their institution's hardware and begin CLARRing.

The second day meeting was held at FCLA and Mary Ann O'Daniel and Marty Johnson were in attendance. Three main topics had been identified for this meeting: the progress in the work on automatic derivation/claiming, delete records in LTLC, and state-wide global changes. The first topic was the hot topic and almost the whole meeting was spent on it.

The program that FCLA is working on for automatic derivation and claiming of LTLC authority records is based on existing NOTIS programs. The program that produces the conflict and error detection report for an institution group (i.e., AM, CF, FA, FI, FS, GC, NF, SF, UF, WF) also produces a report of bibliographic record index headings that do not match against any authority records in the institution group's local authority file. The autoclaim program will take these unmatched bib headings and match them against authority records in LTLC. LTLC authorities will be derived into the local authority file and the heading use codes set (thus the "derive and claim" description of the program).

In response to questions about rules used to match the headings and derive authorities, Marty Johnson of FCLA said the same rules could be applied that apply to the weekly LC authority update (LCA) overlays for each institution. It would be a multi- step process: first records that would be claimed would be pulled, then the LCA criteria would be applied. Mary Ann O'Daniel of FCLA will distribute to each library for review information about the options currently being used for its LC updates. Marty cautioned the group that since the matches will occur on index entries, not the entries as they are in the bibl or authority records, auto derive and claim may not find unique matches on very long headings.

Extensive discussion by the Authority Subcommittee members followed regarding the authorities to claim, how to determine which will be derived, and potential pitfalls.

Alva Stone of FSU Law said that Law does not always derive authorities if there are no cross references. Marty is not sure if he can program for that. If a program can identify the absence of cross references, then it can skip or journal the record. Criteria for programmatic identification of these records would need to be supplied. Nancy Williams of UF pointed out that authorities that don't have cross references now may have them in the future. Deriving them into the local authority file would enable the weekly LCA to update them. Mary Ann commented that file storage space does not need to be taken into consideration in any of the decisions. Authority files are relatively small compared to the other datafiles FCLA now handles.

There was some discussion regarding the BNA records in local authority files. The consensus was that these could be overlaid.

Discussion turned to locally added fields in authority records and whether they could be protected. The libraries need to come to agreement about the fields to be preserved and have a uniform practice to enable machine identification of the fields that should be preserved in the overlay process.

There was discussion regarding pre-AACR2 headings: whether to derive, to derive and journal, or to journal but not derive. There was general agreement that personal names are probably safe to derive. Corporate names are more problematic and derivation may need to be a local option.

Discussion moved on to setting and correcting heading use codes. The program currently looks at only the subject use codes. Marty asked how the codes should be set for names and for subjects. He can provide a list of authority keys derived as a byproduct of the process. Nancy suggested that the names should be set at "cab" and subjects as "bcb." Deborah Crosby of FSU said 151s will be tricky, especially distinguishing between 110s and 151s.

There was agreement that provisional records (E/L 9) should be excluded from the auto derive & claim process. There is not a way to identify pre-order records that use firm fields, so they cannot be excluded.

Alva asked if there will be a list of headings for NACO work; UF also expressed interest. This could be an IG option. Nancy asked for a future report of authorities with no 010s.

There was extensive discussion about series authorities. One option is to identify but not derive and claim them. There may need to be different decisions for the retrospective auto derive & claim over all the data and the ongoing, possibly monthly, auto derive and claim to cover authority updates. Consensus was that the retrospective run should report matched series authorities but not derive & claim them. The ongoing runs should derive and report, with this possibly being an institutional option.

There are two Appendices that provide further information supporting the discussion. [Appendix A](#) is the outline that Mary Ann jotted on the FCLA blackboard. [Appendix B](#) is the information Linda Smith agreed to provide on UNF's automatic authority processing.

The other two topics can be summarized as follows: on delete records in LTLC Nancy asked everyone to think about how we handle delete authorities and how deleted but viewable authorities in LTLC are used. We each had received reports from Mary Ann on the number of deletes in our separate files. Most of us said we would like LUIS number reports on these small d authority records since they are still searchable, viewable, and in need of cleaning up.

The state-wide global changing of headings has been in suspension for almost two years. Since most of us now have some hardware that allows macros or copying/pasting, the old number directive of submitting globals if 10 or more records were involved in a particular library's file was considered probably too low. It seems that we get only a few headings a month that require voluminous file changes, so it was agreed to up the submission number to 50 records affected, to set it up to run monthly, no matter how many headings might be submitted, and to set some regularity. Mary Ann needs to do a little more work on learning how the program runs and its time needs.

Nancy pointed out the beta version of CLARR has a button for Error Detection work and this might prove useful for globals. Once we learn how it might be used and once the OK is given to distribute the more powerful CLARR, state-wide globals may not be continued. It was pointed

out the state-wide global process did clean the headings in all files and brought consistency for the heading form across all data bases. This consistency would be lost if some libraries did not use CLARR or lacked the original authority record to clue them that the authority record was changed and consequently the bibliographic record(s) needed changing. However with automatic derivation/ claiming our individual authority files should be more in sync and changes reported via the weekly changed authority reports.

Authority Committee Meetings Report--Appendix A

Automatic Derive & Claim of Authority Records Notes from 11/21/96 Meeting

CONDITIONS FOR AUTOMATIC DERIVE & CLAIMING

- * Do not derive non-unique personal name authorities (008, byte 32, code b "undifferentiated personal name")
- * Do not look at 9xx (provisional) fields in bibliographic records or at bibliographic records that are encoding level (E/L) 9.
- * Do not derive deleted authority records (leader, byte 5, code d, s, or x).
- * If conference headings are parsed in MHI entries, then match using these criteria:
 - (a) attempt to match on full conference heading first; if no match then
 - (b) match on subfield a of authority record, as long as there is only a subfield a and no other subfields.
- * Do not derive series authority records (008, byte 12, code a, b, c, or z) ; journal instead.
- * Derive AACR2 and AACR2-compatible authorities (008, byte 10, code c or d); do not journal.
- * Derive pre-AACR2 authorities (008, byte 10, code a, b or z); journal.

PRELIMINARY STEP: REVIEW CURRENT LC UPDATE LOAD OPTIONS

- A. Distribute list of possible options to libraries (FCLA task).
- B. Distribute to each library its current profile of option settings (FCLA task).
- C. Review profile (library task).

D. Correct profile coding (FCLA task).

TEST AUTOMATIC DERIVE/CLAIM PROGRAM

A. Test automatic derive/claim without updating

1. Report the authorities that would have been derived, but don't derive (FCLA task).
2. Provide feedback on accuracy and completeness of results (library task).
3. Correct program (FCLA task).

** Repeat 1-3 until results are satisfactory **

B. Run automatic match/derive/claim on first section (to be defined) of MHI index

1. Derive and claim authority records (FCLA task).
2. Review results (library task).

REPORTS

A. Headings that still do not match a local authority record after automatic derive/claim.

B. Series

1. First run through of whole file (retrospective run):
 - * don't auto derive/claim
 - * report all series authorities
2. Ongoing (reevaluate in light of LC proposal to not class series separately?):
 - * auto derive/claim
 - * journal derived series authorities

C. Pre-AACR2 authorities

1. Auto derive/claim
2. Journal authorities derived

Authorities Committee Meetings Report--Appendix B

December 2, 1996

Below is an update to the authority processing statistics from the local UNF New Headings automatic deriving/claiming process. All counts come from the local programs running at UNF. Based upon the relative size of your own database and cataloging activity, these figures may help

in projecting what might be the size of your own institution's authority workflow from authorities processing being developed by FCLA.

These figures include only new headings, no retro work.

Personal and corporate names

Names used as names or used as names and subjects are included. Because of the way the local program works (actually the NOTIS indexes), names used ONLY as subjects are not captured. For topical subjects, see subject section below. These figures exclude series, uniform titles and author/title entries. We plan to include uniform titles in the very near future.

For the period, September, 1995-October 1996 (13 months):

TOTAL NEW NAMES ENTERING DATABASE: 18,030 (1387/month)

Total derived automatically: 14,790 (1138/month)

Total not matched in Itlc: 1,940 (149/month)

Total "matched" but not derived: 1,096 (84/month)

HUMAN REVIEW: 3,036 (234/month)

The 1,096 represent headings for human review, in addition to the 1,940 not found in Itlc. The headings not matched sometimes are the result of simple typos but the majority simply do not have authorities in Itlc. Of the 1,096 found but not derived, we have identified the following problems and sort our headings for review into the categories listed below. Numbers are from the breakdown of the 1,096 headings not found.

Headings which match crosses (4xxs): 101

Headings which match non-unique authorities: 189*

Headings which are "near matches": 424**

Headings which exceed the length of our program to match: 360***

Headings which match but are mistagged (100/151): 22****

*Based upon the fixed field element "Unique" in NOTIS (the element "Name" in OCLC) which identifies personal name authority records which represent more than one person.

**Near matches are headings which are shorter, for example lacking a date, than the new heading.

***Through August, 1996. This category practically eliminated by refinements to our local program.

****A new category, only begun in July, 1996.

Topical subjects

For about the same time period as above, the following were derived/claimed and/or corrected for topics. This local processing includes both main headings and heading subdivisions under name headings.

Subject headings entering the database: 21,104 (1623/month)

Derived as new from above: 2055 (ca. 10%) (158/month)

Not verified 1059 (ca. 5%) (82/month)

Corrected 3938 (ca. 12%) (303/month)

Linda L. Smith

Head, Cataloging Dept. (904) 646-2550

Thomas G. Carpenter Library FAX: (904) 646-2719

4567 St. Johns Bluff Rd. S.

University of North Florida email: lsmith@unf.edu

Jacksonville, FL 32224