

Library of Congress Subject Headings **Pre- vs. Post-Coordination and Related Issues**

*Report for Beacher Wiggins, Director, Acquisitions & Bibliographic Access Directorate,
Library Services, Library of Congress*

Prepared by the Cataloging Policy and Support Office

March 15, 2007

*(Approved by ABA Management June 13, 2007, with annotations on Recommendations [p. 9-13] added
October 2, 2007, revised December 20, 2007)*

The Director for Acquisitions and Bibliographic Access at the Library of Congress requested a review of the pros and cons of pre- versus post-coordination of *Library of Congress Subject Headings* (LCSH). There have been several studies of the pros and cons over the years, most notably by Elaine Svenonius and more recently by Lois Mai Chan and Arlene Taylor/Daniel Joudrey – included as appendices to this report. Briefly stated, pre-coordinated strings provide context, which is needed for “disambiguation, suggestibility, and precision”¹ and browsability. Pre-coordinated strings have a sophisticated syntax that can express concepts better than single words, yet also can be faceted by systems to group topics into categories for post-coordinated displays when desirable. Post-coordinated terms have serious limitations for recall, precision, understanding, and relevance ranking.

Rather than review the issue of pre- versus post-coordination, the issue perhaps should be how subject access can best be achieved for users at a cost that reduces the current expense. We know subject access is important to users and a value added service that libraries provide. Additionally, Karen Markey has recently pointed out that subject headings should figure prominently in the post-Boolean probabilistic catalog, as they contribute to relevance feedback and ranking algorithms.² There are certainly ways to reduce the costs for subject cataloging while retaining the benefits of the pre-coordinated strings of *Library of Congress Subject Headings*, and the Cataloging Policy and Support Office offers recommendations towards that end.

Attached (starting on p. 15) are the following background papers and examples:

Appendix 1: Lois Mai Chan’s “Thoughts on LCSH” of July 31, 2006

Appendix 2: Arlene G. Taylor and Daniel N. Joudrey’s paper sent July 17, 2006 via email to Barbara Tillett, “LCSH Strings – some thoughts”

Appendix 3: Excerpt on “Precoordination or not” from Elaine Svenonius. *The Intellectual Foundation of Information Organization*. Cambridge, Mass.: MIT Press, 2000, p. 188-192.

Appendix 4: Airlie House Conference Recommendations and Progress in Carrying Them Out

¹ Svenonius, Elaine. *The Intellectual Foundation of Information Organization*. Cambridge, Mass.: MIT Press, 2000, p. 189.

² Markey, Karen. “The online library catalog: paradise lost and paradise regained?” *D-Lib*, v. 13, no.1/2 (Jan./Feb. 2007)

Appendix 5: FAST (Faceted Application of Subject Terminology)

Appendix 6: Difference between Google results and a browse display of LCSH;

Comparison of search results of a post-coordinate versus pre-coordinate approach

Appendix 7: Examples of Deconstructing LCSH into Simple Facets

Appendix 8: LC Staff Survey on Subject Cataloging Report

Introduction

The Library of Congress Subject Headings (LCSH) are by far the most widely adopted subject indexing language in the world.³ Since 1898 LCSH has provided a set of terms for a comprehensive, broad range of topics. LCSH has been translated into many languages⁴ and is used around the world by libraries large and small.

It has been both praised and criticized for over a century. Two recent reports have expressed the notion that LCSH should be stopped.⁵ It is claimed that

- it takes too long to train anyone to correctly apply the complex rules associated with LCSH pre-coordinated subject strings,
- that the specific terms and text strings are not understood by end-users (or perhaps even reference librarians or catalogers themselves),
- that subject terms should be automatically provided by systems based on an assigned classification number, and
- that the Library of Congress is too slow to incorporate new terms.

Earlier criticisms focused on

the level of specificity (LCSH is too specific and many of the specific terms are only used once in the OCLC database of many millions of records – users would never find the specific term),

the syndetic structure (LCSH is not rigorous enough for a true thesaurus – and as a result of this criticism, attempts were made in the past to add the thesaural structure of BT, NT, RT to LCSH),

the form of headings and subdivision practices (LCSH is too complicated and inconsistent – and the Airlie House recommendations that addressed this are described further below and in Appendix 4),

lack of currency and existence of prejudices reflected in the terminology, (LCSH has too many outdated terms, not keeping up with current terminology, and has too many terms that reflect prejudice – the response has been for LC to get assistance from many partners in cooperative programs and to respond to recommendations from individuals and groups),

the difficulty of use (LCSH requires an apprentice period to properly learn the principles, patterns, and rules),

etc.⁶

³ Anderson, James D. and Melissa A. Hofmann. “A fully faceted syntax for *Library of Congress Subject Headings*,” *Cataloging & Classification Quarterly*, v. 43, no. 1 (2006), p. 7-38.

⁴ It is the foundation for subject heading systems worldwide in such languages as Lithuanian, Spanish, Japanese, French (both as spoken in France and Canada), etc., either as directly translated or as the source for an ongoing system, (i.e., RAMEAU and Laval’s RVM) and was chosen as the base vocabulary behind OCLC’s FAST project. It is a key vocabulary in the cross-thesaural effort, MACS, and has been the object of many research projects for linking controlled vocabularies.

⁵ University of California Libraries. Bibliographic Services Task Force. Rethinking How We Provide Bibliographic Services for the University of California. Final report: December 2005.

Calhoun, Karen. “The Changing Nature of the Catalog and its Integration with Other Discovery Tools” prepared for the Library of Congress by Karen Calhoun. Final report March 17, 2006.

⁶ Kirtland, Monika and Pauline Cochrane. “Critical views of LCSH-Library of Congress Subject Headings. a bibliographic and bibliometric essay,” *Cataloging & Classification Quarterly*, v.1, no. 2/3 (1982), p. 71-94.

LC's own internal staff survey also highlighted many of these same issues, such as the need for more training, difficulty in interpreting the manuals, and the need for clearer documentation.⁷ Adjustments by the Library of Congress over the years to try to address these criticisms have been incomplete primarily due to the resources available⁸ or due to the community disagreement on direction.⁹

Rather than being based on comprehensive principles (although principles were later documented)¹⁰ or covering the universe of knowledge, the LCSH terms are based on "literary warrant" of the materials cataloged by first the Library of Congress and now also by its partners in SACO (Subject Authority Cooperative of the Program for Cooperative Cataloging). Terms are proposed by catalogers and editorially reviewed by the subject policy specialists in LC's Cataloging Policy and Support Office. Literary warrant has also been both praised and criticized: praised for the ability to reflect the topics of materials being added to library collections and criticized for lacking a universal knowledge structure to anticipate topics or for lacking a way for users to easily contribute to the terminology. The latter issue could be somewhat resolved through the augmentation of systems with complementary social tagging mechanisms, also known as folksonomies or "collabularies."

Pre-coordination or Not?

The need for syntax in a subject language has been studied at least since 1951. The findings of Mortimer Taube's empirical evaluation that year about "precoordination or not," were eloquently described by Elaine Svenonius with her summarization of the pros and cons of costs and benefits (see excerpt in Appendix 3). At the end of her book on the "Intellectual Foundation of Information Organization" she stresses that context (precoordination) "is needed for disambiguation, suggestibility, and precision."¹¹ She goes on to say "the relative uniformity of a subject-language syntax and the type of specificity it provides allow it to create better browsing domains than a natural-language syntax... The syntax used in subject languages offers a means to organize large retrievals conceptually."¹² "A subject-language syntax, then, is superior to a natural-language syntax insofar as it can organize large sets of citations by layering and chunking."¹³

Shubert, Steven Blake. "Critical views of LCSH-ten years later: a bibliographic essay," *Cataloging & Classification Quarterly*, v. 15, no. 2 (1992), p. 37-07.

Fischer, Karen S. "Critical views of LCSH, 1990-2001: the third bibliographic essay," *Cataloging & Classification Quarterly*, v. 41, no. 1 (2005), p. 63-109.

⁷ *LC Staff Survey on Subject Cataloging Report*, submitted by Kay Ritchie and Rosa Alicea, October 25, 2006. (See Appendix 8)

⁸ An example is the change to use thesaural constructs of broader term (BT), narrower term (NT), and related term (RT) that was done programmatically and continues to be adjusted manually as problems are encountered by staff and individual headings are reviewed.

⁹ An example is disagreement on across the board application of the Airlie House recommendation on a consistent order of subdivisions. The community declared exceptions for such things as "history".

¹⁰ Chan, Lois. *Library of Congress Subject Headings: principles of structure and policies for application*. Annotated version. Washington, D.C.: Cataloging Distribution Service, Library of Congress, 1990.

¹¹ Svenonius, p. 189.

¹² Svenonius, p. 190.

¹³ Svenonius, p. 191.

Anderson and Hofmann point out that “A post-coordinate vocabulary of single-concept terms may be effective for electronic searching, but it would be a disaster for browsing, a long-time strength of LCSH. Browsing single concept terms would be like using a terrible back-of-the-book index without subheadings, so that under most headings, the number of references of items would be enormous and therefore difficult if not impossible to peruse....With a growing body of research that recognizes that browsing can be the most effective means for accessing useful documents in some situations, this is no time for LCSH to abandon one of its strongest contributions to effective information retrieval.”¹⁴

The Role of Pre-coordination and Post-coordination in LCSH

Pre-coordination is the combining of elements into one heading in anticipation of a search on that compound heading. Post-coordination is the combining of headings or keywords by a searcher at the time he/she looks for materials in a catalog. 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.

National Coordinated Cataloging Program (NCCP) and Airlie House Recommendations

By the late 1980s, cooperative cataloging programs had stimulated new interest in an old topic: cataloging simplification, both descriptive and subject. Of particular concern to participants in the National Coordinated Cataloging Program (NCCP) was the task of building individual headings by adding free-floating subdivisions. As a direct response to that concern, and in keeping with ongoing efforts to improve the cataloging process, the Library of Congress sponsored an invitational conference at Airlie House, Virginia, in 1991 to address subdivision practice in LCSH. Six recommendations emerged from the debate and discussion of the position papers prepared for the conference. The conference’s endorsement of

- *pre-coordinated LCSH subject strings, and
- *indirect geographic subdivision as well as the conference's recommendations
- *to institute a standard order of subdivisions,
- *to subfield code form subdivisions, and
- *to standardize and simplify subdivision practice

set the agenda for the subject heading revision that has taken place since 1991. The conference recommendations are detailed along with statements of the progress that has been made in carrying them out in Appendix 4 to this report.

¹⁴ Anderson and Hofmann, p. 9-10.

Airlie House suggested a regularized standard order of subdivisions or facets under topical headings:

Topic – Place – Time – Form

Although this order was recommended in general, it was later shown not to work in all fields, and that meaning would be lost, especially in art, literature, music, history, and law. However, later studies showed there was no increase or decrease in understanding by changing the order.¹⁵ It is widely acknowledged that the logic of the LCSH strings is lost on most users and some catalogers,¹⁶ and yet, the syntax given to facets in a pre-coordinated string gives expressive meaning that is important to the understanding and sorting out of concepts. Even though it is an artificial syntax, it is suggestive of meaning, especially once a person learns how to “read” the string - the meaning is very useful.

Alternatives?

FAST. One suggested alternative, in some of the reports saying to abandon LCSH, was FAST (Faceted Application of Subject Terminology), which is being developed by OCLC. The Library of Congress has participated in and continues to support the research efforts that are developing FAST at OCLC. Yet, **FAST is built on LCSH and requires the LCSH process and system to bring in new terminology.** FAST is not totally post-coordinated, as it was discovered some terms (especially topical concepts) would indeed require pre-coordination to provide context to the phrase. FAST has introduced non-standard formulations of uniform titles and geographic names contrary to international cataloging rules and has its own set of rules. (See Appendix 5 for more information about FAST.) The attraction of post-coordinated subject terms seems to be based on the assumption that such systems are easier to apply and result in just as good or better retrieval for end users. There is no evidence that this is true – it has yet to be proven.

New search engines and display capabilities. Marcia Bates has often reminded us that users don’t know the controlled terminology when they begin a search and that they can recognize better than recall terminology.¹⁷ Pre-coordination can display to users different strings in which their “keyword” terms are embedded. Systems like Endeca can mine the terms and show users the universe they are searching in meaningful categories (facets) and offer pathways to trigger ideas of where to find what the user wants. This sort of search engine and display capability takes the best advantage of the pre-coordinated facets of LCSH in combination with all the other subject rich terminology found in the bibliographic records, authority records, classification records, and full-text documents when available (as attached tables of contexts, abstracts, summaries, or the full-text of the resource itself). This can be complemented by the inclusion of social tagging terms (folksonomies) that are provided by end users and linked to the

¹⁵ Franz, L, J. Powell, S. Jude, and K. M. Drabenstott. “End user understanding of subdivided subject headings,” *Library Resources & Technical Services*, v. 38, no. 3 (1994), p. 213-226.

¹⁶ Drabenstott, Karen Markey, Schelle Simcox, and Eileen G. Fenton. “End-user understanding of subject headings in library catalogs,” *Library Resources & Technical Services*, v. 43, no. 3 (1999), p. 140-160.

¹⁷ Bates, Marcia J. “Task Force Recommendation 2.3 Research and Design Review: Improving User Access to Library Catalog and Portal Information; Final Report (Version 3). June 1, 2003. <http://www.loc.gov/catdir/bibcontrol/2.3BatesReport6-03.doc.pdf>

bibliographic records. We can now do this in the MARC 21 format in 653 fields for catalogers to add such data, but more sophisticated techniques with a new OPAC front-end could enable end user input. (See Appendix 2 for more about Endeca capabilities.)

Recent Analyses of Pre- vs. Post-coordination (Pros and Cons)

Lois Mai Chan, Elaine Svenonius, and Arlene Taylor/Danny Joudrey have all written on retaining the Library of Congress Subject Headings pre-coordinated approach (see the Appendices 1, 2, and 3 for texts). They and others have described the pros and cons of pre- versus post-coordination of the Library of Congress Subject Headings. They have suggested ways to simplify the application of pre-coordinated terms and ways to reduce costs of assigning the LCSH subdivisions and terms using automated capabilities – many of which are already available to catalogers, if they would use them, such as the Class Web ability to suggest terms and classification numbers relevant to a topic.

Advantages (Pros) of Pre-coordination

- 1) Flexibility. Pre-coordinate terminology is more flexible in terms of system design. Pre-coordinate indexing can be disassembled for use in post-coordinate systems. Pre-coordinate indexing can be useful for either browse or keyword searching. Post-coordinate indexes are not as useful for browse searching, and individual facets cannot reliably be reassembled to offer syntax or relationships among the searched terms.
- 2) Library Community Support. The NCCP and Airlie House conference specifically recommended that a new “v” subfield for form subdivisions be created, implying that libraries find pre-coordination helpful. This is also supported by Chan, Taylor, Markey, and others noted in the footnotes of this paper.
- 3) Keywords. Pre-coordinate indexing is advantageous even for keyword searching since the pre-coordination provides a proximity feature that can be useful. Terms separated in post-coordinate indexing supply little information about relationships for proximity searches.
- 4) Indicating General Works. Pre-coordinate terminology enables clearer indication of general works. Persons seeking introductory or general works on a topic can use a browse search to find the unsubdivided heading. With a post-coordinate system general and specific works on a topic usually would be displayed together without distinguishing qualifications or subdivisions.
- 5) Relationships. Pre-coordinate indexing enables clearer indications of relationship of topics when that is the point of the work, for example, the commercial relations of France with Germany and Italy. In LC’s current pre-coordinate system this relationship can be clearly expressed, whereas a post-coordinate system would likely display together commercial relations of all countries with relations with either France or Germany or Italy.
- 6) Hierarchical Displays. Pre-coordination of terms enables hierarchical displays for improved browsability.

7) Standard Order Gives Meaning. The consistent use of a standard order to pre-coordinated strings in itself gives meaning to the words used.

8) Relevance of a retrieved search result. Pre-coordinated subject headings can be used to improve the relevance ranking of systems like Google. David Bade of the University of Chicago in a recent posting on the ALA subject headings listserv stated : "... we would do well to point out to other librarians and system designers that subject headings, especially pre-coordinated headings based upon a particular thesaurus and system (i.e. LCSH), have far greater potential for relevant retrieval than any algorithm can or ever will. The chief reason for that is that subject headings are the product of human judgments of exactly the same nature as the relevance judgments made by the searcher, while relevance ranking a la Google and Endeca can never be anything other than statistical analyses of probabilities based upon surrogates."¹⁸

Disadvantages (Cons) Of Pre-coordination

- 1) Too complicated. Lois Chan (Appendix 1) pointed out that pre-coordination requires manual indexing, that is, humans must devise the strings based on elaborate rules for heading construction. We do not yet have a full authority file with all strings that have been used that could be employed for machine-validation and to assist in assignment of heading strings. It takes "too long" to train anyone to correctly apply the complex rules associated with LCSH pre-coordinated subject strings. In addition, the forms of headings and subdivision practices are inconsistent.
- 2) Not easily understood by end-users. Pre-coordination assumes that users understand and can apply the correct syntax when searching. However, anecdotal evidence of users' search patterns suggests that most do not possess this skill. The rules for cross-references - designed to help users find the resources they need - may actually work against the user in some circumstances. Since references are typically provided on only the main heading, and not on the possible subdivisions of non-preferred terms, those who attempt to use subdivisions but who do not know the exact syntax may not be successful in their searching.
- 3) Too expensive. Ms. Chan went on observe that the process of maintaining and verifying pre-coordinated terms is costly, even when those costs are reduced through cooperative efforts.
- 4) Less flexible. Lois Chan stated that pre-coordinated terms are less flexible, presumably because the syntax must be considered. She also observed that there is little value in pre-coordinated strings when browsing is not an option, because users typically do not search by strings, but rather by keywords.

¹⁸ Bade, David. [Response to the [headings@ala.org] listserv discussion, March 12, 2007 with the subject line: "Subject headings as relevance ranking"]

Recommendations of the Cataloging Policy and Support Office

Based on the recommendations from these outside experts as well as in-house expertise, the Cataloging Policy and Support Office recommends the following:

1. **Continuation of pre-coordinated strings.** In order to benefit from the context provided by pre-coordination as well as to maximize the potential for post-coordinated access, continue to assign pre-coordinated subject strings. However, we also recommend incorporating more automated assistance and simplified application rules as described below in further recommendations (also suggested by Chan (see Appendix 1)).

ACTION: CPSO will continue to develop and maintain LCSH for pre-coordinated assignment and continue to recommend automated ways to facilitate clustering, guided searching, faceting on the fly, and machine validation. This includes

- providing the *Subject Cataloging Manual: Subject Headings* to give guidance to catalogers,
- reviewing those guidelines for increased consistency and ease of application;
- continuing the weekly editorial meetings for an open forum to discuss decisions on proposals for new and changed subject headings and related classification numbers;
- assuring editorial oversight of the controlled vocabulary for consistency and maintenance of the syndetic structure;
- encouraging more users of LCSH to submit suggestions for updates and corrections to existing terms and additional lead-in terms as references using the Web form;
- maintaining the authority and bibliographic records that change as a result of approved proposals;
- further developing automated authority record generation and validation to simplify the cataloger's effort and to improve accuracy for new subject headings assigned.

CPSO also will continue to apply the agreed recommendations from the Airlie House conference (see Appendix 4) on new and corrected subject heading strings.

As for a post-coordinated, faceted approach, we expect faceting capabilities from future search engines, taking advantage of mining the LCSH tagged strings. LC does not plan to assign FAST headings to LC records. FAST requires LCSH as its base, so LC needs to continue the LCSH process to maintain the vocabulary. It would be inefficient to introduce an added process for the assignment of FAST subject strings that requires a set of completely different rules (including a non-standard formulation of uniform titles entered under personal author and formulating geographic headings based on Geographic Area Codes rather than cataloging rules). The assignment of FAST terms would take more of LC catalogers' processing time, and any algorithmically generated addition of FAST terms would increase LC's record distribution steps without any demonstrated user benefit or demand. However, the FAST project did confirm the usefulness of strings of topical headings with free-floating topical subdivisions (as was recommended in the Airlie House report – see Appendix 4), and LC is now meeting this need by distributing subject validation authority records (see recommendation 3.b., below). LC will continue to support OCLC's development of FAST for the potential benefit to other libraries.

Other simplification of application of rules and automated assistance are addressed in other recommendations below.

2. **A new search engine front end to the ILS.** Explore and implement a new search engine front end that offers “guided search” or “clustered searching” capabilities to mine existing terminology to augment and lead users to the controlled vocabulary of LCSH and LCC. Also enable users to add their own access terms or “folksonomies” to bibliographic records to augment the uncontrolled access. Reference librarians especially should be encouraged, if not required, to contribute to this effort.

ACTION: The Acquisitions and Bibliographic Access Directorate (ABA) requested approval to use CDS FY08 funds to conduct a Request for Information (RFI) to compare front-end search engines and will continue to pursue this test. Additionally, during the fall of 2007, CPSO represented ABA in a library-wide collaboration with the Office of Strategic Initiatives (OSI) for exploring search and access on LC’s Web site, including innovative subject searching capabilities. CPSO also collaborated with OSI in supporting academic research into visualization methods for subject access and will continue to encourage such research.

As for the potentially complementary social tagging opportunities, CPSO will continue to explore such capabilities with OSI and the ILS Program Office. CPSO also will encourage reference librarians and SACO partners to contribute terminology to use as references to existing LCSH terms to expand front-end, entry-vocabulary to controlled terms.

3. **Simplification of cataloger’s work.** Continue to explore ways to simplify the cataloger’s work in devising subject heading strings, e.g.,
 - a) **Automate LCSH term assignment further:**
 - Increase the use of Class Web to suggest previously used subject headings and class numbers for a topic;

ACTION: Class Web provides automated support to catalogers in subject heading assignment based on past precedents in the LC bibliographic database. Following the anticipated hiring of a Class Web coordinator for CDS, that person will work on reminders to LC staff, training (such as the Instructional Development and Training Division’s ‘Skill Builders’ on Class Web), and additional promotion of Class Web both within and outside LC.

CPSO will continue to include instruction on using the automated LCSH proposal forms as part of LCSH training classes offered through LC’s Instructional Design and Training Division (IDTD). These forms simplify the cataloger’s effort and feed into the maintenance and building of the LCSH correlations in Class Web between subject headings and class numbers (both LCC and Dewey).

CPSO will continue a project to add class numbers to subdivision authority records and other subject authority records as appropriate to enhance the Class Web capabilities.

- Implement software to suggest subject headings and class numbers for all digital texts.

ACTION: In addition to enhancing the Class Web capabilities, CPSO will work with

David Williamson, senior automation specialist for ABA, to explore metadata generation software such as used by the World Bank and experimental systems being studied through the University of North Carolina, School of Information and Library Science's Metadata Research Center; to test and hopefully implement such a capability during FY08.

- Use LCSH to assist with LCC assignment and vice versa. Expand and systematize projects to add class numbers to subject headings that have corresponding class numbers to enable a first step towards machine-assisted assignment of subject strings.

ACTION: Class Web is built with a feature that correlates LCSH and LCC very well. However, not all subject authority records have the necessary classification data for such correlations. To add the class numbers to subject authority records requires manual effort and expertise, but we anticipate a great payoff for future catalogers. Using CPSO methodology, this effort will be assigned part-time to the Database Improvement Unit (DBIU) to further test and to propose a project for funding during FY08. CPSO also proposes to work with the Coop Team to expand the MLA and Music Funnel for SACO to add appropriate class numbers for music. In addition, once the CDS Class Web Coordinator is hired, CPSO will explore how to further automate these capabilities.

b) Automate the verification and validation of strings further:

- For heading and free-floating combinations for popular and frequently assigned headings, provide "validation records," so that machine validation can be performed. These combinations would be coded as non-print, so that the 5-volume print *LCSH* would not become intolerably large. The resulting "validation records" would assist catalogers in subject heading assignment, improve accuracy, confirm valid subdivision combinations/patterns, thereby making future subject heading assignments easier.

ACTION: As for increasing automated generation of authority records for validation of subject heading strings, successful implementation began summer 2007 with SHED staff and programs from David Williamson. CPSO expects to increase the number created each week to approximately 5,000 (it is estimated that several million will eventually be created). These "validation records" will not appear in the printed LCSH "Red books" but will be available in the MARC Distribution Service so other libraries can benefit from the machine-validation potential. The "validation records" are clearly identified by the presence of a 667 field that reads: "Record generated for validation purposes." Notification from CDS was sent to customers in the summer of 2007. As of December 6, 2007, the initial test set of 3,900 records was created and distributed.

c) Build more coded structures within LCSH to enhance the ability for systems to suggest terms:

- Code many subject authority records so that headings can be matched with an appropriate free-floating subdivision list to simplify the construction of subject heading strings and allow for automatic or computer-assisted assignment of subject strings and

machine validation. For example, code the heading ‘Cotton’ as a plant/crop, so that automated systems could provide a linked list of subdivisions usable under plants and crops for perusal by the cataloger. This would involve implementing the 072 Subject Category Code in authority records. (See Chan, *Thoughts on LCSH*, p. 6: “In a 1990 study of 4,233 assigned subject headings, only 590 (13.9%) were enumerated in LCSH; the rest were constructed by catalogers.”)

ACTION: Gary Strawn at Northwestern has done a similar effort and would be contacted to collaborate on a test for LC. Once the CDS Class Web Coordinator is hired, that person will work with David Williamson to explore how to automate this. CPSO will use detailees to target new categories that are needed and to experiment with “Plants” as a start.

- d) **Remove subdivision by language** for all subject headings except Dictionaries, e.g., for Catechisms, where we use Catholic Church--Catechisms--French, thereby removing instances of ‘topical’ subdivisions after forms, and bringing this in line with Airlie House recommendation on the order of subdivisions. Retrieval by language would rely on the language coding in the bibliographic record (041 field and the 008, positions 35-37) rather than on the subject heading subdivision. *This is dependent on systems having the capability to guide users to options to limit a search by language of the resource.*

ACTION: LC’s Integrated Library System (ILS) will soon have enhanced filtering capabilities for both the OPAC and the cataloging module. CPSO will coordinate a test with LC’s reference librarians to check the impact of this recommendation. Before any change in policy, CPSO will prepare notices to users and seek comments and suggestions during FY08. Part of the test also will include checking for any other categories than “dictionaries” that should remain an exception where language is an essential component of the meaning of the subject term.

- e) **Explore giving some form subdivisions** in a separate 655 field as a form heading rather than redundantly with each subject heading, e.g.,

Title: A tourist map of Annapolis.
651 Annapolis (Md.)
655 Maps, Tourist.

The negative impact of this recommendation is loss of collocation by form under the topic in browse displays, so this needs further exploration before deciding to implement it. If we proceed with this recommendation, we would need to:

- identify the form headings that currently exist in LCSH and create authority records for them with the MARC 21 tag 155. This process would be a major undertaking but would allow the creation

of a form headings list and serve to validate such headings in machine environments

- based on the form subdivision list, move form subdivisions to form headings and create authority records for them with the MARC 21 tag 155

ACTION: CPSO expects to submit a report with recommendations for further actions in January 2008 on LC's Genre/form explorations that will end in December 2007. In response to requests from the Motion Picture, Broadcasting and Recorded Sound Division (MBRS), CPSO created approximately 200 genre/form subject authority records tagged 155 as of December 6, 2007. This was a test to integrate into LCSH the genre/form headings from both the *Moving Image Genre-Form Guide (MIGFG)* and *Radio Form/Genre Guide (RADFG)* used by MBRS. Further expansion will depend on decisions noted directly above. At the same time, the Music Library Association has expressed interest in having LC take the lead in identifying the issues related to genre/form access to music material, so music explorations also will be part of the January 2008 report. The recommendations resulting from discussion of that report will be distributed widely for further comment and reaction both within the Library and with the library community at large. The recommendations will also be discussed with appropriate professional organization subject cataloging units (programs already scheduled with ALA/SAC, OLAC, and with AALL during FY08), prior to final LC management approval of any changes in policy.

4. **Follow-up on the August 2006, survey of LC's subject cataloging staff recommendations** for simplifications and improvements to assigning LCSH terms. (See Appendix 8)

ACTION: Many of the recommendations from the August 2006 staff survey are already in place and others are in line with the above recommendations.

5. Test the feasibility of **offering the LCSH vocabulary in a format usable on the Web** to encourage development of applications that we hope would help both catalogers and end-users.

ACTION: For several years CDS and CPSO have discussed the legal and budgetary issues related to making LC's authority records freely available on the Web. During 2007, Beacher Wiggins (Director, ABA) and Kathryn Mendenhall (Acting Director, TechPolicy) approved CPSO proceeding with a test of a free version of *LCSH* in SKOS format (while retaining the subscription for that database in MARC 21 format). Since May 2007, following meetings with Semantic Web representatives, CPSO has explored with OSI a test to convert the LCSH vocabulary into SKOS/RDF and we expect to conduct that test during FY08.

The *Library of Congress Subject Headings* and the process for creating and maintaining them should be improved upon to reduce costs, so this system can continue to serve us for decades to come.

Thoughts on LCSH

Lois M. Chan

July 31, 2006

Introduction

It seems that several separate but related issues must be considered in any discussion of the future of LCSH:

- Keyword vs. controlled vocabulary
- Library of Congress Subject Headings (LCSH)
 - Terminology
 - Application
 - Precoordination
 - Postcoordination

Each of these issues may be examined from different perspectives:

- Theoretical/philosophical soundness
- Usefulness in OPACs and in the Web environment
- Cost in implementation

Keyword vs. controlled vocabulary

This is a broad issue and much has been written or said about it. Keyword is the predominant approach to information retrieval on the Web, and controlled vocabulary, with a much longer history, has been the primary method for subject retrieval in library catalogs. Its use has also carried over to many of the databases.

Whether, in the current environment, there is still a need for controlled vocabulary is the issue that must be dealt with first, because if the library community decided that there is no longer a need or place for controlled vocabulary, then the future of LCSH would be a moot point.

The following comments are based on the assumption that controlled vocabulary will continue to serve useful purposes in subject retrieval, particularly in controlled environments such as OPACs, databases, and knowledge management systems.

Library of Congress Subject Headings (LCSH)

In considering the future of LCSH, it is important to separate two aspects of the system:

- (1) Terminology – the Library of Congress Subject Heading List
- (2) Application – How LCSH is used. Issues involved in application include precoordination vs. postcoordination and the depth (summarization vs. exhaustivity) in assigning headings

Terminology

The word “terminology” is used here to refer to how terms and their relationships are created and displayed. As a subject vocabulary, LCSH has the following features:

- LCSH is the most comprehensive subject vocabulary in the English language
- It has a rich vocabulary covering all subject areas
- It has synonym and homograph control
- It has extensive hierarchical and associative references among terms
- It is the de facto standard controlled vocabulary in library catalogs—it is used extensively by academic and research libraries and many public libraries in the United States and many other countries around the world
- It has been translated and adapted into many languages
- It has served as the model for developing controlled vocabularies in many countries world wide
- It is compatible with other subject vocabularies used in OPACS, such as Sears List of Subject Headings and Medical Subject Heading, and more recently FAST, all of which had their origins in, or were modeled on, the LCSH vocabulary.

Because of its use in cataloging records produced by the Library of Congress and its wide adoption by other libraries, LCSH represents a unique contribution to the library community by the Library of Congress. The impact of LCSH is one of the most far-reaching of all LC’s bibliographic and access services.

This is not to say that LCSH as it stands is an ideal or perfect controlled vocabulary, but it is unique in its comprehensiveness and its rich terminology. To abandon the LCSH list would be a disservice to its users around the world, and would amount to dismantling the foundation on which much of the subject access services around the world is built.

Application

The term “application” is used here to refer to how LCSH has been and is being used to provide subject access to information, especially in library catalogs. Since its inception in late nineteenth century, libraries have followed LC’s practice in applying LCSH. LC’s practice in assigning subject headings places emphasis on subject heading strings that summarize or capture the main content of the resource being represented. The full-string approach to complex subjects is designed:

- to ensure precision in retrieval;
- to facilitate browsing of multiple-concept or multi-faceted subjects in the online catalog; and,
- to provide an “overview of the shape and depth of the book literature,” an overview that is “both *comprehensive* and *comprehensible*.” (Mann, 2006)

At the heart of the issue is the question of precoordination vs. postcoordination. Both approaches have their advantages and disadvantages.

Precoordination

No one has provided more cogent and compelling arguments on the advantages of precoordination, particularly in the research and scholarly environment, than Thomas Mann of the Library of Congress. Some of the major advantages of precoordination are:

- Preserves context
- Preserves relationships between concepts
- Is able to express complex subjects
- Is unambiguous

In the OPAC, designed to accommodate heading strings, the precoordinate approach, in the hands of professional and experienced searchers, has contributed to effective and fruitful retrieval.

On the other hand, the disadvantages of precoordination include:

- Requires manual indexing
- Requires elaborate rules for heading construction (only a very small fraction of headings assigned to bibliographic records are enumerated in the LCSH list)
- Is costly to maintain and verify
(Note: this is most prevalent argument against the precoordinate approach)
- Is less flexible
- Is of little value in environments where browsing the index or the thesaurus is not an option, because users typically do not search by string

These limitations mean that applying LCSH properly in compliance with current policy and procedures entails the following requirements:

- trained catalogers and indexers
- systems with index browsing capability
- systems with online thesaurus display
- sophisticated searchers to fully benefit from the subject strings

Even with elaborate rules, no existing subject access system is totally precoordinate because of the infinite possibilities of combining concepts and facets. (The PRECIS System of the 1970s came closest to a fully precoordinate system). LC's current policy in applying LCSH *prefers* precoordination, but in many instances complex subjects are represented by multiple headings. As a result, elaborate rules have been put in place to ensure consistency in string construction.

Postcoordination

The advantages of postcoordination include the following:

- Is flexible in searching
- facilitates automatic indexing
- Is easy to maintain
- Allows infinite combinations
- Is amenable to search mechanisms outside of OPACs

There are also disadvantages in taking the postcoordination approach:

- False coordination results in low precision
(Note: this is the major argument against the postcoordinate approach)
- Lack of context in headings often hampers the searcher's ability to evaluate the relevancy and usefulness of the retrieved items
- Complex relationships among topics and concepts cannot be fully and precisely represented by stand-alone headings

The Future of LCSH

The pros and cons of precoordination and postcoordination present challenges to those pondering the future of LCSH. As library cataloging data are being integrated into the networked environment, certain features would be desirable in order to improve interoperability between LCSH and other information retrieval systems and to facilitate federated searching:

- improving compatibility in syntax between LCSH and other controlled vocabularies such as other subject headings lists, thesauri, and structured lexical lists
- ensuring the viability of LCSH in search environments outside of OPACs
- minimizing the need to construct heading strings manually by simplifying the rules for heading construction
- developing and improving tools for automatic indexing and cataloging using LCSH
- developing and improving tools for automatic or computer-assisted authority control
- developing new search mechanisms that maximize the features of LCSH, especially the rich relational links among headings

Already, there are encouraging experiments and developments along these lines. For example, iVia, an open source Internet subject portal and virtual library system, is using LCSH to test its content creation feature. Gary Strawn's Toolkit includes batch processing capability for making subject heading changes and corrections once identified. The newly developed Endeca system, originally designed for a much wider application environment, is being implemented in the OPAC of North Carolina State University. With its Guided Navigation feature, Endeca helps OPAC users focus or narrow their searches by exploiting the relational references in the LCSH list and different facets in LC heading strings assigned to cataloging records.

In deliberating the future of LCSH, a few key issues and questions to be raised may include:

- (1) Is it an all-or-nothing or an either-or proposition?
Is the question one of either retaining or abandoning LCSH all together?

- (2) Must all resources be treated equally in subject representation? In other words, is it mandatory that the same level of application be applied to all resources?

Most people would agree that not all library resources are of equal value or merit. For example, many resources are of an ephemeral nature and are not worth the time, effort, and cost of detailed subject representation. Libraries have dealt with this problem by setting priorities. Examples include different levels of bibliographic description adopted by LC and PCC. A scalable or multi-layered approach has also been adopted in other aspects of bibliographic control, e.g., various versions of MARC and the Dewey Classification.

Many would agree that books and other resources of a research and scholarly nature warrant fuller treatment in order to maximize their retrievability. Most OPACs have browsing capabilities, which optimize the usefulness of subject strings. For such resources, precoordinated full subject heading strings may be justified and worth the investment in time effort, and cost.

Incidentally, PCC's core record policy has set a precedent for a tiered approach to subject analysis. A variation to the policy of assigning fewer heading strings to core- or minimal-level records could be the use of multiple headings as enumerated in the list without constructing long strings.

- (3) Is it possible that the high cost in creating and maintaining LCSH headings assigned to bibliographic **records may be reduced by simplifying the application rules?**

- (4) Can technology be called on to facilitate subject analysis to achieve “the best of both” approaches?

An interesting example is National Library of Medicine's LocatorPlus, where individual subject headings belonging to different facets are stored separately in the database and re-assembled into strings for users who preferred precoordinated MeSH heading strings. This would require a fixed citation order of facets within a string. With regard to LCSH, the Airlie House Conference discussed the issue of citation order and came up with a number of prevalent patterns.

To explore some of these questions, it may be useful to revisit some of the research findings regarding LCSH usage, for example:

- In 1984, Cochrane (Cochrane, 1984) recommended breaking up long subdivided subject headings applying a simple set of rules that could also be employed to reconstruct the same long subdivided subject headings that we now use

- In 1986, Pauline Cochrane (Cochrane, 1986) made the reservation about lengthy subdivided subject headings saying "the logic behind the string's construction is lost on most catalog users and some catalogers."

- In 1994, Karen Drabenstott and Diane Vizine-Goetz suggested (Drabenstott and Vizine-Goetz, 1994):

“Standardizing subdivision order would simplify cataloging and save money. Library schools and technical services departments would no longer spend time training people how to order the subdivisions in subject heading strings. Cataloging staff would no longer spend time determining the order of subject subdivisions. They would build strings based on a standardized order of subdivided elements. Library systems staff could introduce computer-based techniques to automatically verify the order of subdivisions in existing strings and in newly assigned strings. Such techniques would reduce the errors that occur in subdivided subject headings due to subdivision order.”

- In a further study in the late 1990s, Drabenstott, Simcox, and Williams conducted a study on "Understanding Subject Headings in Library Catalogs" (Drabenstott et al., 1999), testing the extent to which librarians are able to interpret subdivided LC headings correctly. Among their findings is that about 54 percent of the meanings librarians gave to subdivided subject headings were correct. The researchers made the following recommendation:

“Librarians need to examine both approaches--standardizing subdivision order and breaking up long strings, determine the benefits and drawbacks of each approach to catalogers, cataloging operations, and catalog users, and make decisions about how best to proceed. Maintaining the present practice of lengthy subdivided subject headings is wasteful because, ultimately, the meaning of these strings is lost on most cataloger users and some catalogers. Something needs to be done--now!”

- Another study (Chan and Vizine-Goetz, 1998), also conducted in the late 1990s focusing on LC subject heading distribution and error rate, found that out of a sample of 20,473 LC headings assigned to LC-created records, 14,500 (70.8%) headings were assigned only once each and accounted for 24% of total usage in the sample. On the other hand, 122 (0.6%) headings were assigned more than 50 times each and accounted for 25.8% total usage in the sample. The headings assigned less frequently also showed a higher error rate. In the same study, an analysis on the overlap between enumerated headings and assigned headings was conducted on a subset of the sample. Among 4,233 assigned headings, only 590 (13.9%) were enumerated in LCSH; the rest were constructed by catalogers.

LCSH is an enormously complex system. Deliberations on its future will have far-reaching implications for direct and indirect users both within and outside the library community.

References

- Chan, Lois Mai and Diane Vizine-Goetz. (1998). Toward a Computer-Generated Subject Validation File: Feasibility and Usefulness. *Library Resources & Technical Services* 42(1):45-60.
- Cochrane, Pauline A. (1984). Modern Subject Access in the Computer Age. *American Libraries* 15 (Apr. 1984): 250-52, 254-55.
- Cochrane, Pauline A. (1986). *Improving LCSH for Use in Online Catalogs*. Littleton, Colo.: Libraries Unlimited. p, 62.
- Drabenstott, Karen M. and Diane Vizine-Goetz. (1994). *Using Subject Headings for Online Retrieval: Theory, Practice, and Potential*. San Diego, Calif.: Academic Pr. pp. 113-20.
- Drabenstott, Karen M., Schelle Simcox, and Marie Williams. (Summer 1999). Do Librarians Understand the Subject Headings in Library Catalogs? *Reference & User Services Quarterly*. 38(4):369-87.
- Mann, Thomas. (2006). The Changing Nature of the Catalog and Its Integration with Other Discovery Tools. Final Report. March 17, 2006. Prepared for the Library of Congress by Karen Calhoun: A Critical Review. (<http://guild2910.org/AFSCMECalhounReviewREV.pdf>)

To: Barbara B. Tillett
From: Daniel N. Joudrey and Arlene G. Taylor
Date: July 17, 2006
RE: **LCSH Strings – some thoughts**

In recent months we have been hearing rumors throughout the cataloging community about the potential demise of LCSH. We would like to share some concerns, if that is indeed a course of action being considered by the Library of Congress. While we do not think that the system is perfect, we do believe that full LCSH strings add great value to our catalog records. LCSH strings assist us in fulfilling Cutter's "objects" of the catalog. From the time of Cutter until today, catalogers have been charged with providing access to materials of which the subject is known, to show what the library has on a given subject, and to assist in the choice of a book as to its character. LCSH strings not only serve the finding/search function, but they are key to the collocation/gathering function and to the choice/evaluation function as well. They make the catalog navigable in ways that keywords and single descriptors do not. Many of the arguments made about replacing LCSH tend to focus only on the first objective without a fair consideration of the importance of LCSH to the second and third of Cutter's "objects." We hope the following points and examples help to illuminate the importance of LCSH strings in the organization of information.

While keyword searching is fine for information-seekers, the scholar needs more. If one is looking for something or anything about a topic, a scattershot approach is fine, but the researcher with a specific and narrow focus needs precise tools to find needed and appropriate materials for her/his scholarly endeavors. LCSH strings are the anchor for precision-oriented searching in the online public access catalog. While keyword searching may direct the user to an initial item, it is the subject headings that bring the researcher directly into contact with a larger pool of related materials collocated by subject headings (but not necessarily by classification number). The combination of terms in the string provides a fuller meaning, with deeper levels of specificity that allow for a more useful evaluation of the items returned as search results. The precision available via the strings allows the researcher to choose the relevant materials quickly and effectively.

In addition, the ability to browse the full subject heading strings after a subject search allows the researcher to gain a better understanding of the topic area. The browse screen with full LCSH strings in alphabetical order appears to break the broad topic into much more manageable, comprehensive component parts by dividing the subject into further topical, geographic, chronological, and/or form groupings. The string-based browse display gives the researcher a conceptual map of the topic; it is an invaluable overview of the shape and depth of a subject area that allows the researcher to better identify the major subtopics and the major gaps within the knowledge domain. The LCSH strings are also useful in creating new displays for the catalog. The Endeca-powered OPAC at NCSU (<http://www.lib.ncsu.edu/searchcollection/>) would not be nearly as useful without the LCSH subdivisions that are added to the subject headings. This system uses the

LCSH strings, with all its subdivisions, to categorize the search results into various facets based on, among other things, geographic area, time period, form, and further subject subdivisions. Without the full LCSH strings, the meaning of the subjects in the individual items are lost. The system was designed to take better advantage of the subject information that has already been created in ways that most OPACs do not. Endeca, with its use of LCSH strings, provides the best of both worlds, providing fully understandable LCSH strings within the individual records to offer meaning and context alongside the ability to manipulate, mix, and match individual components from those strings creatively to make our search results all the more navigable, useful, and profitable. We believe the usefulness of the topical metadata is only now beginning to be fully tapped by the systems being developed. Without LCSH strings, the development of this interesting and valuable area of system design would be thwarted before it could even be fully realized.

Subject heading strings provide context and meaning, without which the researcher has very little on which to base relevance judgments. The set combination of concepts within a single string allows the researcher to understand how the various aspects of the subjects interact within the information package. Without that context, only a series of isolated keywords or subject words would exist. That set of keywords, while somewhat helpful in understanding the subject of the entire information package, can lead to false hits and confusion because not all subject words are appropriately matched with the others. If one has a work on multiple topics that take place in various locations over several time periods, there are numerous possible combinations of the component concepts. Many of those combinations, however, could be incorrect. For example, the book *Die silberne Taufschale zu Siegen* is about:

650 _0 Silverwork ‡z Peru.
 650 _0 Art objects ‡z Germany ‡z Siegen.

If the full subject heading strings weren't in place, one might think that this work was about silverwork in Germany or art objects in Peru. This is a relatively simple example, but what about the item *Evitas Geheimnis*? It has the following subject headings:

600 10 Peron, Eva, ‡d 1919-1952 ‡x Travel ‡z Switzerland.
 600 10 Peron, Juan Domingo, ‡d 1895-1974.
 651 _0 Argentina ‡x Ethnic relations.
 650 _0 Voyages and travels.
 651 _0 Argentina ‡x Foreign relations ‡z Switzerland.
 651 _0 Switzerland ‡x Foreign relations ‡z Argentina.
 650 _0 Immigrants ‡z Argentina ‡x History ‡y 20th century.
 650 _0 Bank accounts ‡z Switzerland.
 650 _0 Antisemitism ‡z Europe ‡x History.
 651 _0 Argentina ‡x Emigration and immigration.
 651 _0 Germany ‡x Emigration and immigration.
 650 _0 War criminals ‡z Germany ‡x History ‡y 20th century.
 650 _0 War criminals ‡z Argentina ‡x History ‡y 20th century.

650_0 Nazis ‡z Argentina ‡x History.

After removing the duplicate headings, there are 20 separate concepts represented.

1895-1974

1919-1952

20th century

Antisemitism

Argentina

Bank accounts

Emigration and immigration

Ethnic relations

Europe

Foreign relations

Germany

History

Immigrants

Nazis

Peron, Eva

Peron, Juan Domingo

Switzerland

Travel

Voyages and travels

War criminals

This means there are 2,432,902,008,176,640,000 possible combinations of these terms. There are quite a few chances to get some false drops in this set of terms.

As an aside, neither of the previous examples would be retrievable to an American researcher using English keywords without some work by an English-speaking cataloger to determine aboutness and to add to the record some kind of search terms in English. Even a researcher who reads German is not likely to think of searching an American catalog using German terms. And other researchers may find a foreign-language work to be so germane that they wish to have it translated.

Works in the English language, of course, also benefit from terminology being coordinated into strings. *Atlas of selective sentinel lymphadenectomy for melanoma, breast cancer, and colon cancer*, for example, has the following subject headings:

650_0 Lymphatic metastasis ‡x Diagnosis ‡v Atlases.

650_0 Lymph nodes ‡x Dissection ‡v Atlases.

650_0 Melanoma ‡v Atlases.

650_0 Breast ‡x Cancer ‡v Atlases.

650_0 Colon (Anatomy) ‡x Cancer ‡v Atlases.

650_0 Tumors ‡x Classification ‡v Atlases.

Eleven separate concepts are represented, not all of which can be usefully combined:

Atlases
Breast
Cancer
Classification
Colon (Anatomy)
Diagnosis
Dissection
Lymph nodes
Lymphatic metastasis
Melanoma
Tumors

Note that a keyword search for “Colon Classification” would retrieve this record.

Finally, the dismantling of LCSH would destroy the work that has been accomplished in the area of form/genre, an area of particular value in public libraries as well as in special collections. Assigning a single heading of Atlases, Dictionaries, Catalogs, or Statistics renders those headings meaningless.

Excerpt from

“Precoordination or Not”

in Elaine Svenonius. *The Intellectual Foundation of Information Organization*.

Cambridge, Mass.: MIT Press, 2000, p. 188-192.

{Request for copyright permission in progress}

“The pros and cons of coextensivity are the pros and cons of a more basic issue – whether subject languages need to employ a syntax at all. This issue was first raised in 1951 when the potential of the computer in information retrieval was first discussed. Mortimer Taube, an early visionary, proposed replacing the relatively complex grammatical syntax used to construct subject headings with a simple logical syntax.²⁹ Taube called this simple logical syntax *coordination*; it is known now as *Boolean coordination*. To explore the implications of his proposal, Taube developed a uniterm subject language – a syntax-free language whose vocabulary consisted almost entirely of one-word terms. He then evaluated his language with respect to other subject languages, empirically and analytically.³⁰ The empirical evaluation took the form of a retrieval experiment – an experiment that in embryonic form prefigured the landmark Cranfield experiments of several years later. The analytical evaluation consisted of comparing subject heading and uniterm languages according to sixteen evaluation criteria, among them simplicity, speed in cataloging or indexing, vocabulary size, vocabulary growth and obsolescence rates, cost, universality, compatibility, specificity, and suggestibility.

Now, nearly fifty years later, the issue of “precoordination or not” or whether to employ a subject-language syntax is still unresolved. As with many bibliographic dilemmas, this one comes down to costs versus benefits. The pros and cons can be summarized using some of Taube’s original criteria as well as other criteria that derive from the collocation and navigation objectives:³¹

- *Specificity* Cutter’s injunction to use the most specific heading possible is the foremost rule in indexing. The concept of specificity is an elusive one.³² In part this is because there are different ways to achieve specificity. The specificity achievable by a postcoordinate language using the Boolean syntax of intersection and coordination is different in kind from the specificity that can be achieved by a precoordinate language using a subject-language syntax. Taube argued that the former is superior to the latter, at least for scientific subjects. But he did not demonstrate this, and, in fact, it {is} unlikely it could be demonstrated since a common criticism leveled against postcoordinate languages is their inability to express relationships more specific than those expressible by the AND, NOT, and OR operators.³³ Precoordinate languages have a more sophisticated syntax and as such are more expressive. (PRECIS, for instance, boasts of its ability to distinguish the biting of dogs by children and the biting of children by dogs.)
- *Precision* This post-Taube measure is the degree to which a subject language is capable of deselecting irrelevant documents in retrieval. To the extent that precision is a function of the specificity and expressive power of a subject language, precoordinate languages also score better here. Generic Boolean relationships foster false drops: a search for information on the history of philosophy using the search prescription History AND Philosophy may also retrieve material on the philosophy of history.

- *Contextuality* Related to specificity and expressive power, contextuality is the ability of a subject language to recognize distinctions in meaning. Since the task of vocabulary disambiguation is not limited to homonyms and polysemes but extends to all words whose meaning is in part contributed to by context, context is needed for making distinctions in meaning. The context shown by Boolean intersections of terms are relationally indiscriminate and therefore cannot disambiguate sufficiently to achieve acceptable precision. Precoordinate strings, by contrast, offer a wide variety of contexts that can be used to pinpoint meaning.
- *Suggestibility* Often users do not know what to do when they retrieve too many citations. They find it difficult to think of appropriate qualifiers. A precoordinate language addresses this problem by displaying to users the different strings in which their search terms are embedded. The association of ideas triggered by such displays helps users to imagine what they cannot clearly articulate a priori. It was only in the matter of suggestibility that Taube conceded that subject-heading languages with their long phrases and topical subdivisions might be superior to his uniterm language.

Natural-Language versus Subject-Language Syntax

Context is needed for disambiguation, suggestibility, and precision. Immediately, the question of exploiting the richness of natural-language contexts in retrieval arises: why incur the expense of creating artificial subject-language contexts? Natural-language contexts have long been used in indexing, at least since the middle of the last century when Samson Low and Andrea Crestadoro introduced title-term indexing.³⁴ a precursor of keyword indexing, introduced a century later by Hans Peter Luhn.³⁵ There are various forms of keyword indexing. Relevant to the present discussion is keyword in context indexing (KWIC). In a KWIC index document titles are algorithmically rotated to bring each of its keywords to a selected eye-catching position. Simple in its conception and execution, KWIC spawned a progeny of more complex string index languages whose syntaxes are algorithmic permutations of natural-language syntax.³⁶ Using an index string generator, they transform a natural-language input into an output consisting of a set of index strings. Examples of such languages are Michael Lynch's Articulated Subject Indexing and Timothy Craven's NEPHIS.³⁷

KWIC-derived languages were designed for printed indexes and are infrequently used today. However, most retrieval systems exploit natural-language contexts in their displays. Typically, a first-stage response to a user's query is to display his search terms in a syntagmatic context, such as in title statements, abstracts, or other text fragments. Contexts such as these are helpful in suggesting additional search terms and in narrowing searches to improve precision. An important question is whether they are any more or less helpful than the more pricey context afforded by subject-language syntax.

It can be argued that the relative uniformity of a subject-language syntax and the type of specificity it provides allow it to create better browsing domains than a natural-language syntax. In card-catalog days subject headings were subdivided to ensure that not too many cards would file behind a single heading. This was a simple and expedient means to improve browsability and also precision. The bunching of large numbers of cards behind a single heading in a card catalog has its online analog in large retrieval sets. These must somehow be managed. One way to do this is to order retrieved text fragments according to the frequencies of occurrence of search terms in them. But the

resultant ordering, being statistical rather than semantic or conceptual, does not, like subdividing a card file, create a browsable domain.

The syntax used in subject languages offers a means to organize large retrievals conceptually. It does this by chunking, compressing, and layering subject descriptions of the retrieved items. A keyword-in-*title* search on a widely used term like Art would in many online catalogs result in an unmanageably large retrieval. A more focused strategy would be to perform a keyword-in-*subject heading* search on the term. (Even this more focused search would in the UCLA ORION 1 catalog retrieve around 8,000 subject headings.) The focused strategy offers a better guarantee of relevance, since the indexing is intellectual rather than mechanical. Moreover, and more to the point here, is that what is retrieved is conceptually browsable. The syntax used to create headings can also be used to order them in a systematic manner. While the ordering of subject headings in most catalogs today is distorted by the use of computer filing, it is in principle possible to order them automatically and neatly by geographical, form, chronological, and topical facets. The conceptual ordering could be further refined by grouping topical headings into larger classes – for instance, by organizing them by discipline.³⁸ Browsability could be enhanced by consolidating headings containing like information. For instance, headings such as Art, Modern – Nineteenth century – Austria; Art, Modern – Nineteenth century – Belgium; and Art, Modern – Nineteenth century – Brazil could be consolidated as Art, Modern – subdivided by geographical area, which then can be opened when needed by a mouse click.³⁹ A subject-language syntax, then, is superior to a natural-language syntax insofar as it can organize large sets of citations by layering and chunking. The effect of assigning precoordinate subject headings to documents is to pack them, or rather their surrogates, into an orderly sequence of Chinese boxes.

The creation of a meaningful order is equally as important in information organization as the grouping of documents into classes. In the pre-computer age filing rules could be relied on to order bibliographic records in a systematic manner. Computer filing can produce orders based on statistical, alphabetical, and numerical data but cannot without assistance produce orders that are semantically meaningful. Imposing structure on the material to be ordered is a way of providing that assistance, which is what bibliographic languages, author-title language, as well as subject languages can do by virtue of their syntax.

A second answer to the subject-language versus natural-language syntax question is the answer given to the more familiar question of free-text versus controlled-vocabulary searching. The answer turns on the need for normalization: a searching vocabulary needs to be normalized and so does the syntax used to create synthetic search expressions. Without normalization there will be recall, precision, and navigation failures in retrieval. This does not have to be demonstrated empirically, since it is logically deducible from the fact that natural languages contain synonymy and homonymy at the level of both vocabulary and sentence structure. Ridding a natural language of its surface-level anomalies amounts to reducing the variety exhibited by phenomenal language to its seminal or deep structure. The reductive analysis pursued in linguistics for the purpose of explaining and generalizing about natural languages, in bibliographic practice, serves the more practical purpose of intelligent information organization.”

29. Taube (1953).
30. Taube (1953).
31. Svenonius (1995).
32. Svenonius (1971).
33. Farradane (1970).
34. Crestadoro (1856).
35. Luhn (1959).
36. Craven (1986).
37. Lynch (1969); Craven (1982).
38. Massicotte (1988).
39. Massicotte (1988); Svenonius and McGarry (1991).

- Craven, Timothy C. (1982). "Automatic NEPHIS Coding of Descriptive Titles for Permuted Index Generation." *Journal of the American Society for Information Science* 33 (2): 97-101.
- Craven, Timothy C. (1986). *String Indexing*. Library and Information Science Series. Orlando: Academic Press.
- Crestadoro, Andrea. (1856). *The Art of Making Catalogues of Libraries, or, A Method to Obtain in a Short Time a Most Perfect, Complete, and Satisfactory Printed Catalogues of the British Museum Library*. London: Literary, Scientific and Artistic Reference Office.
- Farradane, Jason E. L. (1970). "Analysis and Organization of Knowledge for Retrieval." *Aslib Proceedings* 22 (12): 607-616.
- Luhn, Hans Peter. (1959). Keyword-in-Context Index for Technical Literature (KWIC INDEX). Technical Report RC-127. Yorktown Heights, N.Y.: IBM.
- Lynch, Michael F. (1969). "Computer-Aided Production of Printed Alphabetical Subject Indexes." *Journal of Documentation* 25 (3): 244-252.
- Massicotte, Mia. (1988). "Improved Browsable Displays for Online Subject Access." *Information Technology and Libraries* 7(4): 373-380.
- Svenonius, Elaine. (1971). "The Effect of Indexing Specificity on Retrieval Performance. Ph.D. dissertation, University of Chicago.
- Svenonius, Elaine. (1995). "Precoordination or Not." In *Subject Indexing: Principles and Practices in the 90's: Proceedings of the IFLA Satellite Meeting Held in Lisbon, Portugal, 17-18 August 1993*, edited by Robert P. Holley, Dorothy McGarry, Donna Duncan, and Elaine Svenonius, 231-235. Munich: Saur.
- Svenonius, Elaine and Dorothy McGarry. (1991). "More on Improved Subject Displays." *Information Technology and Libraries* 10: 185-191.
- Taube, Mortimer. (1953). *Studies in Coordinate Indexing*. Washington, D.C.: Documentation Incorporated.

Airlie House Conference Recommendations and Progress in Carrying Them Out

Recommendation #1: *Under topical headings (as opposed to name or place headings), place, chronological and form subdivisions shall be applied as needed and on an individual basis, based on the judgment of the cataloger as to their appropriateness to the item being cataloged. If the cataloger chooses to apply subdivisions, the subdivisions should always appear in the following order: topical, geographic, chronological, form. This is not to suggest that each type of subdivision shall always be present under each heading; it is simply to specify a standard order for them when they are assigned. When a non-topical subdivision element is expressed or implied in a topical main heading, then as a general rule the main heading should be used rather than a subdivision under a topical heading. A topical heading should not receive geographic subdivision if it is used as a topical subdivision after geographic subject headings.*

The Library of Congress identified two critical elements in this recommendation: 1) facilitation of the assignment of geographic subdivisions, and 2) standardization of the order of subdivisions in topical subject heading strings. To facilitate the assignment of geographic subdivisions, new topical headings for which geographic treatment is possible are established with authorization for geographic subdivision. Authorization for geographic subdivision is added to records for existing headings on a case-by-case basis as they are modified for any reason. Projects have been undertaken to add geographic subdivision authorization to specific types of headings. To achieve the recommended standard order of **[topic]--[place]--[chronology]--[form]** where it can be applied, new topical subdivisions for which geographic orientation is feasible are now established with authorization for further geographic subdivision. On a case-by-case basis, topical subdivisions not previously divided by place are being authorized for geographic subdivision. A significant percentage of the free-floating topical subdivisions have been authorized for geographic subdivision, but many highly used subdivisions remain unsubdivided because of the volume of bibliographic file maintenance that changes to them would entail. To bring the expression of more concepts into the recommended standard order of **[topic]--[place]** and to eliminate exceptional practices, over seventy subdivisions on the list of free-floating subdivisions used under names of places were discontinued in favor of dividing the corresponding main headings by place.

The Library of Congress team that investigated this recommendation recognized that a significant change to current practice would occur if the proposed order were applied in special areas, such as art, literature, and history. They also recognized that potential loss of meaning could result in some categories of headings. Therefore, they recommended that those areas be studied further. In 1993, the ALCTS CCS Subject Analysis Committee (SAC) established a Subcommittee on the Order of Subdivisions in Library of Congress Subject Heading Strings to pursue alternatives to the recommended order in special areas, such as art, literature, and history. The Subcommittee submitted its final report to SAC in 1996. Its basic finding was that many of the goals of the Subject Subdivisions Conference would not be met by a wholesale implementation of the

proposed standard order, but its feasibility studies supported the hypothesis that some level of default order was possible if several broad areas of exceptions were identified and specified. The report included specific recommendations for changes regarding application of the subdivision --**History** and the application of geographic and chronological subdivisions with art headings. LC subsequently clarified use of the subdivision --**History** and made the changes recommended by SAC in the order of subdivisions in art headings in 2001.

Recommendation #2: *The developing “national authority file” should contain authority records for topical headings and for topical heading-topical subdivision(s) combinations. Further non-topical elements in any given string will not normally be established, unless such a record is desirable for cross-reference purposes. Authority records for headings containing subdivisions governed by pattern and free-floating lists will not require formal editorial review. Topical subdivision records and coding showing relationships between headings and topical subdivisions would be desirable features to help create the file and to assist validation. The conference encourages the Library of Congress to continue and expand its programs by which other libraries contribute to the national subject authority file.*

OCLC’s Office of Research has supported investigation of the feasibility of a machine-generated subject validation file that would contain full strings of all types of headings, including name, topical, and geographic headings. In 1994, a Task Group on the Subject Authority File was established by the predecessor of the Program for Cooperative Cataloging (PCC). It was charged to define and evaluate the functional requirements and uses of the subject authority file and its relation to LCSH, including evaluating the feasibility and need for representing entire character strings in the authority file and the necessity of creating authority records for subject subdivisions to allow for better online control of heading-subdivision combinations. Its final report was issued in November 1994.

The Library of Congress implemented the 18X fields for subdivision authority records in 1999 and undertook a project to create subdivision authority records for the more than 3,900 free-floating subdivisions from 1999-2003. Those records are used to generate an introductory section on free-floating subdivisions in the annual printed edition of *Library of Congress Subject Headings*, to validate main heading/free-floating subdivision combinations in LC’s online catalog, and to control subject headings in OCLC’s Connexion system.

The Library of Congress has expanded the means by which cooperating libraries may contribute subject heading proposals in the Subject Authority Cooperative Program (SACO).

Recommendation #3: *Chronological subdivisions under topical headings should relate to the coverage of the content of the item and not to its date of issue. The Library of*

Congress should investigate using numerals as dates or date ranges in chronological subdivisions. The Library of Congress should consider the pros and cons of the use of free-form chronological subdivisions. Vendors and utilities should be encouraged to develop range searching capabilities.

The Library of Congress agreed that chronological subdivisions used under topical headings should relate to coverage of the contents of the items rather than to their date of issue and cancelled nearly 200 subdivisions that had been used to represent date of issue rather than contents. The Library has no plans to use numerals for dates or date ranges in chronological subdivisions in all cases, nor to change chronological subdivisions following the subdivision --**History** where alphabetic characters appear as the initial elements. However, words have been deleted from over 100 chronological subdivisions established under names of places where the words were deemed not necessary. The Library decided not to pursue the use of free-form date subdivisions.

Recommendation #4. *The question of whether subdivisions should be coded specifically to improve online displays for end users should be considered by organizations such as the Network Development/MARC Standards Office at the Library of Congress, MARBI and SAC of the American Library Association, and the utilities, among others. In particular, the Library of Congress should investigate implementing a separate subfield code for form subdivisions.*

In 1995, a proposal was approved to define a new subfield code \$v for form subdivisions in the MARC authority and bibliographic formats. In 1999, the Library of Congress implemented use of \$v to identify form subdivisions functioning as forms in subject heading strings in authority and bibliographic records. Subdivision authority records include codes that identify the free-floating list(s) in the *Subject Cataloging Manual: Subject Headings* that authorize their use with particular categories of headings.

Recommendation #5: *The current policy of indirect geographic subdivision should be continued. The Library of Congress should investigate including the indirect form of geographic headings in authority records for geographic names.*

The Library of Congress agreed to continue its practice of indirect geographic subdivision. In 1992, a 781 linking field to use for recording the geographic subdivision form of geographic headings was added to the MARC authority format. The Library of Congress implemented the 781 field in 1999, when this data began to be added to new geographic subject authority records. A project to add 781 fields to existing geographic subject authority records was completed in 2003. Also in 2003, NACO libraries were given the option of adding 781 fields to the geographic name authority records they create or revise. In 2006, OCLC completed a project to enhance over 83,000 name authority records for jurisdictions by programmatically adding either 781 fields showing their geographic subject subdivision forms or notes about subject usage in 667 fields.

Recommendation #6: *The Conference strongly recommends that the Library of Congress simplify subdivisions in the Library of Congress subject headings system. Target areas for simplification include the reduction of overly fine distinctions, the consolidation of lists, and increased consistency in syntax. The Conference acknowledges the potential impact of such changes on existing files, but the changes are of such importance that they should be made in spite of possible disruptions to existing databases. The Conference urges the Library of Congress to pursue vigorously enhancements to its automated systems that could compensate for disruptions caused by these changes.*

Simplifying the form and application of subdivisions continues on many fronts, and specific changes to headings and subdivisions are regularly announced. Changes that are deemed desirable are made in spite of their impact on existing databases. Hundreds of subdivisions have been updated to more current forms, replaced by existing subdivisions or phrase headings, or discontinued. The formation of the Database Improvement Unit in the Cataloging Policy and Support Office and the use of customized software has aided in updating bibliographic records for one-to-one changes. Documentation on subdivisions and lists of free-floating subdivisions in the *Subject Cataloging Manual: Subject Headings* have been streamlined and improved and a fifth edition of the manual was published in fall 1996.

FAST (Faceted Application of Subject Terminology)

FAST (Faceted Application of Subject Terminology) is a research project that has been in development at OCLC since 1998. Currently, authority records for FAST headings are available on an OCLC web site in Site Search. A user interface providing both search and browse capabilities remains to be developed.

FAST is derived from LCSH and the LCNAF and depends upon those systems. FAST generally retains terminology from LCSH but applies it using a different syntax. The basic component categories of topic, place, time period, and form, which in LCSH are generally expressed by separate elements that can be combined to form pre-coordinated subject heading strings, are faceted into separate FAST headings based on the Dublin Core metadata set. Other FAST facets include personal names, corporate names, titles as subjects, and meeting names/events.

Topics:

Topical main headings and main heading/topical subdivision combinations are retained as is from LCSH. These include single term headings as well as headings that are pre-coordinated in adjective noun phrases, prepositional phrases, and heading/subdivision combinations to represent complex or multi-element topics.

LCSH guidelines call for establishing heading/subdivision combinations for unique topics or combining authorized free-floating subdivisions with designated headings to create needed subject headings. In FAST, authority records have been created for topical heading/topical free-floating subdivision combinations that have been assigned in LC and member-supplied bibliographic records in WorldCat. A program was developed to validate authorized combinations, but it is possible that records for some invalid combinations may have been created if they were assigned more than a few times.

The reference structure of LCSH is generally retained on topical headings carried over from LCSH, but no references have been made on the authority records for the main heading/free-floating subdivision combinations. That means that a user would generally have to know how they would be constructed in order to find them.

Geographics:

Geographic headings for jurisdictions from the name authority file and geographic features from LCSH have been converted to FAST geographic headings.

FAST geographic headings are structured hierarchically based on entries in the *MARC Code List for Geographic Areas (GACs)*. That means that geographics that are entered directly in that list are also entered directly in FAST, for example,

Maryland, Mississippi River, United States, Europe. Headings for localities within jurisdictions from the GAC list are entered subordinately, for example, **Maryland--Baltimore**. A geographic feature that is in more than one jurisdiction and does not have its own entry in the GAC list is entered subordinately under the next higher level in the GAC list. For example, a river in two states of the United States is entered under **United States**. The qualifiers that identify those specific states in the LCSH heading are not necessarily retained in the FAST heading. A basic principle that LC has tried to follow is that the same form of heading is used for geographic names for both descriptive and subject access (with the exception of earlier names of jurisdictions that have changed their name). FAST geographic headings used for subject access may differ from the geographic headings that are constructed according to AACR2 and used for descriptive access in LC/PCC cataloging.

Constructing FAST geographic headings has taken much longer than originally expected because many geographic situations are complicated. OCLC staff working on FAST have needed to develop special rules for constructing headings that do not fit in straightforward categories, such as historical entities, bodies of water that are open to the sea (as opposed to inland bodies of water), etc. Some of the resulting headings are not necessarily intuitive. As members of the Subject Analysis Committee (SAC) Subcommittee on FAST have noted, in some cases users need to know the structure of the GAC list and the FAST rules in order to know how some types of geographic headings are constructed.

Time periods:

Most FAST chronological headings consist of numerics rather than words, that is, a single date or date range, for example, **1984** and **1900-1999** for the 20th century. Some chronological headings may include terms, for example, those expressing open dates, **Since 1951**; those with an ending date, **To 1856**; and headings for geologic time periods, **From 140-190 million years ago** (for the Jurassic Period). FAST authority records for chronological headings have only been created when it has been necessary to make cross-references from equivalent or deconstructed LCSH headings. Chronological headings are assigned to reflect actual time coverage in the resources being cataloged. While assigning specific dates may provide greater chronological accuracy, some might argue that the task of determining the specific dates to assign could sometimes take longer or require closer examination of the resources than selecting headings that include pre-established periods, as is done with LCSH, for example, **France--Politics and government--1981-1995**.

Forms:

FAST form headings have been created based on existing LCSH form subdivisions. Because LCSH form headings have not yet been differentiated from topical headings with the MARC 21 form/genre tag of 155, the full range of possible forms/genres is not yet represented in FAST. In some cases, terminology from LCSH form subdivisions was restructured hierarchically to create FAST

form headings, for example, **Catalogs--Video catalogs** for the concept of a catalog listing video recordings.

In deconstructing some LCSH formulations, OCLC staff have made decisions to treat as form some subdivisions that LCSH treats as topical, for example, **History**, but these decisions may be reconsidered.

Other facets:

Personal names: FAST authority records have been created for headings for personal names that have been assigned as subject headings in bibliographic records in WorldCat.

Corporate bodies: FAST authority records have been created for headings for corporate names that have been assigned as subject headings in bibliographic records in World Cat. It should be noted that names of jurisdictions appearing in \$a subfields of headings for subordinately entered government corporate bodies may differ from an authorized FAST geographic heading, for example **Baltimore (Md.). Board of Health**.

Uniform titles: FAST authority records for this facet have been created for titles that have been assigned as subject headings in bibliographic records in WorldCat. However, because a basic principle of FAST is that different facets cannot be combined in one heading, the standard author/title citation form cannot be used. OCLC staff are considering other ways of constructing these headings that would differ from current library conventions, for example, using the title followed by name of the author/creator in parentheses, *Hamlet (Shakespeare, William)*; *Final Act (Conference on Security and Cooperation in Europe)*.

Meetings/Events: OCLC personnel are considering a separate facet that would include 1) headings for events established as topical headings in phrase form in LCSH and carried over to FAST; 2) reconstructed phrase headings for events that are established in subdivision form in LCSH, most frequently following headings of the type **[place]--History**; and 3) headings for conferences, meetings, and organized events such as competitions that are now established as name headings with the tag 111. These FAST headings will all have the same tag and consistent construction.

The FAST project has encountered challenges in breaking up headings in LCSH that combine elements from facets that are kept separate in FAST, for example, headings for historical events and time periods that are established as subdivisions under the names of places and clearly convey their meanings, for example, **United States--History--Civil War, 1861-1865**; **Great Britain--History--Elizabeth, 1558-1603**, as well as headings that combine personal, corporate, geographic, or uniform title headings with topical or form subdivisions, for example, **Napoleon, Emperor of the French, 1769-1820--**

Friends and associates; Jesus Christ--Seven last words; Catholic Church--Influence; Library of Congress--Administration; France---Foreign relations; Bible--Canon.
Headings from separate facets in FAST need to be assigned concurrently to express these concepts, and in some cases, context and/or specific meanings could be lost.

APPENDIX 6

Difference between Google Results and a Browse Display of LCSH Comparison of Search Results of a Post-Coordinate versus Pre-Coordinate Approach

Google search for "porous materials"

Web Results 1 - 10 of about 419,000 for "porous materials". (0.17 seconds)

[Welcome to PMI: Porous Materials Inc | Porosity Home](#)
USA. Manufacturers of precision testing instruments for a range of industrial applications, including nonwovens and textiles. The front page is the entrance ...
www.pmiapp.com/ - 9k - [Cached](#) - [Similar pages](#)

[SpringerLink - Publication](#)
Journal of **Porous Materials**. Publisher: Springer Science+Business Media BV, Formerly Kluwer Academic Publishers BV. ISSN: 1380-2224 (Paper) 1573-4854 ...
www.kluweronline.com/issn/1380-2224/contents - [Similar pages](#)

www.springerlink.com/openurl.asp?genre=journal&...
[Similar pages](#)

www.springerlink.com/link.asp?id=102944
[Similar pages](#)

[IngentaConnect Publication: Journal of Porous Materials](#)
Journal of **Porous Materials** logo Springer logo. Publisher: Springer. 32 issues are available electronically. Volume 12; Number 4, October 2005 · Number 3, ...
www.ingentaconnect.com/content/klu/jopo - [Similar pages](#)

www.swetswise.com/link/access_db?issn=1380-2224
[Similar pages](#)

Sponsored Links

[Need Porous Materials](#)
Find **porous materials** online
Comprehensive list of manufacturers
www.SourceTool.com

[PTFE \(Teflon\) Mesh](#)
PTFE Mesh, for filtration, labs etc
high temp. & chemical resistance
www.internetmesh.net

[Porous Materials?](#)
See how easy it is to measure pore size, volume, surface area, etc.
www.quantachrome.com

[Washable School Glue](#)
Adhesives - Children's Glue
Art Supplies at Misterart.com
www.misterart.com

Voyager keyword search for "porous materials". Results are arranged alphabetically by main entry.

Titles Index

Font: Arial Unicode MS

Name: Main Author, Creator, etc.	Full Title	Date
	Access in nanoporous materials / edited by Thomas J. Pinnavaia and M.F. Thorpe.	1995
	Adsorbts "s"ii "a" : tekstura dispersnykh i poristykh materialov / A.P. Kamaukhov.	1999
	Adsorption and transport at the nanoscale / Nick Quirke.	2006
	Advances in microcrystalline and nanocrystalline semiconductors, 1996 : symposium held December 2-6, 1996, Boston, Massachusetts, U.S.A. / editors, Robert W. Collins ... [et al.].	1997
	Advances in porous materials : symposium held	1995

OK Cancel Search Select All Clear All

221 Records Found Search: Command="porous materials"

Browse display in Voyager

Auth/Ref	70	Porous materials
	1	Porous materials Acoustic properties Mathematical models.
	1	Porous materials Bibliography. [from old catalog]
	1	Porous materials Computer programs.
	29	Porous materials Congresses.
	1	Porous materials Design and construction Congresses.
	1	Porous materials Deterioration.
	1	Porous materials Deterioration Congresses.
	2	Porous materials Experiments.
	1	Porous materials Fatigue Mathematical models.
	8	Porous materials Fluid dynamics.
	1	Porous materials Fracture.
	1	Porous materials Fracture Mathematical models.
	2	Porous materials Handbooks, manuals, etc.
	1	Porous materials Industrial applications.
	1	Porous materials Industrial applications United States.
	17	Porous materials Mathematical models.
	2	Porous materials Mathematical models Congresses.
	8	Porous materials Mechanical properties.
	4	Porous materials Mechanical properties Congresses.
	2	Porous materials Mechanical properties Mathematical models.
	3	Porous materials Optical properties Congresses.
	3	Porous materials Periodicals.
	10	Porous materials Permeability.
	2	Porous materials Permeability Congresses.
	10	Porous materials Permeability Mathematical models.
	3	Porous materials Permeability Mathematical models Congresses.
	2	Porous materials Testing.
	1	Porous materials Testing Congresses.
Authorized	15	Porous materials Thermal properties
	1	Porous materials Thermal properties Bibliography.

Conclusion: The browse display is superior in terms of organizing materials for understanding and retrieval.

Keyword search for "women" and "sex" as subjects: 7954 hits

Titles Index

Font: Arial Unicode MS

Name: Main Author, Creator, etc.	Full Title	Date
	131-nin no onnatachi no kokuhatsu : Ishihara Tochiji no "babaa hatsugen" saiban kara miete kita mono / Ishihara Tochiji no "Babaa Hatsugen" ni Ikari Shazai o Motomerukai hen.	2005
	2015 : politique nationale de promotion de la femme : pour un développement équilibré homme-femme.	2000
	21-seiki e danjo byōdō suishin Tōkyō puran. Heisei 3-nendo jissai saimoku : josei mondai kaiketsu no tame no Tōkyō-to kōdō keikaku / [henshū Tōkyō-to Seikatsu Bunkakyoku Fujin Seishōnenbu Fujin Keikakuka].	1991
	21-seiki e danjo byōdō suishin Tōkyō puran. Heisei	1992

OK Cancel Search Select All Clear All

7954 Records Found Search: Command=SKEY women AND SKEY sex

The subject headings on any record give a much better idea of what the book is about.

050	0	0	‡a HQ1799.5 ‡b .A12 2000
245	0	0	‡a 2015 : ‡b politique nationale de promotion de la femme : pour un développement équilibré homme-femme.
246	3	0	‡a Politique nationale de promotion de la femme
260			‡a [Antananarivo] : ‡b République de Madagascar, ‡c [2000]
300			‡a 86 leaves ; ‡c 30 cm.
500			‡a "Octobre 2000."
504			‡a Includes bibliographical references (leaves 85-86).
650	0		‡a Women ‡x Government policy ‡z Madagascar.
650	0		‡a Women ‡z Madagascar ‡x Social conditions.
650	0		‡a Women in development ‡z Madagascar.
650	0		‡a Sex role ‡z Madagascar.

The precoordination is important: The government policy is about women, not about sex role or women in development. The book is about social conditions of women, not of Madagascar in general.

Example of Increase in Understanding with Pre-coordination:

1. Title: Advanced editing and finishing techniques in Final Cut Pro 4

Post-coordinated elements created by deconstructing subject headings

Data processing
Digital video
Editing
Final cut (Electronic resource)
Macintosh (Computer)
Programming
Videotapes

Pre-coordinated Subject headings assigned to the same book:

1. Video tapes–Editing–Data processing.
2. Digital video–Editing–Data processing.
3. Macintosh (Computer)–Programming.
4. Final cut (Electronic resource)

2. Title: Between two oceans

Post-coordinated elements created by deconstructing subject headings

Armed Forces
Foreign service
Great Britain
History, Military
Singapore

Pre-coordinated Subject headings assigned to the same book:

1. Singapore–History, Military.
2. Great Britain–Armed Forces–Foreign service–Singapore

Examples of Deconstructing LCSH into Simple Facets

The following examples were provided to the Headings listserv by Adam L. Schiff on Monday, February 12, 2007.

- 650 0 Architecture, Industrial \$z France \$z Arc-et-Senans.
- > 650 0 Glass construction \$z France \$z Paris.
- > 650 0 Deconstructivism (Architecture)\$z Spain \$z Bilbao.
- > 650 0 Architecture \$z France \$y 20th century.
- > 650 0 Library architecture \$z Japan \$z Sendai-shi
- > (Miyagi-ken) \$y 20th century.
- > 650 0 Architecture, Romanesque \$z France \$z Conques (Aveyron)
- >
- > If we break these up, we get something like this (I haven't
- > checked the FAST authority file to see if these are in the form found there):
- >
- > 648 7 20th century. \$2 fast
- > 650 7 Architecture, Industrial. \$2 fast
- > 650 7 Glass construction. \$2 fast
- > 650 7 Deconstructivism (Architecture) \$2 fast
- > 650 7 Architecture. \$2 fast
- > 650 7 Library architecture. \$2 fast
- > 650 7 Architecture, Romanesque. \$2 fast
- > 651 7 France \$z Arc-et-Senans. \$2 fast
- > 651 7 France \$z Paris. \$2 fast
- > 651 7 Spain \$z Bilbao. \$2 fast
- > 651 7 Japan \$z Sendai-shi (Miyagi-ken) \$2 fast
- > 651 7 France \$z Conques (Aveyron) \$2 fast
- >
- > A keyword search on either of the two subject systems would retrieve this
- > work (which is a compilation of 6 films on a DVD). But once you retrieve
- > the record and view it, how is a user to have any clue whether this is
- > about glass construction in Bilbao, Paris, or Sendai-shi? A contents note
- > with the titles of the six films might help in this particular case.
- >
- > Here's another example from our OPAC:
- >
- > 650 0 Human rights \$z Korea (North)
- > 650 0 Refugees \$z Korea (North)
- > 650 0 Koreans \$z China.
- > 651 0 United States \$x Foreign relations \$z Korea (North)
- > 651 0 Korea (North) \$x Foreign relations \$z United States.
- >
- > 650 7 Human rights. \$2 fast
- > 650 7 Refugees. \$2 fast

- > 650 7 Koreans. \$2 fast
- > 650 7 Foreign relations. \$2 fast
- > 651 7 Korea (North) \$2 fast
- > 651 7 United States. \$2 fast
- > 651 7 China. \$2 fast
- >
- > From the FAST system, you cannot tell if this is about human rights in
- > Korea, the U.S., or China. Which country the refugees come from or are
- > in. Who's having foreign relations with whom.
- >
- > Even without any knowledge of LCSH structure and syntax, I think the
- > strings that would display to a catalog user are much more understandable.
- > I will grant that LCSH is not as clear as it could be, since a place in \$z
- > could mean something that is FROM that place or something IN that place,
- > i.e.
- >
- > Refugees \$z Korea (North)
- >
- > in LCSH can mean either or both refugees from North Korea or refugees in
- > North Korea. I'd rather see LCSH refined in some way to make the
- > semantics more clear, perhaps by using an additional word that could be
- > used in display, e.g.
- >
- > Refugees \$i from \$z Korea (North)
- > Koreans \$i in \$z China
- >
- > (I made up a subfield \$i just as one idea that quickly came to mind. It
- > could be ignored in some displays or for searching purposes).
- >
- > Adam
- >
- > ^^^
- > Adam L. Schiff
- > Principal Cataloger
- > University of Washington Libraries
- > Box 352900
- > Seattle, WA 98195-2900
- > (206) 543-8409
- > (206) 685-8782 fax
- > aschiff@xxxxxxxxxxxxxxxxxxxxx

**LC STAFF SURVEY ON SUBJECT CATALOGING
REPORT**

Submitted by Kay Ritchie and Rosa Alicea
October 25, 2006

In August 2006, CPSO conducted a survey in order to better understand the ideas and views of LC cataloging staff on Library of Congress Subject Headings (LCSH), Library of Congress Classification (LCC), and related training issues. The goal of the survey was to utilize the data collected from the staff to make feasible improvements in various areas of subject cataloging. These improvements should help to make subject cataloging an easier, more efficient process, along with achieving a high level of quality and productivity. The following is a summarized report of the survey that may help CPSO to implement the ideas and suggestions received from the cataloging staff.

The survey was the result of work done by the following detailees to CPSO: Kay Ritchie and Rosa Alicea, responsible for survey development and distribution, collection and analysis of the survey data, and summary report; along with Thomas Bishop and Elmer Klebs, who were also responsible for initial survey development.

Initially, cataloging directorate team leaders were asked to identify subject catalogers, including staff from Geography and Map, Prints and Photograph, MBRS, and the Folklife Center. 50% responded to the survey out of the 288 staff members who were identified by their team leaders for participation.

The survey was divided into two main areas: training and cataloging. Training issues were addressed from the trainee perspective, but also from those who have served as trainers. The second portion of the survey focused on cataloger's ideas and comments in a variety of areas. The survey was structured so that catalogers could comment on the various identified areas of subject cataloging, along with having suggested areas where participants could check off items from a form in each category. As they were received, individual survey forms were each given an arbitrary identification number in order to manage the processing of the data. Responses were compiled using EXCEL software to analyze the data, with participant's comments compiled by general topic into WordPerfect documentation.

Staff were mainly identified in three groups by years of experience: 1-4, 5-14, and 15+ years. In addition, for purposes of analysis, the type of cataloging assignment was also identified: whole book, subject only, and, other (e.g., team leaders who had previous subject cataloging experience). No technicians were identified as participants for the survey.

TYPE OF ASSIGNMENT AND LENGTH OF SUBJECT-RELATED EXPERIENCE

The total number of survey participants identified by years of experience, and, type of cataloging assignment is the following:

15+ years of experience			
Whole book		Subject only	Other
30%		4.0%	11%
5-14 years of experience			
Whole book		Subject only	Other
27%		00%	4.0%
1-4 years of experience			
Whole book		Subject only	Other
25%		1.0%	00%

AREAS AND FOCUS OF SUBJECT CATALOGING WORK

Participants were asked to identify the extent and type of their subject cataloging activities. The total results for all three groups are as follows:

	Never	Rarely	Occasionally	Often	
Extensively					
Assign subject hdgs	1.0%	1.0%	4.0%	20%	73%
Propose subject hdgs	15%	15%	39%	20%	11%
Assign class numbers	8.0%	2.0%	5.0%	18%	67%
Develop class numbers	25%	18%	31%	15%	11%
Shelflisting	24%	7.0%	8.0%	13%	47%

The focus of subject-related work that was compiled indicated that: 68% of survey participants are working in specific subject areas; 32% are working in all subject areas, i.e., A-Z classes.

TRAINING FOR SUBJECT CATALOGING

Participants were asked to identify, in as many categories as needed, the types of subject and/or classification training that they found useful as applied to their actual day-to-day work. The following types of training environment were identified as being the most useful:

- 92% One-on-one training
- 63% Formal classroom training

Participants were asked about specific areas that need improvement in subject cataloging training. The overall responses include the following:

- 51% Training was not extensive enough; could have used more time to master new concepts, cataloging tools, reference sources, etc.
- 24% Fundamental cataloging principles were not adequately covered
- 44% More time should be spent on subject analysis

- 39% More emphasis should be given to the rules covering subdivisions of subject headings
- 35% More in-depth coverage should be given to subject authorities
- 33% More emphasis should be given to reference sources used in subject cataloging

In addition, other participants' comments include the following:

- ◆ Training should provide more in-depth coverage of class schedules (their organization, etc.) and class proposals
- ◆ Training should provide more in-depth coverage of the Subject Cataloging Manual (SCM)

RULES FOR CREATING SUBJECT HEADINGS AND CLASSIFICATION NUMBERS

Participants were asked to rate levels of difficulty in three areas: constructing subject heading strings, assigning classification numbers, and, proposing new subject headings. The four possible survey choices were: easy, somewhat difficult, difficult, and very difficult.

The overall results are:

Constructing new subject heading strings

56% Somewhat difficult

15% Difficult or very difficult

Rules for assigning classification numbers

52% Somewhat difficult

9.0% Difficult or very difficult

Proposing new subject headings

41% Somewhat difficult

28% Difficult or very difficult

Participants were also asked what caused them the most problems in subject cataloging. By general category, the results are as follows:

LC Classification/Tables

25% LCC needs better structure and is often confusing to use

20% Instructions are often not clear within the classification schedule

20% Indicated that they had trouble knowing when to propose a new class number

Other participants' comments are:

- ◆ The overall structure of Class Web is often confusing because it is hard to see the indentions (i.e., the continuation of areas and relation of one part of the schedule to another)
- ◆ More current terminology should be used. For example, some catalogers don't understand that a "CF." does not mean to use the other number; they regard it as a "see reference" not a "consult" reference. It's better to use "see also" instead of assuming that everyone knows Latin or academic usage
- ◆ Tables are generally problematic, and need clarification to be understood more readily
- ◆ The enhanced browser in Class Web is not consistently applied for all tables in the schedules

Subject classes identified by participants as causing problems are chiefly: Class H schedule and Tables; Class K schedule and Tables; Class N schedule; and, Class P schedule.

Subject Cataloging Manual (SCM)

37% Indexing needs to be improved

38% Some memos are confusing, and more examples are needed

Other participants' comments include:

- ◆ Need complicated examples instead of just easy ones for geographic subdivisions
- ◆ The SCM documentation is too dense

Subdivisions of Subject Headings

24% Rules for adding subdivisions are not clearly written or are confusing

38% Applying the rules for free-floating subdivisions is difficult

20% Applying the rules for indirect geographic subdivisions is difficult

Several participants commented on a need for better consistency, for example:

- ◆ Some subdivisions need to be more broadly applied to achieve better consistency
- ◆ For the sake of efficiency, subdivisions like "sociological aspects" and "social aspects" could have their usage given in the free-floating subdivisions manual, instead of having to consult individual memos
- ◆ Problematic application of designations for subfield "v" vs. subfield "x" - Where to place indirect geographic names within strings

Subject Heading Proposals

23% Responded that they don't make new proposals because it's too time-consuming and/or problematic

31% Responded that the instructions are not clear, had trouble with the proposal workflow process, and/or they had trouble knowing when to propose a new subject heading.

Broader, Narrower, and Related Terms in Subject Authorities

- 18% More examples are needed for the sake of clarity and efficiency in creating and updating the authorities

Several participants commented that some problems are due to the nature of the language, for example, sometimes a term can be simultaneously used as a broad/narrow term. Rather than trying to establish elaborate hierarchies, a “see also” reference would be sufficient for the user.

Subject Heading Authority Changes

- 15% Considered the workflow process too time-consuming or problematic
- 13% Would like to have more examples of actual authority records included in the instructions
- 11% Reported that instructions are not clearly written and/or they have trouble using the online tool

Keeping Current with Changes, Updates, Weekly Lists, LCSH Development, etc.

- 56% Responded that keeping current is moderately difficult

The two reasons most often given were: not enough time in the workday to spend on it, and, the pressure to produce cataloging statistics doesn't allow it.

Subject Terminology that is Arcane or Outdated

- 26% Agreed that terminology used in subject headings is arcane and/or outdated

Scope Notes in Subject Headings

- 57% Agreed or strongly agreed that more scope notes are needed in subject authorities. If present, they are sometimes not written in a clear manner.

Participants were also asked to identify their single greatest problem in subject cataloging.

The two general areas that were most often given are:

- ◆ Subject analysis, i.e., knowing when to propose a new heading or develop a new class, what headings to assign to an item
- ◆ A-Z cataloging concerns, i.e., wanting to do a good job at cataloging an item, but not having enough subject expertise in a certain area to feel confident about it

Quality of Subject Cataloging

Participants were asked to identify their personal impression of the quality of subject cataloging.

The choices given in the survey form were: excellent (have no concerns about it); good overall (only minor concerns); problematic in certain areas that concern me; and, the worst that I've ever seen—a great need for improvement overall.

Collectively, the participants responded to this question as follows:

- 3.0% Quality is excellent; have no concerns about it
- 41% Quality is good overall; have only minor concerns
- 41% Problematic in certain areas that concern me
- 14% Worst that I've ever seen; there is a great need to improve overall quality

Ways to Improve the Quality of Subject Cataloging

Participants were asked to recommend ways to make improvements in subject cataloging quality.

- 35% Responded that more extensive training is needed for catalogers
- 33% Responded that refresher courses and brown bag lunches, etc. be offered on special topics related to subject cataloging
- 33% Would like subject cataloging forums with topic of special interest offered
- 34% Agreed that regular feedback be given on errors being made in order to improve quality
- 28% Responded that “cheat sheets” and other handouts be created to help make workflow more efficient and achieve higher quality
- 22% Agreed that technology should be better utilized in the directorate to help eliminate errors
- 20% Agree that quality control should be implemented within the team environment

Other suggestions for improving quality are:

- ◆ Would like to see the software create the strings and the cataloger only has to enter data elements without having to worry about the order
- ◆ Collapse documentation so that there are fewer places to consult for answers to questions
- ◆ Improve the efficiency and scope of Validator processing, e.g., currently Validator incorrectly approves a 650 field with a geographic name in the subfield “a” position, instead of it being a 651 field
- ◆ Utilize spell check in Validator to help eliminate typos
- ◆ Give regular feedback to staff for errors being made

- ◆ Make the rules for indirect geographic subdivisions more consistent for the sake of efficiency
- ◆ Consistently apply the enhanced browser in Class Web to all tables
- ◆ Improve the Class Web display so that it is easier to read the hierarchy, i.e., alignment, better spacing, various fonts, bolding, grid background, etc.
- ◆ Add guide card information to the shelflist to improve accuracy and efficiency
- ◆ Add more complex examples to the SCM documentation
- ◆ Improve the indexing of the SCM
- ◆ Provide a presentation on scope notes, including technical aspects, and information contained in SCM H400 documentation
- ◆ Would be helpful if the Weekly List was cumulatively indexed
- ◆ Archive the CPSO specialist's meeting minutes
- ◆ Create "office hours" so that catalogers can consult with CPSO specialists
- ◆ Simplify the subject and class proposal workflow processes
- ◆ Offer refresher courses and provide general training for A-Z cataloging to familiarize staff members with basic subject headings and peculiarities of certain classes
- ◆ Offer a refresher course in the basic principles of subject analysis, so that catalogers now expected to do A-Z classes are able to apply their skills to a broad range of materials, subjects, and formats. With a better understanding of the principles of subject analysis, catalogers will be better able to scrutinize incoming copy cataloging records and more efficiently process them.