DOCUMENT RESUME

ED 275 326 IR 051 673

AUTHOR Mandel, Carol A.
TITLE Classification Schedules as Subject Enhancement in Online Catalogs. A Review of a Conference Sponsored by Forest Press, the OCLC Online Computer Library Center, and the Council on Library Resources.
INSTITUTION Council on Library Resources, Inc., Washington, D.C.
PUB DATE Oct 86
NOTE 28p.; For the final report of the Dewey Decimal Classification Online Project, see ED 266 805.
PUB TYPE Reports - Research/Technical (143) -- Collected Works - Conference Proceedings (021)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cataloging; *Classification; Indexing; Libraries; Machine Readable Cataloging; *Online Catalogs; *Online Searching; Relevance (Information Retrieval); Search Strategies; *Subject Index Terms
IDENTIFIERS Authority Files; *Dewey Decimal Classification; *Library of Congress Classification

ABSTRACT

This paper presents a synthesis of the ideas and issues developed at a conference convened to review the results of the Dewey Decimal Classification Online Project and explore the potential for future use of the Dewey Decimal Classification (DDC) and Library of Congress Classification (LCC) schedules in online library catalogs. Conference discussion centered around the themes of subject search enhancements for the next generation of online catalogs, the future role of class number searching in online catalogs, and the feasibility of using machine-readable LCC and DDC schedules online. Six broad conclusions for the future are outlined: (1) all operational online catalogs should include the subject search features that have already proven necessary; (2) subject search strategies should be explored; (3) it is worthwhile to build the DDC into a classification authority file, available in machine-readable form for use as a cataloger's tool and in online catalogs; (4) using the Dewey Online Catalog (DOC) as a prototype, it is worthwhile to continue to refine the design of a DDC online catalog; (5) displays of related terms will be a valuable search enhancement in future online catalogs; and (6) LCC will eventually be made machine-readable. (KM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CLASSIFICATION SCHEDULES AS SUBJECT ENHANCEMENT IN ONLINE CATALOGS

A Review of a Conference
Sponsored by Forest Press,
the OCLC Online Computer Library Center,
and the Council on Library Resources

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jane A. Rosenberg

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CLASSIFICATION SCHEDULES AS SUBJECT ENHANCEMENT IN ONLINE CATALOGS

A Review of a Conference
Sponsored by Forest Press,
the OCLC Online Computer Library Center,
and the Council on Library Resources

by Carol A. Mandel

Council on Library Resources, Inc.
1785 Massachusetts Avenue, N.W.
Washington, D.C. 20036

October 1986

ERIC

The conference described in this report was funded as part of the Council on Library Resources' Bibliographic Service Development Program, a program administered by CLR and funded by several parties: the Carnegie Corporation, the Commonwealth Fund, the Ford Foundation, the William and Flora Hewlett Foundation, the Lilly Endowment, the Andrew W. Mellon Foundation, the Alfred P. Sloan Foundation, and the National Endowment for the Humanities.

INTRODUCTION

In January 1986, Forest Press, the Online Computer Library Center (OCLC), and the Council on Library Resources (CLR) brought together 30 librarians and information scientists working in the areas of classification, online subject access, and online library catalogs. The purpose of the invitational conference was twofold: (1) to provide an opportunity to review the results of a highly significant research undertaking, the Dewey Decimal Classification Online Project; and (2) to explore the potential for future use of the Dewey Decimal (DDC) and Library of Congress (LCC) classification schedules in online library catalogs. As background for the meeting, the participants had the more than 500-page final report of the DDC Online Project[1] and a brief description of a project using DDC numbers for online access in OHIONET member libraries. In addition, several presentations were made at the conference.

During the course of the lively one and one-half-day meeting, the participants' discussion ranged around three themes:

1. What subject search enhancements should we work to incorporate into the next generation of online catalogs?

2. What is the future role of class number searching in online catalogs?

3. How can we use machine-readable LCC and DDC schedules online?

The following paper is intended to present a synthesis of the ideas and issues developed at the conference. It does not provide minutes of the

discussions, nor does it attempt to summarize the wealth of valuable data and analyses published in the final report of the DDC Online Project. The latter is readily available in the project report's own excellent Executive Summary--a document that is required reading for anyone with a serious interest in using classification schedules online or in enhancing online catalogs. The objective of the following summary is to share the ideas of conference attendees with others in the profession and to stimulate others to continue to work toward exploiting the subject-rich content of classification schedules for use in online retrieval.

BACKGROUND ON THE DDC ONLINE PROJECT AND THE CONFERENCE

The DDC Online Project was a two-year investigation conducted by staff of the OCLC Office of Research under the direction of Research Scientist Karen Markey. With support from the Council on Library Resources and Forest Press (publisher of the DDC), the project studied the use of DDC class numbers and terms from the DDC Schedules and Relative Index in an experimental online catalog. The project sought the answers to three questions:

1. Is the machine-readable DDC (a product developed from print tapes) suitable for implementation in an online catalog as a searcher's tool for subject access, browsing, and display?

2. When used as a searcher's tool, does the online DDC improve the performance of subject searchers at an online catalog?

3. Do subject searchers prefer an online catalog in which the DDC is available?

To investigate these questions, the DDC Online team constructed two parallel online catalog systems--one without DDC enhancements (the Subject Online Catalog, or SOC) and one that included terms, class numbers, and hierarchical arrays of terms from the DDC (the Dewey Online Catalog, or DOC). Databases from four libraries were loaded into four separate SOC/DOC installations and tested by patrons and staff at each site. Between 8,000 and 12,000 MARC records in subject areas selected by the libraries were processed into the experimental online catalog:

1. Library of Congress--Economics (330-339), Commerce (380-382), and Management. Database size: 11,865 records.

2. New York State Library--N.Y. state geography (917.47-917.4799) and history (974.7-974.799); U. S. colonial history (973.1-973.2). Database size: 8,144 records.

3. Public Library of Columbus and Franklin County--Sports, recreation, performing arts (790-799). Database size: 9,719 records.

4. University of Illinois at Urbana-Champaign--Mathematics (510-519). Database size: 7,603 records.

The DOC enhancements included a number of resources previously unavailable in online catalogs: subject-rich terms from the DDC Relative Index and Schedules (captions and notes) included in keyword searches; DDC Relative Index entries available for alphabetical browsing; and DDC Schedule outlines available for

browsing, accessible by either keywords or class numbers. The project team thoroughly analyzed the results of SOC and DOC searches, using failure analyses to determine the value of each of the DDC enhancements in subject searching. They gathered additional evaluative information from analyses of precision and recall, from analyses of the unique subject content of the enhancements, and from interviews with SOC and DOC users. Project results were disseminated in draft form at the end of 1985. The final report was published in February 1986.

The project's research design and findings provide a wealth of new information on subject searching in online catalogs and on the online use of a DDC Schedule database. The project's sponsors realized that the final report did not mark the conclusion of an investigation as much as it provided both an exciting stimulus and a solid base for further work. To help assess and articulate the next phase of development in using classification schedules online, the Council on Library Resources organized an invitational meeting focused on the topic.

Cosponsored by Forest Press, OCLC, and CLR, the Conference on Classification Schedules as Subject Enhancement in Online Catalogs was held at OCLC headquarters in Dublin, Ohio, January 27-28, 1986. (A list of the 30 invited participants is included as Attachment A.) The conference agenda included both prepared presentations and group discussions. Formal presentations were made by: William Mischo on recent subject searching enhancements in online catalogs; Anh Demeyer on the construction of the Dewey Online Catalog; Karen Markey on the results of the DDC Online Project; and Lois Chan on the potential of LCC as an online retrieval tool. In addition, Peter Paulson offered Forest Press's perspective on the future of DDC online and Carol Mandel summarized the conference discussions.

In the following summary, the substance of the presentations is synthesized along with the discussion they sparked. The Chan paper, which provides a valuable, well-organized analysis of LCC characteristics in relation to online retrieval, is available as a separate publication.[2] Details of the Demeyer and Markey presentations can be found in the DDC Online Project final report. The present summary is organized around the three themes that emerged during conference discussions: (1) subject searching enhancements for the next generation of online catalogs; (2) class number searches in online catalogs; and (3) classification schedules as online tools.

SUBJECT SEARCHING IN THE NEXT GENERATION OF ONLINE CATALOGS

Requirements for the current generation

The methodology of the DDC Online investigation was carefully designed to assess the contributions to successful subject searches made by specific features of SOC and DOC. In addition to evaluating DDC enhancements, the in-depth analyses of subject searches confirmed the value of significant features that should be (but are not yet universally) incorporated into currently operational online catalogs. These features are enumerated below.

1. Catalogs designed for end-user searching must compensate for users' documented inability to match exactly the phrases that form Library of Congress Subject Headings (LCSH) subject terms, even when the user is "on the right track." Capitalization and punctuation must be normalized. Some spelling errors can be handled through implicit truncation when no match occurs (e.g., dropping plural endings). Another hedge against spelling discrepancies is an alphabetical

display of subject headings that file (or would file if there is no match) next to the term input by the searcher. Finally, keyword searching of subject headings and titles is essential to effective subject retrieval.

2. Online catalogs should provide more than one subject searching option, including keyword and exact subject searches, call number searches, and browsing of subject indexes. The options are complementary, not redundant. Each option will contribute additional unique retrievals to the same information quest. Also, whether by temperament or need, individual search styles in any given setting will range from quick to methodical. The online catalog must satisfy a variety of approaches.

3. Displays of retrieved records should be brief, easy to interpret, and easy to manipulate (e.g., backward and forward scrolling). No matter how well-framed a search strategy may be or how accurate its results, the effective end result of a search is the subset of records that the user chooses to display. When search results are small (e.g., 1-5 hits), 95 percent of users will display all of the retrieved records. When retrieval results are larger, users become discouraged at viewing search results unless the hits are easy to browse and evaluate.

4. Online catalogs should permit users to refine large retrieval sets. Techniques used successfully in many online catalogs include: (1)

the ability to add further Boolean arguments to the search argument and (2) the ability to limit results by characteristics such as date, language, and location. (More sophisticated techniques are discussed in the section on navigated searches.)

5. Even when a subject search argument matches an indexed term, online catalogs should permit the searcher to choose whether to go directly to displays of search results or to view a list of subject terms that includes the matched term. The former option must be provided for the many searchers who are in a hurry, while the latter is necessary because users may not be aware that a more specific term (e.g., the subject plus a subdivision) exists and would better meet their needs.

The next generation: related topic displays

In existing online catalogs, the only available term displays are alphabetical, either by whole term or by keyword. Yet when users browse lists of subjects, they are seeking terms substantively related to their search argument. Even sophisticated searchers (including conference attendees) instinctively ponder "dumb" alphabetical listings produced by The Computer, seeking meaningful relationships. But such relationships have not yet been built into online catalogs.

Identifying and articulating term relationships is an intellectual and editorial task; such relationships are recorded in thesauri and the captions of classification schedules. In online catalogs, related term displays could be devised from appropriately coded subject authority files or from another online

file (such as a classification schedule) integrated with the online catalog. The LCSH authority tapes provide a database from which to derive displays of terms that are broader, narrower, or related to a user's search argument. However, as presently constructed, LCSH cannot be used to depict topical outlines or hierarchies. The LC class schedule, which is enumerative rather than hierarchical, also does not consistently supply an easily grasped overview of a subject area. The DDC is a potential source of logical and topical outlines for online browsing.

The DDC Online Project provided an opportunity to supply users with displays of topically related terms and to study users' reactions to this aid. On the surface, the results are disappointing, since subject outline searches led to the retrieval of relevant items in only 29 percent of the cases. However, further examination reveals that DOC's subject outline search, while a worthy first attempt, contained a few significant design flaws. The most notable was one that led users to broader topics in the subject outline by matching terms to class numbers and directing users to the most general (or shortest) class numbers of those matched. A redesigned DOC would direct users to the class number and caption with the highest number of postings matched during the grouping procedure. The DDC outlines themselves also proved to be imperfect tools, sometimes lacking useful terminology or sufficient levels of hierarchy.

Conferees concluded that attempts to provide users with a display of related topics should not be judged solely on their ability to deliver relevant hits. The subject outline is not a tool for the quick and dirty searcher, but for the methodical user; it should be evaluated in terms of users' satisfaction with its

contribution to their development of search strategies. Also, it may be a tool that requires user training. If it can add significantly to the success of subject searching for more methodical, well-trained searchers, it will be a powerful enhancement for online catalogs. The DOC subject outline search was a valuable first attempt to present a display of related topics to online catalog users. It will be important to build on this research, formulating retrieval algorithms, displays, and specialized browsing capabilities that exploit the concept/term relationships that can be derived from LCSH and DDC online.

The next generation: navigated searches

The careful failure analyses in the DDC Online investigation highlighted users' lack of expertise in formulating search strategies. Users often initiate a subject search by simply guessing about an approach to their topic. In many cases (about one-third of the time) their search argument bears little resemblance to their own expressed information need; often they search using a term that is broader than the topic they seek. A typical example was the user who searched the term "Olympics" in order to find material on Russian sports. Even if a user enters an appropriately descriptive term, the next hurdle is matching this term to the indexing language of the catalog. Expert searchers (usually reference librarians) have a bag of tricks for formulating search arguments. Can some of these tricks be built into online catalogs?

Navigation aids include instructions and displays that help users refine searches (tools that bring out the expertise of the user), as well as transparent operations that try successive search strategies in response to a single command. (One conference participant noted that users respond to transparent assistance in

two ways: half the users are pleased with the results; the other half spend three hours attempting to understand the source of the results.) Conferees described some existing or experimental examples of navigated searches. The online catalog at the University of Illinois can use a match on title keywords to direct users to the first subject heading that appears on the matching title. Or it can take a user's keyword search, identify the subject headings most common in the retrieval set, and display these to the user. An experimental online catalog, OKAPI, uses the following transparent strategy in response to a user's subject search: attempt an exact phrase match; if no hit, attempt a keyword search; if no hit, search keywords separately and retrieve the two matches with the fewest postings. Some of these strategies can be built into microcomputer "front-ends" to online catalogs in the form of chained commands, assuming the indexes have been designed to support them. This approach is likely to increase in use as microcomputers are employed as gateways directing users to a selection of online databases, including the online catalog. There is a clear research agenda as various navigated search strategies must be tested, evaluated, and promulgated for use in online catalogs.

CLASS NUMBER SEARCHING IN ONLINE CATALOGS

In the libraries of antiquity or in contemporary computer systems, classification schemes have been fundamental to information retrieval. Classification schemes are not merely methods for self-arrangement. The classified catalog is the core tool of many European libraries, and even North American subject catalogs are alphabetico-classed--i.e., subjects are broken down by subtopics within alphabetical lists. Classification schemes provide a different approach from subject headings, pulling together broader concepts to which dozens or even hundreds of separate specific topical headings might apply.

Such broad approaches are desirable, e.g., for browsing, for readers who have not yet refined their topics, for SDI services, or for selecting subsets of records for downloading into another system. A classification number may sometimes provide the most precise access point, as in cases where only a specific aspect of a subject is wanted (e.g., frogs from a culinary aspect). Both the Library of Congress Classification and the Dewey Decimal Classification offer useful, but different approaches. The hierarchical structure of the DDC is particularly powerful for browsing and refining concepts. The enumerative structure of the LCC tends to scatter its specific numbers, but some truncation of LCC numbers is also useful. And in a few cases, the LCC collocates subjects in a manner not available in Dewey--for example, law by jurisdiction. The LCC and DDC shelflists have long been essential tools of the library staff. The online catalog provides the opportunity to make call numbers and other class numbers available to library patrons as well.

The online catalog obviates the need to choose among alphabetical, classed, or some form of alphabetico-classed catalogs; all can exist within the same tool when the records and indexes are structured thoughtfully. Cochrane and Markey[3] reviewed 42 online catalogs that were operating in 1985 and determined that almost three-quarters of them provided access through class or call number. However, many of these call number searches were designed as a follow-on to a bibliographic retrieval in order to obtain circulation status or other item-level information. Even without the availability of an online classification schedule, class number searching could be more fully exploited as an access point, in search strategies, and in an online shelflist.

A number of libraries are using call numbers as rudimentary subject access points in circulation systems that are serving as pseudo-online catalogs. Other libraries, including some using OHIONET's TLM system, use call numbers to access uncataloged or briefly cataloged materials. The class number access point can be taken much further with fully cataloged materials. Viewed as potential subject access points, class numbers need not be restricted to one per record. In fact, class numbers can be assigned and indexed independently from call numbers, freeing changes in classification from the labor of call number and label changes. MARC records distributed by the Library of Congress often contain both LCC and DCC numbers, and may even contain a second, alternate LCC number as well. The advantages of DDC numbers for hierarchical browsing indicate their potential for use even in libraries where LCC provides shelf arrangement. While few libraries are clamoring for the added workloads of assigning multiple class numbers and training patrons to use them, the area is ripe for pilot projects involving specialized collections and applications.

Even when it is not a searcher's first choice access point, the class number is a key element in formulating search strategies. Many library users employ the catalog to find a shelf location where they can go to browse. Experienced searchers use class numbers on relevant records to produce other records that may contain further search terms. A well-designed class number browse can help in refining a search. A class-ordered display of retrieved records could aid in selecting relevant items. Most retrieval displays are unsorted, or in alphabetical or chronological order. Yet the results of a subject search may be more meaningful when shown in classified order. A class number used in a Boolean search argument could add considerably to the precision of a search--for example, by restricting the

search term "gold" to records containing the classification number for jewelry-making. It has been argued that the best way to pose a precise search in large databases is to seek those records that contain the combination of the most precise DDC number, the precise LCC number, and the most specific LCSH term. The subject search strategies navigated by the next generation of online catalogs should certainly take advantage of class numbers contained in bibliographic records.

The online catalog also needs to improve its functionality as an online shelflist. Most online catalogs cannot display items in true shelf order, with records sorted by shelf location (e.g., locations such as "Reference" or "Bibliography") and book numbers correctly sorted within class areas. Thus most online catalogs cannot support essential library staff uses of the old 3x5 shelflist, such as book number assignment or inventory control. Good shelflist displays (i.e., properly sorted; easy to browse forward and backward; call number and author and title information displayed together) are equally important to library users when the shelf itself is not accessible. The online catalog can supply the only "spine" browse for microforms, items in storage, or materials at remote sites.

Despite the potential power as an access point, class number searches currently constitute only about 10 percent of all online catalog searches. Cochrane and Markey[3] argue that class number searching is underutilized because users need tools to aid them in identifying and selecting useful class numbers. In the DDC Online Project, class number searches on DOC led users to an outline that helped them interpret the meaning of numbers--a feature praised in the evaluative interviews. Even more important, users need to be led to desired class/call

numbers through words; the link between a user's natural language and a shelf location (or an online shelflist browse) is a critical function of an online catalog. In existing online catalogs, this link is built through the keywords and class numbers contained in the same bibliographic records, and is therefore limited. Work in developing DOC revealed that only the first subject heading in a given record serves as a useful link to the class number on that record. Subsequent subject headings are most frequently used to bring out other aspects of the work than that represented by the class number. Additional online tools could make the link from terms to class numbers far more powerful. One such potential tool is the LCSH authority file; subject authority records contain separate fields for appropriate LCC and DDC numbers. (Currently, the LCC number field is not consistently used or maintained, and the DDC number field is not used at all.) Navigated searches could take advantage of this link by using matches to subject authority records to help suggest class numbers for searching. The richest potential tools for class number search enhancements are, of course, the DDC and LCC schedules themselves.

CLASSIFICATION SCHEDULES AS ONLINE TOOLS

What do we mean when we envision a classification schedule in machine-readable form? Broadly speaking, the concept ranges from a print tape coded for photocomposition to a database of classification authority records, each carefully coded to designate individual elements (e.g., notes and captions), structure, and relationships to other classification records. (For a thoughtful proposed specification for a MARC classification record format, see Cochrane and Markey.[3]) Machine files are not subject to the space constraints of printed tools and can contain considerably more coding and articulation of information than their printed counterparts. While there is added work involved in encoding the data, the work

would be done by a central agency for use by many libraries. The class schedule database used in the DDC Online Project is fairly well developed, with a separate record for each number and captioned range used in the project. Each record contains specific content designations for the elements relating to the class number, such as captions, class elsewhere notes, and examples notes. Similar records were also used for DDC Relative Index terms leading to class numbers used in the project. No such product yet exists for LCC. The Gaylord Company has keyed a word-processed file of the LCC Schedule L, including index, which is updatable and suitable for viewing on optical disk. (The company is examining the results of the project and may choose to follow it up with further work.) This LC text database contains no "coding" other than indentations and returns. Experience gained from keying the file indicates that machine manipulation of the file would, at best, produce only a first, very rough cut at content designation. An arduous task of record-by-record interpretation and editing would be required before the LCC Schedule could become a machine-readable authority file.

The potential use of online classification schedules will depend, in part, on how they are encoded. The machine-readable databases are developing from the production, editing, and distribution process of the schedules. Thus the initial use of the computerized schedules is as an editor's and classifier's tool. Forest Press has used the DDC file to create an online access and editing system for producing the 20th edition. The Press plans to continue work to restructure the print tapes into a logical database and to edit the Relative Index for online searching. This includes eliminating ambiguous abbreviations, editing for consistency in indentation, and adding index terms for the many class numbers that lack them. The Press hopes to be able to make this database available as a

cataloger's support tool. The potential optical disk LCC is also intended as a cataloger's tool, although it would function as an electronic reference book rather than as an authority file.

The potential use of a classification schedule in online catalogs falls into three areas: (1) as an authority file; (2) as a source of additional searching vocabulary; and (3) as a tool built into the search strategy options offered by the catalog. Looking first at an authority file, one can see such a file already developing from the Forest Press DDC tapes. The tape consists of class number records that include captions, notes, references to related numbers, and references linking new and changed numbers with those previously used. While the initial use of this tape will be as a resource file, it is also possible to envision its future use as an authority file within an online catalog, linked to class numbers in bibliographic records and used for file maintenance and as a source of links to other terms, references, and explanatory information. A regularly updated Forest Press product could be used to keep class numbers in catalogs current. However, unless libraries begin to use class number access points that are separate from call numbers, it is unlikely that they will want to see class numbers automatically "flipped" in the same manner as updated name headings. One current problem with the DDC file as an authority is the lack of records for synthesized numbers that are not explicit in the schedule; Forest Press is currently exploring the possibilities for assembling and parsing these numbers by machine. Should the LCC schedules be developed into authority files, numbers currently built from auxiliary tables would present a similar problem.

While there is little (if any) current demand for online catalogs with classification authority files, there is considerable demand for augmented subject vocabulary in online catalogs, an enhancement that would serve both quick and methodical users. The DDC Online Project demonstrated that terms from the DDC captions and Relative Index were valuable in keyword searching. In fact, the DDC provided more unique terms to DOC than LCSH, although not all of these unique terms were good indicators of a book's content. (About 25 percent of DDC schedule terms were unsatisfactory as subject terms, lacking in subject content or containing awkward wording.) Chan's assessment of LCC indicates that it too would add considerable new vocabulary beyond LCSH as well as thousands of proper names (not necessarily in AACR2 form!). The use of keywords from class schedules is particularly appealing in systems that include records containing class numbers but lacking controlled vocabulary terms, such as catalogs derived from circulation systems or large databases of minimal level records. However, given what we now know about subject searching and the need for multiple approaches, Relative Index terms cannot be viewed as a substitute for subject indexing of most library materials. Also, terms derived from call number links will cover only one aspect of the subject content of a book. The DOC results do confirm that added vocabulary from classification schedules can help to match users' search terms to records in the catalog.

The DDC Online Project has brought the DDC close to readiness for use in vocabulary augmentation. It identified editorial work that needs to be done to make the language of the captions and Index more useful, and the DDC editors have embarked on this work. DOC also tested strategies for moving from DDC keywords to bibliographic records, and the research team has identified recommended

refinements. This is an area that merits further research, both testing and strategies for matching terms with numbers and, for the benefit of the many LCC libraries, investigating the use of DDC terms with the DDC class numbers in LC MARC records rather than with specific call numbers. A catalog effectively using DDC keyword searching could be ready by the time the 20th edition is available on tape.

The use of LCC for vocabulary augmentation is less imminent. The LCC indexes are unintegrated and cumbersome; even if they were made machine-readable they could not be effectively indexed and stored in online catalogs in their current unedited form. Hierarchical relationships between numbers and captions are evident in indentions in the schedules, and these indentions would have to be coded in a machine-readable format for LCC to maintain and express relationships between captions and LCC numbers. Caption terminology would also require editing. Both LCC and DDC terms would be added to keyword indexes as uncontrolled vocabulary; such vocabulary brings the risk of too-large retrievals and false drops. While DOC demonstrated that this problem is manageable for DDC terms, similar careful research would be required to monitor the impact of added LCC terms on retrieval results. This is a field wide open for future research.

Ultimately, class schedules can provide future online catalogs with additional tools for navigating searches. The schedules can be viewed as arrangements of concepts that are linked to numbers that locate books. Relatively few end users will initially enter class numbers into the keyboard; instead systems will use searchers' terms to match class numbers and use these numbers "behind the screen" to provide search options such as browsing subject outlines, refining subjects by aspect, browsing a shelf arrangement, and viewing related terms. The

existence of the DDC in many translations even makes it possible to imagine the DDC numbers as a switching language in a multilingual system using terms from translations of the captions and Index. As we look toward building expert systems into online catalog searches, classification schedules appear to be a promising source of "expertise."

CONCLUSION

Conference participants agreed that the DDC Online Project will stimulate considerable future research and development. In general, the next broad steps seem to be taking shape as follows:

1. All operational online catalogs should include the subject search features that have already proven necessary: e.g., normalization, component word searching, browse displays, a mix of subject searching options, search limitation, and the other features enumerated at the beginning of this paper.

2. Subject searching requires multiple, complementary approaches. The next stage of online catalog development is the exploration of subject search strategies: built-in sequences of arguments and algorithms that overcome "no hits," refine too-large retrievals, and aid in the selection of search terms.

3. It is worthwhile to continue to build the DDC into a classification authority file, available in machine-readable form for use as a catalogers' tool and in online catalogs.

4. Using DOC as a prototype, it is worthwhile to continue to refine the design of a DOC online catalog, learning more about the use of augmented vocabulary and assisted search strategies.

5. Displays of related terms will be a valuable search enhancement in future online catalogs. It is worthwhile to continue to explore the use of displays of related terms in subject searching, improving on DOC's subject outlines and developing useful related term displays from the LCSH tapes.

6. LCC will eventually be made machine-readable, at least for ease of editing, updating, and electronic publication. As changes in the LCC editorial process are considered, they should be viewed in light of an evolution (however protracted) toward an LCC authority file.

REFERENCES

The three works referred to in this paper are all highly recommended reading for those interested in classification in online catalogs.

1. Markey, Karen, and Anh Demeyer. Dewey Decimal Classification Online Project: Evaluation of a Library Schedule and Index Integrated into the Subject Searching Capabilities of an Online Catalog. Dublin, Ohio: OCLC, 1986.

2. Chan, Lois Mai. "Library of Congress Classification as an Online Retrieval Tool: Potentials and Limitations." To be published in Information Technology and Libraries, Fall 1986.

3. Cochrane, Pauline A., and Karen Markey. "Preparing for the Use of Classification in Online Cataloging Systems and Online Catalogs." Information Technology and Libraries 4, no. 2 (1985): 91-111.

Attachment A
CLASSIFICATION SCHEDULES AS SUBJECT ENHANCEMENT IN ONLINE CATALOGS

A Conference Sponsored by Forest Press,
OCLC, and the Council on Library Resources

OCLC Headquarters
Dublin, Ohio
January 27-28, 1986

List of Participants

Julianne Beall
Decimal Classification Division
Room 507, Madison Building
Library of Congress
Washington, D.C. 20540
202/287-6986

Marcia Bates
Univ. of California, Los Angeles
Grad. School of Lib. & Info. Science
405 Hilgard Avenue
Los Angeles, California 90024
213/825-4352

Lois M. Chan
University of Kentucky
College of Library & Info. Science
Lexington, Kentucky 40506
606/257-5942

John Comaromi
Decimal Classification Division
Room 507, Madison Building
Library of Congress
Washington, D.C. 20540
202/287-6991

Walter Curley
Forest Press
8 Bass River Parkway
South Yarmouth, Massachusetts 02664
617/394-5038

Barbara C. Dean
Council on Library Resources
1785 Massachusetts Avenue, N.W.
Washington, D.C. 20036
202/483-7474

Anh Demeyer
OCLC
c/o 11 5060 Sheri Street
Naperville, Illinois 60565
312/983-5178

Ron Diener
OHIONET, Inc
1500 West L. [illegible]venue
Columbus, OH. [illegible]221
614/486-2966

Jeff Gerckens
Office of Research
OCLC
6565 Frantz Road
Dublin, Ohio 43017
614/764-6085

Sarah E. How
Research Libraries Group, Inc.
Jordan Quadrangle
Stanford, California 94305
415/329-3583

Charles Husbands
Office for Systems Planning and
Research - University Library
Harvard University
Cambridge, Massachusetts 02138
617/495-3724

C. Lee Jones
Council on Library Resources
P.O. Box 248
Buchanan Dam, Texas 78609
512/793-6118

Judith Kramer-Greene
Forest Press
85 Watervliet Avenue
Albany, New York 12206
518/489-8549

Gary Lawrence
Office of the President
Library Plans and Policies
Library Studies & Research Division
Room 7, University Hall
Berkeley, California 94720
415/642-2370

Carol Mandel
Central University Library
CO75-G
University of California, San Diego
La Jolla, California 92037
619/452-3061

Karen Markey
OCLC
6565 Frantz Road
Dublin, Ohio 43017
614/764-6082

Penny Mattern
Oxford Project
OCLC
6565 Frantz Road
Dublin, Ohio 43017
614/764-6044

William Mischo
Engineering Library
University of Illinois
Urbana, Illinois 61801
217/333-3576

Pat Molholt
Art & Architecture Thesaurus
Folsom Library
Rensselaer Polytechnic Institute
Troy, New York 12181
518/266 8300

Brian Nielsen
Reference Department
University Library
Northwestern University
Evanston, Illinois 60201
312/491-7658

James M. O'Brien
Suburban Library System
125 Tower Drive
Burr Ridge, Illinois 60521
312/325-6640

Peter J. Paulson
Forest Press
85 Watervliet Avenue
Albany, New York 12206
518/489-8549

Mary K. Pietris
Subject Cataloging Division
Library of Congress
Washington, D.C. 20540
202/287-5342

Gail Persky
Elmer Holmes Bobst Library
70 Washington Square South
New York University
New York, New York 10012-1091
212/598-2498

Phyllis Bova Spies
Local Systems
OCLC
6565 Frantz Road
Dublin, Ohio 43017
614/764-6266

Elaine Svenonius
Univ. of California, Los Angeles
Graduate School of Library and
Information Science
405 Hilgard Avenue
Los Angeles, California 90024
213/825-4352

Arnold Wajenberg
University of Illinois
214 Library
1408 West Gregory Drive
Urbana, Illinois 61801
217/333-6411

Hans Wellisch
College of Library and
Information Services
Room 4105
Undergraduate Library Building
University of Maryland
College Park, Maryland 20742
301/454-3785

Nancy Williamson
School of Library and Information
Science
University of Toronto
Claude Bissell Building
140 St. George Street
Toronto, Ontario
Canada M5S 1A1
416/978-7079

Jennifer Younger
Memorial Library
University of Wisconsin
Madison, Wisconsin 53706
608/262-4907