

Loading and Using 2004 MeSH in MARC Format

Karen Jensen, with
notes from Pat Riva
(McGill University)

McGill's Original MeSH File

- Created in 1993/94 by Blackwells North America
- No local updates, only a few manual changes to keep the file up-to-date
- 19,000 records, including descriptor and descriptor/qualifier combinations

McGill's Final MeSH File

- 523,146 records
- Catching up with ten years of work
- Staff time spent: 3 months
- Time spent on indexing and automatic updates: 2 weeks

Benefits

- Users will find similar works under the same MeSH terms
- Reference staff will find it easier to teach subject searching
- Cataloguing staff have direct access to the latest MeSH terms
- Easy to update the MeSH file each year

Planning

2004 MeSH file contains:

- 22,568 descriptor records (150, 151, and 155)
- 83 qualifier records (180 subdivisions)
- 500,495 descriptor/qualifiers (150 \$a \$x)
- 523,146 total records (127 MB)

Descriptor/Qualifier Combinations

We decided to load these because:

- Many were in our catalogue already
- NLM issues the file as a single entity
- Faster to catalogue with them
- P_auth_03 designed to work with more authority records

Ongoing Updates

- MeSH is updated annually in the fall
- Ftp the full base and update files from NLM's Web site
- <http://www.nlm.nih.gov/mesh/filelist.html>
- Acknowledge the file source and year in your OPAC

Ongoing Updates

- In help files under Browse: *“Subject indexes in MUSE are created using a controlled vocabulary. McGill uses two major sets of subject headings: Library of Congress (LC) and the U.S. National Library of Medicine’s 2004 MeSH for health sciences subjects.”*
- P_manage_31 vs. p_manage_36

Loading the Full 2004 MeSH File

- Aleph table updates
- Sketch of initial load process
- Convert MARC records
- Modify MARC record file
- Fix & check catalog records
- Check input file against database
- Load catalog records

Indexing Considerations

- MGU11 will re-index itself online using the new index tables
- New index tables include Short-doc and see also 550 links between authority records (Z103)
- Alpha long headings (p_manage_17) ran for seven hours

Indexing Considerations

- To allow the SUM index to link to new MeSH authorities, let MGU11 Z07 (queue for records waiting to be indexed) finish before turning MGU11 ue_11 (messaging) back on
- Manage_34 took four to five days to run
- Manage_18 needs rep_change 4011 (post-patch 7) to force a Z105 to be written for each authority record

P_Manage_18

- Run p_manage_18
- Specify “Indexing = FULL” with the MGU11 ue_11 stopped and MGU11 ue_01 running
- Has MGU11 ue_01 processed all the Z07s?
- Do “select count(*) from z07;” to verify

P_Manage_18

- One MGU11 Z105 record has been created for each input authority record
- Start MGU11 ue_11
- Ensure MGU01 ue_11 and ue_08 are also running

Post-load Authority Cleanup

- Delete obsolete MeSH authorities first
- Format a list for manual review
- Review changed MeSH authorities, COR report
- Delete the COR by global change
- Publication types global change
- Conflicts and ambiguous 4xx
- Bibliographic records under a 4xx

Continuing to Clean Up the Bibliographic File

- OCLC Lacey authority processing in April 2003
- NLM changes allowable qualifiers
- P_auth_03 needed
- Problem with p_auth_03, PRB 6970
- 600/610/611/630 headings in our SUM index

Fixing Problems Created in 1999

Alternatives to p_auth_03:

- Keyword searching in Subjects
- Global changes
- *Application of MeSH for Medical Catalogers*
- <http://www.nlm.nih.gov/tsd/cataloging/catmesh.html>

What Were NLM's Major Changes?

- Form subdivisions renamed publication types
- Publication types entered as 650 \$a \$v or 655
- Obsolete form subdivisions
- Age groups and pregnancy qualifiers changed and made free-floating

Conclusion

- Easier to use bibliographic and authority records now
- LocatorPlus helps with manual review
- We still need p_auth_03 for ongoing work
- Questions are welcome:
karen.jensen@mcgill.ca and
pat.riva@mcgill.ca