

# Author-Assigned Keywords versus Library of Congress Subject Headings

## Implications for the Cataloging of Electronic Theses and Dissertations

By C. Rockelle Strader

*This study is an examination of the overlap between author-assigned keywords and cataloger-assigned Library of Congress Subject Headings (LCSH) for a set of electronic theses and dissertations in Ohio State University's online catalog. The project is intended to contribute to the literature on the issue of keywords versus controlled vocabularies in the use of online catalogs and databases. Findings support previous studies' conclusions that both keywords and controlled vocabularies complement one another. Further, even in the presence of bibliographic record enhancements, such as abstracts or summaries, keywords and subject headings provided a significant number of unique terms that could affect the success of keyword searches. Implications for the maintenance of controlled vocabularies such as LCSH also are discussed in light of the patterns of matches and nonmatches found between the keywords and their corresponding subject headings.*

The usefulness of controlled vocabulary has been debated for a number of years. The question has come even more to the forefront with the popularity of online tools such as Google and the use of keywords as users' primary search strategy. For libraries, the debate also centers on whether controlled vocabularies, such as Library of Congress Subject Headings (LCSH), are worth the time (and associated expense) of assigning and adding to bibliographic records in catalogs and databases. Studies on the issue focus primarily on users as seekers of information and examine keyword terms as used in searches. Few studies exist that examine the use of keywords assigned by authors of online documents. The present study is intended to contribute to the literature on this issue of keywords versus controlled vocabularies in online catalogs and databases.

**C. Rockelle Strader** (strader@osu.edu) is Catalog Librarian, The Ohio State University Libraries, Columbus.

The author presented a report about this project to the Association for Library Collections and Technical Services Cataloging and Classification Research Interest Group at the American Library Association Midwinter Meeting in Denver, Colorado, January 25, 2009.

Submitted February 27, 2009; returned April 2, 2009 with request to revise and resubmit; resubmitted May 28, 2009; reviewed and accepted for publication June 26, 2009.

### Literature Review

Several studies have addressed the uses of controlled vocabulary versus keywords in users' catalog searches. A representative selection will be reviewed here to

provide context for the current project. Carlyle conducted a study matching catalog users' search terms with LCSH in which 47 percent of the search terms matched exactly.<sup>1</sup> When including partial matches, word order variations, and spelling variations, the figure rose to 74 percent. Only 5 percent of users' search terms could not be matched at all. The remaining 21 percent were matches that required two or more LCSH terms to cover the search term. In this study, users' searches were done through subject search fields, not general keyword searches, which were not available at the time of the study. Carlyle concluded that a maximum 74 percent match rate was not an acceptable performance for LCSH and that further analysis of LCSH vis-à-vis user language was needed. The study is important because it defined levels of matching and called both for better matching against cross-references and for making LCSH semantically more flexible.

Frost investigated the utility of keywords taken from titles as "entry vocabulary" to subject searches by examining the degree of match between title keywords and controlled vocabulary.<sup>2</sup> Matches could be exact over the entire heading in direct order (11 percent of Frost's sample), in any order (30 percent), exact main heading only (12 percent), exact in subdivision (5 percent), truncated variant in main heading (14 percent) or subdivision (1 percent), or no match at all (27 percent). Thus matches of some type occurred in 73 percent of the titles in her sample, leaving the remaining 27 percent with no matches at all. Frost concluded that keywords and subject headings are complementary.

Ansari replicated Frost's study using medical dissertations written in Farsi.<sup>3</sup> Her findings were very close to Frost's; 70.3 percent of Ansari's terms were matches of some type and 29.7 percent did not match at all, compared to Frost's 73 percent and 27 percent, respectively. Ansari also concluded that keywords and descriptors are complementary and that keywords for which there is no matching descriptor should be considered for addition to indexing lists.

Voorbij conducted a study of title keywords and subject descriptors using somewhat different criteria for comparison.<sup>4</sup> His focus was on comparing the descriptors to the keywords rather than comparing keywords to descriptors. His aim was to determine how well subject descriptors enhanced bibliographic records. The comparison defined matches in thesaural or semantic terms instead of using Frost's more literal use of LCSH construction (i.e., main headings and subdivisions) and spelling. Voorbij categorized the results as exact match, synonym, broader term, narrower term, related term, some relation but difficult to determine, and no match. The first three categories, constituting 59.6 percent of the results (629 of 1055 descriptors), were not considered enhancements to the record. The remaining 426 descriptors (40.4 percent) were examined for the degree to which they enhanced the bibliographic record. Initially all 426 were considered as "possibly enhancing"

the bibliographic record; this included 24.4 percent in the "no match" category. Further subjective examination determined that within the remaining 426 descriptors, 342 (33.0 percent of the sample) could be said to "slightly enhance" the bibliographic record, and 241 (23.2 percent) could be regarded as "considerably enhancing" it. Like Frost, Voorbij concluded that title keywords and descriptors are complementary, noting that descriptors help to reduce irrelevant hits and boost precision as well as to group synonymous terms.<sup>5</sup> He further acknowledged that adding descriptors is an expensive activity that must be subjectively weighed against the value of precision and collocation.

Gross and Taylor examined transaction logs of users' searches to see if controlled vocabulary provides additional keywords and consequently enhances both recall and precision in keyword searches of a catalog.<sup>6</sup> Findings indicated an increase of up to 30 percent in the recall of relevant documents by the use of controlled vocabulary; about one-third of the keyword searches examined would have failed if the controlled terms had not been present. This percentage is similar to that of the "no match" category in Frost's and Voorbij's studies of title keywords and controlled vocabularies.<sup>7</sup>

Garrett studied the impact of adding subject headings to records in the Eighteenth Century Collections Online database of full-text documents.<sup>8</sup> Preliminary results indicated that some 60 percent of searches would have failed if subject headings had not been present in the record. Terms, such as "sanitation," that are common now were not used in the original documents and would not be retrieved without the cross-reference structure provided by controlled vocabularies.

Little has been written about author-assigned keywords. Two studies touch on them: one by Kipp and one by Gil-Leiva and Alonso-Arroyo.<sup>9</sup> Kipp compared user tags with author-assigned keywords and indexer-assigned descriptors for 165 journal articles. Matching was done on a hierarchical scale (similar to Voorbij's) of thesaural relationships, including same, synonym, broader term, narrower term, related but not in thesaurus, and not related. The focus of the study was on user tags and did not break out statistics specifically related to author-assigned keyword matches. In this study, 44.5 percent of all terms fell into the category of "related but with some ambiguity in the relationship . . . as well as relationships that were not formally in the thesaurus."<sup>10</sup> Kipp concluded that tags, as well as keywords and descriptors, can be valuable as additional access points.

Gil-Leiva and Alonso-Arroyo performed a matching study of author-assigned keywords and indexer-assigned descriptors for journal articles in four databases.<sup>11</sup> This study found an average of 24.59 percent for exact matches of keywords with descriptors and up to 45.66 percent when adding "normalized" matches (terms similar in meaning).

By inference, some 54 percent of the keywords did not match, a far greater rate of nonmatch than that found in the studies related above. The authors concluded that keywords are valuable sources of information for indexers.

The debate between controlled vocabularies and keywords may be framed in terms of the issues involved with the formation (and subsequent maintenance) of new controlled terms for use by catalogers and the use of uncontrolled terms by users. As noted above, keywords may be used as guides for the creation of controlled terms, which could affect the maintenance of controlled vocabularies such as LCSH. LCSH is maintained on the principles of “literary warrant.” Historically, literary warrant for LCSH meant that terms were derived from the materials held by the Library of Congress and has since been expanded to include contributions by Subject Authority Cooperative (SACO) member libraries.<sup>12</sup> The standard for controlled vocabularies, ANSI/NISO Z39.19-2005, states that “the word or phrases chosen should match as closely as possible the prevailing usage in the domain’s literature.”<sup>13</sup> Contrasting literary warrant is “user warrant,” which is defined by the ANSI/NISO standard as “generally reflected by the use of terms in requests for information on the concept or from searches on the term by users of an information storage and retrieval system.”<sup>14</sup> The ANSI/NISO standard presents literary warrant and user warrant as complementary guiding principles for turning keywords into controlled terms on the basis of current literature as well as the use of terms by users who may or may not be familiar with the discipline in which they are seeking information.

## Research Method

This study investigated the following questions:

- How well do author-assigned keywords match LCSH (either the established heading or a “see from” reference)?
- Conversely, how well do LCSH match keywords used by authors of electronic theses and dissertations (ETDs)?
- How many keywords are unique to their respective bibliographic records? Do these keywords add sig-

**Table 1.** Categories of Match

<b>Exact Match</b>	Exact match Exact match of cross-reference
<b>All Present</b>	All present, but not in exact order All present, but not in exact order, in cross-reference
<b>Partial Match</b>	Partial match Partial match of cross-reference
<b>Needs 2 LCSH</b>	KW covered by 2 LCSH, but if either LCSH were missing there would be only partial match Part of KW in main LCSH, while remainder is covered by cross-reference of another LCSH KW covered by cross-references of 2 LCSH
<b>Variant</b>	Variant, separated from “n” to accommodate possibility of truncation, etc. Variant of cross-reference Variant is abbreviation (e.g., chemical symbol such as CO <sub>2</sub> for carbon dioxide)
<b>No Match</b>	No match/not present

nificantly more relevant terms that may increase the likelihood of their respective ETDs being found?

- Likewise, how many LCSH are unique within their respective bibliographic records; that is, how many LCSH are assigned for which there are no corresponding author-assigned keywords? Do LCSH add significantly more unique terms that may aid in the retrieval of the ETDs to which they are assigned?
- What are the implications for the way LCSH is used? What conclusions may be drawn regarding the construction or maintenance of LCSH?

Answers to these questions may corroborate the results of the studies related above and may further be used to draw conclusions regarding the use of both cataloger-assigned terms and author-assigned keywords for enhancing catalog searches.

The current project’s data set consisted of 285 eligible ETDs submitted by Ohio State University (OSU) doctoral candidates to the OhioLINK ETD Center and their associated bibliographic records in OSU’s online catalog. Eligible titles were those for which automatic e-mail notification of availability was received by catalogers in OSU Libraries’ Cataloging Department between June 1 and October 31, 2005, had author-assigned keywords, and had full text available at the time of cataloging. The cataloging of these titles was finished in 2006. Following interruptions due to a major building renovation, data collection and analysis were conducted in late 2007 through mid-2008.

The data were collected by visual inspection of the metadata page for each eligible ETD in the OhioLINK ETD Center and its bibliographic record in OSU Libraries’ online catalog, as well as the authority record for each LCSH as found through OCLC’s Connexion Client. These data included the author-assigned keywords in the ETDs,

LCSH supplied in the bibliographic records, and “see from” references as indicated in the authority files. The data were recorded in Excel spreadsheets for collocation, counting, and comparisons. A total of 1,681 author-assigned keywords and 1,181 LCSH terms were collected.

To address the research questions presented earlier, the collected keywords and associated LCSH terms were assessed to answer the following working questions:

- How many keywords exactly matched LCSH, that is, could be placed in the 600, 610, 611, 650, or 651 MARC fields (fields for controlled vocabulary)?
- How many keywords were LCSH “see from” references?
- How many keywords could or could not be converted to LCSH, that is, could be placed only in a 653 field (field for uncontrolled terms)?
- How many LCSH terms had or did not have corresponding author-assigned keywords?
- How many keywords and LCSH terms could or could not be matched to corresponding words in titles and abstracts?

To categorize and codify the data, the categories of match in table 1 were used. Where more than one interpretation existed of how a keyword could be matched with a corresponding LCSH and vice versa, a rule was established to prefer the category of match in the order (top to bottom) shown in table 1.

## Results and Discussion

The results of the comparisons of keywords and LCSH with each other and the matching of both in titles and abstracts yielded some patterns as well as several differences. As noted in the previous section, the

**Table 2.** Average, Mode, Maximum, and Total Keywords and LCSH Per Title

	Average	Mode	Max	Total
<b>KW/title</b>	5.9	5	57	1681
<b>LCSH/title</b>	4.1	4	13	1181

**Table 3.** Raw Counts of Keywords and LCSH Matches

Raw Counts of Keywords		Number
Category		
Keyword exactly matched by LCSH		333
Keyword exactly matched only in LCSH cross-references (4xx in authority record)		90
All keywords in LCSH but not exact word order		50
All keywords only in LCSH cross-references but not in exact order		7
Keyword partially matched by LCSH		365
Keyword partially matched only in LCSH cross-references		50
All keywords covered by 2 LCSH		26
All keywords covered by 2 LCSH including cross-references		10
All keywords covered only in cross-references of 2 LCSH		2
Variant form/spelling of keywords found in LCSH		145
Variant form/spelling of keywords found in LCSH cross-references		8
Variant is an abbreviation (e.g., chemical symbol)		14
Keyword not matched or covered in any form		581
Total		1,681
Raw Counts of LCSH		Number
Category		
LCSH exactly matched keyword		347
Cross-reference exactly matched keyword		84
LCSH completely covered keyword but not in exact order		47
Cross-reference completely covered keyword but not in exact order		6
LCSH partially matched keyword		324
Cross-reference partially matched keyword		43
LCSH is/contained variant of keyword		119
LCSH is/contained abbreviation of keyword		9
Cross-reference is/contained variant of keyword		6
LCSH did not match any keyword		196
Total		1,181

total number of keywords was 1,681 and the total number of LCSH was 1,181. The average number of keywords per title was 5.9 (mode, 5), while the average number of LCSH per title was 4.1 (mode, 4). However, there was a stark contrast between the maximum number of keywords (57) that were assigned to a title and the maximum number of LCSH (13); see table 2.

Table 3 shows the raw counts of keyword and LCSH matches. The percentages of the six broad categories—exact match, all present (in a single heading), all present (needing two LCSH), partial match, variants, and no match—of keyword matches to LCSH are presented in table 4 and include the matches to cross-references, to more than one LCSH, and to abbreviations. Tables 3 and 4 summarize the data that address the issue of how well author-assigned keywords match LCSH and serve to answer the first three working questions, that is, how many keywords matched LCSH, how many keywords matched only cross-references, and how many keywords did not match LCSH. A total of 44.49 percent of the author-assigned keywords did not match cataloger-assigned LCSH (34.56 percent had no matches; 9.93 percent were variant forms); see table 4.

One explanation for the large percentage of terms not covered by cataloger-assigned LCSH is that LCSH has not kept up with current research. This issue of maintenance has been a recurring criticism of LCSH over the years.<sup>15</sup> LCSH typically are established from evidence of a new topic found in the piece in hand, that is, from literary warrant. This is usually a monograph in hand, since articles and chapters are generally not cataloged.<sup>16</sup> However, in some disciplines, such as the physical sciences and medicine, the journal literature is the primary publication environment for new research, and dissertations in those fields could be among the first comprehensive monographic treatments of a topic that has been otherwise extensively discussed.

Further, the distinction is becoming blurred as articles and chapters are added to bibliographic databases such as WorldCat. Although these resources are placed in research databases to aid discovery, they usually are not formally cataloged and thus are not considered as sources for new controlled terms. Yet they typically contain current terms of the disciplines in which they are written and which may or may not be familiar to users who need those resources. These terms are

**Table 4.** General Categories of Keyword Matches to LCSH

<b>Keywords Matched to LCSH (including cross-references)</b>	
<b>Category</b>	<b>%</b>
Keyword matched exactly by LCSH	25.16
Keyword matched, but not in order (single heading)	3.39
Keyword matched, but not in order (needing two LCSH)	2.26
Keyword partially matched	24.69
Keyword were variant forms	9.93
Keyword not found in LCSH at all	34.56
Total	99.99*
<b>Keywords Matched to LCSH Cross-References Only (4xx)</b>	
<b>Category</b>	<b>%</b>
Keyword matched cross-reference exactly	5.35
Keyword matched cross-reference in any order	0.42
Keyword partially matched cross-reference	2.97
Variant of keyword matched cross-reference	0.48
Keyword covered by 2 LCSH, in one or both cross-reference	0.71
Total % of keyword matches in any form to LCSH cross-references	9.93

\* Does not equal 100% because of rounding.

**Table 5.** General Categories of LCSH Matches to Keywords

<b>LCSH Heading Matched to Keywords</b>	
<b>Category</b>	<b>%</b>
LCSH matched keyword exactly	36.49
LCSH matched keyword, not in order	4.49
LCSH partially matched keyword	31.08
LCSH was variant form	11.34
LCSH did not match any keywords	16.60
Total	100.00
<b>LCSH Cross-References Matched to Keywords</b>	
<b>Category</b>	<b>%</b>
LCSH cross-reference matched keyword exactly	7.11
LCSH cross-reference matched keyword, not in order	0.51
LCSH cross-reference partially matched keyword	3.64
LCSH cross-reference was variant	0.51
Total % of LCSH cross-references matched to keywords	11.77

uncontrolled keywords that users may be likely to search on first. This use of terms for the purpose of searching is the essence of user warrant.<sup>17</sup> As full-text access to articles and chapters becomes increasingly easier and ubiquitous, should these resources be considered as valid sources for controlled terms?

Another explanation for the unmatched keywords could be the use of different terminology for similar concepts, an issue not examined in this study. In other words, a match may not have occurred because of a lack of cross-reference in a related or semantically equivalent term, implying a

different need for the maintenance of LCSH. This implication corroborates Carlyle's conclusion about the need for the maintenance of cross-references to reflect changing user language.<sup>18</sup>

The large nonoverlap also could imply that some keywords may be spurious or not topical in nature. For example, one keyword that was used, "MD/PhD," does not describe the topic of the document, but rather the type of degree program in which the author was enrolled.

Other keywords, such as "grounded theory," may not have been matched because of the cataloger's judgment of the relevance of the term to the topic of the given ETD. The cataloger may have considered such terms to be methodological and not topical. However, in some cases discussion related to such terms in the document was significant, and the terms in question could be seen to warrant inclusion in the bibliographic record.

The question of how well LCSH terms match keywords used by ETD authors was also addressed by the first three working questions as well as the specific working question of how many LCSH did or did not have corresponding author-assigned keywords. The data to address these questions are presented in table 5, which shows the broad categories of LCSH matches to keyword. The bottom half of the table shows the LCSH cross-reference matches to keyword.

As shown in table 5, 36.49 percent of the cataloger-assigned LCSH matched author-assigned keywords exactly and only 16.60 percent did not match any keywords while 31.08 percent were partial matches and 11.34 percent were variant forms of the keywords. The low total of variant matches and nonmatches could imply that keywords are used to guide the catalogers' assignment of LCSH, consistent with the findings of Ansari, and Gil-Leiva and Alonso-Arroyo.<sup>19</sup> Keywords, as assigned by the authors, could be seen to reflect the current use of

**Table 6.** Keyword and LCSH Matches in Title and Abstract

Keyword Matches in Title and Abstract		
Category	% in Title	% in Abstract
Keyword exactly matched	26.23	54.61
Keyword matched, but not in order	2.8	10.11
Keyword partially matched	16.06	15.94
Variant of keyword	11.12	8.74
Keyword not present at all	43.78	10.59
Total*	99.99	99.99
LCSH Base Matches in Title and Abstract		
Category	% Base in Title	% Base in Abstract
LCSH exactly matched	14.14	26.84
LCSH matched, but not in order	2.12	10.75
LCSH partially matched	21.42	16.93
Variant of LCSH	12.36	14.39
LCSH not present at all	49.96	31.08
Total*	100.00	99.99
LCSH Subdivision Matches in Title and Abstract		
Category	% Subdivision in Title	% Subdivision in Abstract
LCSH subdivision exactly matched	15.14	31.13
LCSH subdivision matched, but not in order	0	3.62
LCSH subdivision partially matched	10.66	17.06
Variant of LCSH subdivision	6.4	8.96
LCSH subdivision not present at all	67.8	39.23
Total	100.00	100.00

\* Some totals do not equal 100% because of rounding.

terms in a field and can be used as points of entry for both users and catalogers. Where keywords can be translated into existing LCSH, the controlled vocabulary and cross-reference structure can then allow for meaningful sorting and organization (or “triage,” as Sclafani describes it) of search results.<sup>20</sup>

In light of the professed advantages of cross-references, however, the effect of cross-references in this study was not as great as expected, although still noticeable. The total percentage of keyword matches in any form to LCSH cross-references was 9.93 percent (table 4), while the total percentage of matches of LCSH cross-references to corresponding keywords was 11.77 percent (table 5).

To answer the final working question (regarding uniqueness of terms within the bibliographic record), data were collected on the presence of the keywords and subject headings in their respective titles and abstracts. As with the keyword to LCSH matching procedure, exact and partial matches were counted as well as singular and plural differences and other variants that could affect user-search results. However, the LCSH matching procedure was varied for this portion of the study. In the previous parts of the study, base terms and subdivision strings were kept together, but for this part of the study the base terms and subdivisions were treated separately. This was done for two reasons. First, subdivided LCSH are not natural language phrases as keyword phrases were in this population of documents; exact matches over entire subdivided LCSH did not occur. Second, most of the assigned LCSH (712, or 60.29 percent) were not subdivided; that is, they were base terms only, and consequently the subdivisions were separated out to allow for the collocation of the data across all collected base terms. The percentages for the subdivisions are derived from the remaining 469 LCSH (39.71 percent) that contained them. Table 6 shows the percentages of matches that were found in titles or abstracts.

While conducting this study, the investigator learned that ETD authors were discouraged from using or relying on the titles of their works when selecting keywords. The degree to which this practice affected the results is unknown. The fact that 43.78 percent of the keywords had no match in the title and another 11.12 percent had only a variant match may reflect this instruction. Conversely, no correlation may exist. This possibility is consistent with the finding that 49.96 percent of assigned LCSH were not matched in the title and 12.36 percent were present as a variant. Further, titles are inherently limited in wording, and consequently contain a restricted number of words that could be repeated in keywords and LCSH assigned to the work

A notable result occurred when keywords and LCSH were matched against abstracts, which are included in the bibliographic records for OSU ETDs. Author-assigned keywords exactly matched words in the abstract 54.61 percent

**Table 7.** Selections from 57 Author-Assigned Keywords from One Document

artists of picture books	picture books in art education
artists of picturebooks	picture books in education
design	picture book design
the history of design	picturebook design
how picture books work	the postmodern in picturebooks
how picturebooks work	the postmodern in picture books
illustration	the post modern in picture books
the history of illustration	text and image
meaning in picture books	text and image relationships
meaning in picturebooks	the history of children's literature

of the time, while cataloger-assigned LCSH exactly matched only 26.84 percent of abstract words. Keyword nonmatches occurred 10.59 percent of the time, and cataloger-assigned LCSH nonmatches occurred 31.08 percent of the time. Put another way, only about one-tenth of the keywords and roughly one-third of the assigned LCSH are unique to the bibliographic records. This result corroborates Gross and Taylor's findings in which more than one-third of the user searches that they examined would have failed if LCSH were not present in the records found.<sup>21</sup> In terms of the discoverability of bibliographic records, the use of LCSH significantly complements keywords by providing further unique terms for searching and matching, even in the presence of enhancements such as abstracts.

The data gathered in this study suggest that authors performed rather effectively (when compared to assigned LCSH) in providing relevant keywords. A total of 65.44 percent of author-assigned keywords matched exactly, partially matched, or were variant forms of LCSH. Indeed, as noted above in relation to table 5, only 16.60 percent of the cataloger-assigned LCSH did not have corresponding author-assigned keywords. Authors, however, were not always concise about assigning keywords. One author assigned 57 keywords (the maximum noted in table 2), many of which are redundancies to capture variants. Table 7 shows a sample of these redundancies found in that record. One could consider this as an exemplar to demonstrate the value of controlled vocabulary.

## Conclusion

In this study, LCSH demonstrated their potential to provide unique access points for approximately one-third of searches,

even in the presence of bibliographic enhancements such as abstracts. Keywords provide a similar benefit, although not as strong, since they more often duplicate terms that appear in abstracts. Abstracts in the bibliographic records for ETDs are the norm for the OSU online catalog, but elsewhere this is likely not the case. Consequently, both LCSH and keywords provide significant numbers of unique terms that may increase the discoverability of ETDs in a catalog where abstracts are not present. Evidence of this can be seen by the number of nonmatches (i.e., unique terms) in the title-only comparisons of LCSH (49.96 percent) and keywords (43.78 percent). LCSH has the added benefit of collocating ETDs with like materials in other formats in the catalog.

The currency of research as found in dissertations represents a challenge to controlled vocabularies such as LCSH. Literary warrant, as it is currently practiced, makes it difficult for such systems to keep up with the pace of new research. Keywords may compensate for this lagging behind, which is inherent in the maintenance of controlled vocabularies, by serving as entry points into the catalog and as guides for the assignment of controlled terms that have already been established. This study corroborates the findings of much of the research on controlled vocabulary and uncontrolled keywords, showing that they are complementary tools for helping users find the materials that they need.

### References and Notes

1. Allyson Carlyle, "Matching LCSH and User Vocabulary in the Library Catalog," *Cataloging & Classification Quarterly* 10, no. 1/2 (1989): 37–63.
2. Carolyn O. Frost, "Title Words as Entry Vocabulary to LCSH: Correlation between Assigned LCSH Terms and Derived Terms from Titles in Bibliographic Records with Implications for Subject Access in Online Catalogs," *Cataloging & Classification Quarterly* 10, no. 1/2 (1989): 165–79.
3. Mariam Ansari, "Matching Between Assigned Descriptors and Title Keywords in Medical Theses," *Library Review* 54, no. 7 (2005): 410–14.
4. Henk J. Voorbij, "Title Keywords and Subject Descriptors: A Comparison of Subject Search Entries of Books in the Humanities and Social Sciences," *Journal of Documentation* 54, no. 4 (Sept. 1998): 466–76.
5. Frost, "Title Words as Entry Vocabulary to LCSH."
6. Tina Gross and Arlene G. Taylor, "What Have We Got to Lose? The Effect of Controlled Vocabulary on Keyword Searching Results," *College & Research Libraries* 66, no. 3 (May 2005): 212–30.
7. Frost, "Title Words as Entry Vocabulary to LCSH"; Voorbij, "Title Keywords and Subject Descriptors."
8. Jeffrey Garrett, "Subject Headings in Full-Text Environments: The ECCO Experiment," *College & Research Libraries* 68, no. 1 (Jan. 2007): 69–81.
9. Margaret E. I. Kipp, "Complementary or Discrete Contexts in Online Indexing: A Comparison of User, Creator, and Intermediary Keywords," *Canadian Journal of Information and Library Science* 29, no. 4 (Dec. 2005): 419–36; Isidoro Gil-Leiva, and Adolfo Alonso-Arroyo, "Keywords Given by Authors of Scientific Articles in Database Descriptors," *Journal of the American Society for Information Science and Technology* 58, no. 8 (June 2007): 1175–87.
10. Kipp, "Complementary or Discrete Contexts in Online Indexing," 429.
11. Gil-Leiva and Alonso-Arroyo, "Keywords Given by Authors of Scientific Articles in Database Descriptors," 1176.
12. Library of Congress, Cataloging Policy and Support Office, *Library of Congress Subject Headings: Pre- vs. Post-Coordination and Related Issues* (Washington, D.C.: Library of Congress, 2007): 3, [www.loc.gov/catdir/cps/po/pre\\_vs\\_post.pdf](http://www.loc.gov/catdir/cps/po/pre_vs_post.pdf) (accessed Feb. 25, 2009).
13. National Information Standards Organization (NISO), *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies* (Bethesda, Md.: NISO, 2005): 16.
14. Ibid.
15. In addition to Carlyle, "Matching LCSH and User Vocabulary in the Library Catalog," see, for example, Hope O. Olson and John J. Boll, *Subject Analysis in Online Catalogs*, 2nd ed. (Englewood, Colo.: Libraries Unlimited, 2001): 40; *Library of Congress Subject Headings*.
16. NISO, *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*, 16; *Library of Congress Subject Headings*, 4; Birger Hjørland, "Literary Warrant (and Other Kinds of Warrant)," (Aug. 20, 2008), [www.db.dk/bh/Lifeboat\\_KO/CONCEPTS/literary\\_warrant.htm](http://www.db.dk/bh/Lifeboat_KO/CONCEPTS/literary_warrant.htm) (accessed Feb. 25, 2009).
17. NISO, *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*, 16.
18. Carlyle, "Matching LCSH and User Vocabulary in the Library Catalog."
19. Ansari, "Matching Between Assigned Descriptors and Title Keywords in Medical Theses," 414; Gil-Leiva and Alonso-Arroyo, "Keywords Given by Authors of Scientific Articles in Database Descriptors," 1179.
20. Fredrick Sclafani, "Guest Essay: Controlled Subject Heading Searching Versus Keyword Searching," *Technicalities* 19, no. 9 (Oct. 1999): 15.
21. Gross and Taylor, "What Have We Got to Lose?" 223.