

UCLA

UCLA Previously Published Works

Title

The Gravitational Pull Away from Pre-Coordinated Subject Headings

Permalink

<https://escholarship.org/uc/item/96z3s6mv>

Journal

Technicalities, 36(3)

Author

Riemer, John J

Publication Date

2016-05-01

The Gravitational Pull Away from Pre-Coordinated Subject Headings

Technicalities Column

March 9, 2016

by John J. Riemer

Head, UCLA Library Cataloging and Metadata Center

On a number of fronts in recent years, I have perceived the library community beginning to edge away from its historic commitment to pre-coordinated subject headings. I would like to use this space to provide observations and points for professional discussion of this issue.

Overtures from the National Library of Medicine

In fall of 2015, NLM announced the results of its consultation of the library community about the necessity of its continuing to distribute its cataloging records containing the Medical Subject Headings in pre-coordinated format:

“In June 2015, NLM proposed discontinuing distribution of its bibliographic records with artificially reconstructed subject strings in CATLINE and SERLINE products. This proposal was sent to the Regional Medical Libraries, MEDLIB, and MEDCAT discussion lists, as well as to subscribers of the NLM MARC21 files. Responses to the proposal were unanimously in favor of discontinuing the distribution of subject strings and having the records in the distributed files match the records as they appear in LocatorPlus and the NLM Catalog.”¹

Internally, NLM had long provided subject terms in its online catalog and in PubMed in a deconstructed format that pulls out the geographic and form facets. As part of the workflow for distributing its cataloging records externally, NLM had been running a program to assemble the pre-coordinated strings. As of December 2015, that step was being discontinued. The following example² illustrates the difference this makes:

Old form (with string)

650 12 \$a Health Policy \$z Africa \$v Congresses
650 12 \$a Health Policy \$z Caribbean Region \$v Congresses
650 22 \$a Cross Cultural Comparison \$z Africa \$v Congresses
650 22 \$a Cross Cultural Comparison \$z Caribbean Region \$v Congresses

New form (deconstructed)

650 12 \$a Health Policy
650 22 \$a Cross Cultural Comparison
651 2 \$a Africa
651 2 \$a Caribbean Region
655 2 \$a Congresses

The fall announcement concluded with this comment: “NLM wishes to thank all of those who commented on the proposal. Your input was very helpful in confirming *the direction the community is moving towards in subject analysis* [emphasis added].”³

Back in 2005, NLM had put forward a proposal to make this change. The UCLA response⁴ was probably in the majority, opposing the move. Of the four reasons provided in that response, the only one still carrying any weight was the statement that pre-coordinated subject strings were more specific and give the user better context in a subject browse display. The other arguments focused on the discrepancy between legacy and new cataloging, the burden it would require for a library to go against the grain and attempt independently to sustain the provision of pre-coordinated headings, and the potential disruptions caused to authority vendors.

As part of its 2015/2016 implementation, not only would NLM be discontinuing pre-coordination, but it would redistribute all of its past cataloging in the new format.⁵ OCLC planned to load those records and overlay existing records in WorldCat. “Following that reload, we do plan to convert existing MeSH in records in WorldCat that did not get touched by the NLM reload.”⁶ At least one authority vendor was prepared to receive legacy records from libraries and to deconstruct the Medical Subject Headings within them.⁷

The Emergence of FAST

Around the time NLM was beginning to deconstruct MeSH for its internal audience, an ALA Subcommittee on Metadata and Subject Analysis in 1999 was recommending the creation of a simplified syntax for subject elements used in metadata records for digital resources. The subcommittee’s recommendation in the report specifically mentioned the possibility of breaking out the non-topical elements from the LCSH strings.⁸

OCLC traces the origins of Faceted Application of Subject Terminology (FAST) to needs perceived during 1999-2000 OCLC Cooperative Online Resource Catalog (CORC) project. “The broad purpose of adapting the LCSH with a simplified syntax to create FAST is to retain the very rich vocabulary of LCSH while making the schema easier to understand, control, apply, and use.”⁹ During the training I received in Dublin Core data element usage from Erik Jul and others at OCLC during that CORC project, I recall FAST being touted as a happy medium between totally-uncontrolled keywords at one end of the spectrum and difficult-to-learn and -apply pre-coordinated subject strings at the other end. FAST preserved the virtue of controlled vocabulary control in its usage of topical, geographic, and form facets.

“FAST has evolved into an eight-facet vocabulary with a universe of approximately 1.7 million headings across all facets.”¹⁰ In descending order of usage, the eight facets are:

- Topics
- Geographic names
- Form/Genre
- Personal names
- Corporate names
- Chronological
- Events
- Titles¹¹

For much of FAST’s existence, the institutions applying the vocabularies did so for selected categories of material that might not otherwise receive controlled vocabulary terms.¹² FAST implementations generally represented a usage that complemented and co-existed with pre-coordination.

LCSH/FAST Benefits and Drawbacks Are Leveling Out

A seemingly radical idea has emerged in the past couple of years—the possibility of making a complete changeover to a system like FAST. Probably the stage was set for considering something like this by a gradual leveling of the advantages and disadvantages of the two approaches. The payoff for adhering to a pre-coordination approach has been declining and its upkeep is increasingly a challenge. At the same time the perceived benefits of an approach like FAST are rising.

1. The LCSH rules for combining facets into the correct order are complex. The Airlie House conference a quarter century ago aimed to rectify that with the recommended standard order of **[topic]--[place]--[chronology]--[form]**, eliminating exceptional practices.¹³ The application of free-floating lists of subdivisions is governed by the category(ies) the main topic falls under and designated

pattern headings.¹⁴ Topics that are “inherently legal” do not receive normally free-floating subdivisions like “—Law and legislation” or “—Legal status, laws, etc.”¹⁵

2. Even with LCSH’s rules for the order of facets in the pre-coordinated headings, there are limits to how much can be included in a single heading:

“LCSH is a system in which untold numbers of headings can be constructed from individual elements that represent facets, such as topic, place, time, form, and language, and various aspects of topics. Although LCSH is primarily a pre-coordinate system, practice with many complex or multi-element topics requires post-coordination in order to achieve coverage. There are numerous cases in which elements cannot be combined in single headings, even with subdivisions. In those situations, an array of headings may be assigned, that, taken together, are coextensive with the topic of an item.

“So, LCSH itself requires some degree of post-coordination of the pre-coordinated strings to bring out specific topics of works.”¹⁶

3. The traditional browse displays in catalogs built upon carefully-constructed headings are not always included in today’s discovery tools. The relatively new WorldCat Local catalogs, exemplified by the University of California’s Melvyl catalog, do not include heading browse. The facets offered for search navigation are quite popular.¹⁷

4. Even within the LCSH system, form and genre aspects of subject analysis are being pulled out from subject headings and provided as stand-alone terms. At least a dozen different controlled vocabularies now fall under the umbrella of the Library of Congress Genre/Form Terms projects, which are separate lists from LCSH.¹⁸ “It is particularly important to maintain the distinction between subject (what a work is about) and genre/form analysis (what a work is).”¹⁹

One of the most recently-developed of these vocabularies is the Library of Congress Medium of Performance Thesaurus (LCMPT) terms. The thinking has been that we should pull out this facet from

existing headings. “Medium of performance is now recognized as a separate and distinct bibliographic facet that should have its own vocabulary and vocabulary structure.”²⁰

5. Subject headings containing subdivisions added to established terms are difficult to make machine actionable. Following one of the Airlie House conference recommendations, LC and OCLC collaborated in the 1990s to explore the feasibility of creating for machine validation a file consisting of all the valid subject headings currently in use (combinations of editorially-established terms plus added subdivisions, as found in OCLC WorldCat bibliographic records).²¹

Beginning in 2007, LC’s Policy and Standards Division began to create and distribute subject authority records derived from headings found in LC’s bibliographic records. As of March 2012, 78,000 validation records had been created.²² Such records have been included in LC’s Linked Data Service.²³ For such a file to include all theoretically possible valid formulations, its size would need to expand exponentially. The unavailability of identifiers for each and every validly pre-coordinated subject heading is a definite liability in a linked data age.

6. The ease of assigning terms from a controlled vocabulary has been considerably enhanced by the SearchFAST user interface.²⁴ As a user starts to enter a term, autosuggest facilitates making a choice and authorized terms are offered in boldface. Increasingly, libraries find it desirable and necessary to enlist others without a cataloging background in metadata creation efforts,²⁵ and this guided data entry approach is enormously helpful. As examples of this reliance on others, digital resources inherently lend themselves to collaborative description efforts and the creators of research data sets are needed to provide the metadata for what is being shared.

SearchFAST has been incorporated in recently-developed metadata creation tools like OCLC's Low Barrier Metadata Creation tool and Zepheira's BIBFRAME editor.²⁶ In working with our emerging pool of collaborators, it seems much more practical to opt for a system that does not involve complex pre-coordination.

FAST as Potential Replacement for Pre-coordinated Vocabulary Usage

At the most recent ALA Midwinter Meeting in Boston, I sat in on a discussion on ways to sustain the viability of a vendor's continuing to provide MARC cataloging for monographs acquired by partner libraries. The pair of ideas geared toward lowering the production cost of the metadata included scaling back on name authority work (set aside since it would undercut a linked-data future) and changing from pre-coordinated LCSH to FAST. The vendor was going to explore the latter suggestion. That move is something Stanford University is "actively considering" as well.²⁷

The British Library has been studying how it should expand the usage of FAST in its operations since late 2014 and has recently brought forward recommendations.²⁸ After a small team of catalogers tested FAST and abridged Dewey on a range of publications in various languages over four months, the first two recommendations are:

1. To adopt FAST selectively to extend the scope of subject indexing of current and legacy content.
2. To implement FAST as a replacement for LCSH in all current cataloging, subject to mitigation of the risks identified.

The Library states that it currently "is consulting with stakeholders concerning the potential impact of these proposals. *No firm decisions have yet been taken regarding either of these standards.*"²⁹

Implementation Issues

The major risk the British Library identified was that FAST is currently being maintained by OCLC as a research project and not a service we can rely on to be sustained. This seems like something that could be remedied in our community, particularly if other institutions indicate they are seriously thinking of adopting the BL's second recommendation.

It is often noted that FAST is completely derived from LCSH. Shouldn't the Program for Cooperative Cataloging explore a future of SACO work that includes making it possible to propose new terms for either or both vocabularies?

There seems to be a low risk that libraries would get caught between a newly-embraced FAST practice and a large quantity of legacy LCSH data. Beginning in September 2013, OCLC has made a systematic pass throughout all of WorldCat to add the FAST heading equivalents to all records containing LCSH.³⁰ An estimated 100 million records would be affected. Cornell is one institution that has carried out the same process locally in order to better support faceting in its discovery system.

Conclusion

Some proposals for significant changes in the library community bring with them major controversy and a real sense of loss of something valuable. Is this true for the long-standing practice of pre-coordinated subject strings? I am anticipating more of us will be asked in the near future to think about and respond to proposals to change over to FAST, and we will have to decide whether that warrants a strenuous objection or little more than a ripple of concern. Before that discussion comes to you, I wanted to offer the above background information and perspective.

1 Diane Boehr, Head of the Cataloging and Metadata Management Section, NLM, writing September 10, 2015 to the MEDCAT and MEDLIB-L discussion lists.

2 Diane Boehr, email message to PCCLIST, November 30, 2015.

3 Boehr announcement, September 10, 2015.

4 Cynthia Shelton, AUL for Collection Management and Scholarly Communication, memo to Diane Boehr, August 15, 2005. "UCLA strongly prefers that NLM continue its current practice of distributing precoordinated subject strings in bibliographic records."

5 Diane Boehr. "Change to Subject Headings in Distributed NLM CATLINE and SERLINE Products," *NLM Technical Bulletin*, no. 406 (September-October 2015)
https://www.nlm.nih.gov/pubs/techbull/so15/so15_cataloging_unstringing_subject_heading.html

6 Cynthia M. Whitacre, OCLC's Manager, WorldCat Quality, email correspondence to Valerie Bross, UCLA Library ERM/Continuing Resources team leader, October 27, 2015.

7 "Backstage Library Works launches new service: Medical Subject Headings (MeSH) deconstruction" March 1, 2016 press release, in *Library Technology Guides*, <http://librarytechnology.org/news/pr.pl?id=21365>

8 American Library Association, Association for Library Collections & Technical Services, Cataloging and Classification Section, Subject Analysis Committee, Subcommittee on Metadata and Subject Analysis. "Subject Data in the Metadata Record: Recommendations and Rationale." July 1999. See section 3.2.3. Syntax.
http://www.ala.org/alcts/resources/org/cat/subjectdata_record

9 OCLC Office of Research. "FAST (Faceted Application of Subject Terminology)" web page, updated as of August 23, 2013. <http://www.oclc.org/research/themes/data-science/fast.html>

10 OCLC FAST web page.

11 Mixer, Jeffrey, and Eric R. Childress. FAST (Faceted Application of Subject Terminology) Users: Summary and Case Studies. Dublin, Ohio: OCLC Research, August 2013. Figure 2. FAST facets use by adopters, page 10.
<http://www.oclc.org/content/dam/research/publications/library/2013/2013-04.pdf>

12 Mixer and Childress. FAST Users, Table 1. Agencies known to have adopted FAST, page 14.

13 Subject Subdivisions Conference, 1991, Airlie, Virginia. *The Future of Subdivisions in the Library of Congress Subject Headings System: Report from the Subject Subdivisions Conference May 9-12, 1991*, edited by Martha O'Hara Conway. Washington, D.C.: Library of Congress, Cataloging Distribution Service, 1992. (The text of the conference's original recommendations can be seen online in Library of Congress. Five-year Progress Report on Subject Subdivisions Conference Recommendations, December 1996. <http://www.loc.gov/catdir/marvel/subdiv96.html>)

14 Library of Congress. Subject Headings Manual, section H 1146, Subdivisions Controlled by Pattern Headings. "The principle of free-floating subdivisions controlled by pattern headings was officially incorporated into **Library of Congress Subject Headings** in 1974." "**General rule.** Use any subdivision that is established under a pattern heading as a free-floating subdivision under any other heading belonging to its category if it is appropriate and if no conflicting

heading is established in the subject authority file in either phrase or main heading-subdivision form.”

15 LC Subject Headings Manual, section H 1705, Legal Materials: Law and Legislation and Other Subdivisions. Thus, Domestic relations—Law and legislation; Families—Law and legislation; and Aliens—Laws, status, etc. are all invalid formulations.

16 Library of Congress Cataloging Policy and Support Office. “Library of Congress Subject Headings: Pre- and Post-Coordination and Related Issues, March 15, 2007, revised December 20, 2007, page 5.
https://www.loc.gov/catdir/cpso/pre_vs_post.pdf

17 OCLC. “Features of WorldCat Local,” <https://www.oclc.org/en-US/worldcat-local/features.html> and “WorldCat Local at a Glance,” <https://www.oclc.org/en-US/worldcat-local/about.html> (includes “Faceted browsing by author, format, year, audience, topic and more”)

18 Library of Congress Policy and Standards Division. Frequently Asked Questions about Library of Congress Genre/Form Terms for Library and Archival Materials (LCGFT), revised June 6, 2011. The initial set of vocabularies is listed on page 2. https://www.loc.gov/catdir/cpso/genre_form_faq.pdf

19 LCGFT FAQ, response to Q6, page 3.

20 Library of Congress. “Medium of Performance Terms for Library of Congress Medium of Performance Thesaurus for Music: Draft Authorized Vocabulary for Medium of Performance Statements in Bibliographic Records Agreed on by the Library of Congress and the Music Library Association,” February 2013, page 1. “Thus, for music, LCSH will continue in the direction begun with the LCGFT initiative, to represent exclusively what a musical resource is about, leaving what the musical content of the resource is to LCGFT [Library of Congress Genre/Form Terms for Library and Archival Material], and what the medium of performance is to the new medium of performance thesaurus.”
http://www.loc.gov/catdir/cpso/medprf_lcmla.pdf

21 Lois Mai Chan and Diane Vizine-Goetz. “Toward a Computer-Generated Subject Validation File: Feasibility and Usefulness.” *Library Resources & Technical Services* v. 42, no. 1 (January 1998), page 46
<http://www.ala.org/alcts/sites/ala.org.alcts/files/content/resources/lrts/archive/42n1.pdf>

22 Library of Congress. *Library of Congress Subject Headings*, 34th edition, 2012, page xv. Available online via Google Books <https://books.google.com/books?id=5nAqnr3gsNUC>

23 Library of Congress. “Library of Congress Subject Headings.” Modified April 26, 2011.
<http://id.loc.gov/authorities/subjects.html>

24 OCLC Office of Research. SearchFAST <http://fast.oclc.org/searchfast/>

25 Karen Calhoun. “CORC and Collaborative Internet Resource Description: a New Partnership for Technical Services, Collection Development and Public Services, in Karen Calhoun, John J. Riemer, editors. *CORC: New Tools and Possibilities for Cooperative Electronic Resource Description* (New York: Haworth, 2001) pages 131-142. Also published as *Journal of Internet Cataloging* 4(1/2) 2001. (See Figure 2, page 139, showing “Context of a widely distributed internet resource description system)

26 ALA Technical Services Directors of Large Research Libraries Interest Group, 2015 Midwinter Meeting notes, Friday, January 30, 2015, page 4: “Subject entry is supported by auto-suggested FAST.”

<http://connect.ala.org/node/238409>. In Zepheira's BIBFRAME editor, FAST is one of the "external services" supporting data entry <http://editor.bibframe.zepheira.com/static/>.

27 Phil Schreur, Assistant University Librarian for Technical and Access Services, communication over ALA "Big Heads" listserv, February 1, 2016.

28 British Library. FAST/Dewey Review Group. "Consultation on Subject Indexing and Classification Standards applied by the British Library," April 2015 <http://www.bl.uk/bibliographic/pdfs/british-library-consultation-fast-abridged-dewey.pdf>

29 FAST/Dewey Review Group report, page 3.

30 Enriching WorldCat with FAST [2013?] (6 pages)
<https://www.oclc.org/content/dam/support/worldcat/records/faceted-application-of-subject-terminology.pdf>