

---

# Preface

Subject analysis in online library catalogs is limited to the assignment of one or two subject headings per book. Most headings are actually an assemblage of several separate components. The sequence of components is not fixed because the meaning of subject headings is often dependent on the sequence of their individual components.

When searching online catalogs, system users are faced with a bewildering task. The cataloging records they hope to retrieve are assigned a limited number of subject headings. The sequence of subject heading components is difficult to predict. The subject searching approach featured in the online catalog has its own particular limitations and idiosyncrasies. The few online catalogs that offer more than one subject searching approach provide little, if any, guidance as to which approach is appropriate in view of the particular subject query entered. Furthermore, users express their subject queries in a variety of unpredictable forms, such as single words, phrases, questions, and even whole sentences.

The purpose of this book is to describe subject analysis and subject searching in online catalogs, to discuss their limitations on retrieval, and to demonstrate how such limitations can be overcome through system design and programming to enable systems to respond to user queries with the subject searching approach most likely to succeed.

The book begins with an introductory chapter on subject analysis. It describes the Library of Congress Subject Headings (LCSH) system and focuses on system characteristics such as vocabulary, structure, and syntax that impede retrieval.

Chapters 2 and 3 discuss the subject information stored in machine-readable files of authority records and bibliographic records. Chapter 4 investigates the relationships between records in these two files. The next five chapters address the subject queries users enter into online bibliographic systems. Chapter 5 surveys studies of user queries. Chapter 6 gives readers a broad picture of user queries and their closeness to authorized subject headings in the authority file. Chapters 7 through 10 focus on characteristics of user queries that are likely and those that are not likely to produce retrievals. For example, users are quite successful entering subject terms that match topical subject headings in the authority file; they are rarely successful retrieving bibliographic records when they enter queries that contain a topical element and a

personal name element. We make suggestions about system responses to user queries that are intended to increase precision, encourage browsing, and thus produce more useful retrievals.

Chapters 11 and 12 detail system responses and the characteristics of user queries best suited to specific responses. A new approach—the exact approach—is described in Chapter 11. It handles user queries that are exact or almost exact matches of catalog headings and references. This approach summarizes in a single screen the many options available for the matched heading, for example, availability of related terms, scope notes, and subdivided forms. The exact approach can be implemented in existing online catalogs, but certain enhancements are required in both bibliographic and authority files to make this approach work more effectively. Chapter 12 discusses several subject searching approaches now available in operational systems and the characteristics of user queries best suited to these approaches.

Chapter 13 focuses on search trees, which are defined as a set of paths with branches or choices that enable online systems to carry out the subject search approach most likely to succeed. Search trees control system responses and determine appropriate subject searching approaches to user queries. Users do not explicitly choose a particular approach. Rather, systems respond with an approach based on the extent to which queries match the catalog's controlled vocabulary and produce retrievals. The benefit of incorporating search trees into online bibliographic systems is that it places on the system the responsibility of determining which approach produces the best results.

The concluding chapter, Chapter 14, is a think piece on the future of subject headings for online retrieval. The authors speculate about the viability of the LCSH system for online retrieval, discuss planned changes in the system, and hypothesize about changes that would make the system more responsive to end-user needs and behavior.

Each chapter begins with a brief introductory section that gives readers a broad picture of the scope, issues, and subject coverage of the chapter and concludes with Synthesis and Summary sections that pull together the various points made in the chapter, interprets them for readers, and summarizes important details.

The analyses presented in this book span multiple years of effort; therefore, some LC subject headings or references used in examples may no longer represent current forms. We chose not to amend these terms since the premises put forth in the book remain valid despite ongoing changes to the LCSH system.

We hope that this book becomes an indispensable tool for online system designers who are grappling with developing new systems or refining existing ones. Practitioners in institutions using or considering the installation of an online catalog can consult this book to generate system specifications and to evaluate vendor proposals. The book will appeal to researchers in library systems, information retrieval, and user behavior because its recommendations are based on the results of empirical studies of library practice and end-user behavior. This book also addresses the interests both of advanced students in library and information studies schools and of

instructors in library automation, information retrieval, cataloging, indexing, and user behavior.

*Using Subject Headings for Online Retrieval* is truly unique. Although several books discuss subject searching capabilities and the problems of subject access, none specifically recommends which capability is best suited to the wide variety of subject queries users enter into online bibliographic systems.

This book demonstrates how the limitations of subject analysis and subject searching can be overcome through system design and programming to enable systems to respond to user queries with the subject searching approach most likely to succeed in providing relevant information. Despite our efforts to generate a new subject access design that does not require editorial review of existing subject access tools, we recognize that computers cannot solve all the problems of subject retrieval. Editorial changes that we suggest may be more difficult for the library and information science community to implement than automated changes because they require a concerted effort by the community to support the Library of Congress in making the revamping of LCSH a top priority.

As a new century approaches, the LCSH system will be called on to handle new forms of communication that are not represented on paper. If LCSH is not up to the task, it may be supplanted by an entirely new subject analysis tool. The data, analyses, and recommendations in this book are intended to motivate the library and information science community to think deeply about the future of LCSH and its potential as a subject access tool in online systems.