Guidelines for Indexes and Related Information Retrieval Devices

James D. Anderson

Abstract: This technical report provides guidelines for the content, organization, and presentation of indexes used for the retrieval of documents and parts of documents. It deals with the principles of indexing, regardless of the type of material indexed, the indexing method used (intellectual analysis, machine algorithm, or both), the medium of the index, or the method of presentation for searching. It emphasizes three processes essential for all indexes: comprehensive design, vocabulary management, and the provision of syntax. It includes definitions of indexes and of their parts, attributes, and aspects; a uniform vocabulary; treatment of the nature and variety of indexes; and recommendations regarding the design, organization, and presentation of indexes. It does not suggest guidelines for every detail or technique of indexing. These can be determined for each index on the basis of factors covered in the technical report.

A Technical Report Sponsored by
the National Information Standards Organization

NISO Press

Bethesda, Maryland, U.S.A.

About the NISO Technical Report Series

The National Information Standards Organization is accredited by the American National Standards Institute (ANSI) to develop and maintain technical standards for libraries, information services, and publishers. NISO technical reports address a variety of topics that are related to NISO's standardization interests. Conclusions or recommendations do not represent a consensus of the NISO membership and the material presented is not normative in nature. Questions or comments about the NISO Technical Report Series may be addressed to:

Contents

Foreword

This technical report was originally drafted as a proposed NISO standard on indexes to replace ANSI Z39.4-1984 Basic Criteria for Indexes. In the course of balloting the proposed standard, the NISO membership failed to reach sufficient agreement on the text, so it is published here as a technical report to serve as a current resource on indexing.

The document originated when NISO's Standards Development Committee charged NISO Standards Committee YY with revising the 1984 standard. The purpose of the revision was to broaden the scope of the standard, addressing all types of indexes (electronic and print, displayed and non-displayed, human and automatic); the 1984 standard had focused on print indexes created by human indexers.

The draft that resulted was balloted according to NISO procedures for developing standards. Following the ballot, the Committee responded to a wide range of suggestions, making adjustments to accommodate them, but the basic organization and focus of the draft remained unchanged. The draft was then balloted a second time. The second ballot resulted in two "No" votes which carried objections to the draft's broader level of detail and a number of recommendations.

At that point the draft was withdrawn from further deliberation because the added recommendations would have completely altered the character of the proposed standard. Indeed, if the new effort to include all types of indexes and indexing were dropped, there would be no compelling need to change the 1984 standard. Subsequently, the decision to issue the draft as a technical report was made. The 1984 standard is now withdrawn.

The text that appears here is the draft that was balloted a second time with corrections and changes suggested by representatives from the American Library Association, the American Society for Information Science, the American Society of Indexers, the American Theological Library Association, and the Association of Jewish Libraries.

The members of NISO Standards Committee YY who prepared the draft standard on which this report is based were James D. Anderson, Rutgers University (Chair); Barbara Anderson, Knight-Ridder Information, Inc.; Catherine Grissom, U.S. Department of Energy; Barbara Preschel, Public Affairs Information Service; Deborah Swain, IBM and Society for Technical Communication; and Hans Wellisch, University of Maryland.

This technical report is the second in the NISO Technical Report Series. It is not a national standard, and its material is not normative in nature. Comments may be addressed to the National Information Standards Organization, 4733 Bethesda Avenue, Suite 300, Bethesda, MD 20814.

Patricia Harris
Executive Director
National Information Standards Organization

Preface

This technical report is intended for everyone concerned with indexes used for information retrieval — professional indexers working with every kind of document, database producers, publishers of indexes and of documents containing indexes, designers of electronic index displays and indexing algorithms, librarians and catalogers, students, and other users of indexes.

It is particularly designed for persons who have substantial knowledge and experience related to indexes. Others will want to use it to help set goals and determine criteria for indexes, while using the more detailed guides and textbooks listed in the bibliography for background and instruction.

Because indexes range from simple lists to very complex tools for locating information, and because types of indexes and methods of indexing are quite varied, the report reflects this complexity and variation by dealing with the principles of indexing. In doing so, it emphasizes three processes essential for all indexes: comprehensive design, vocabulary management, and the provision of syntax.

The Committee that prepared this technical report has tried to address complex issues in the most straightforward manner possible, but we recognize that in addressing such a complex process as indexing, simple language cannot always suffice. Our task was complicated by our charge to address in a single document all types of indexes used for information retrieval. The technical terminology associated with various types of indexes is not always compatible. We attempted to bring some uniformity to the terminology of indexing, spanning the several disciplines and professions most directly concerned.

The world of information retrieval indexes is changing rapidly. Publishers who in the past produced only print-on-paper books are now issuing books on electronic disks, replacing the traditional back-of-the-book index with an electronic index. Other non-traditional indexes are in use for other media. This report speaks to the fast-changing context of indexes and indexing by identifying generic criteria that apply to all types of indexes, such as criteria related to vocabulary management and syntax, as well as recommendations that apply only to particular types of indexes.

To help persons interested only in particular types of indexes, the report includes four guides, one each for:

- print indexes for single documents (including back-of-the-book indexes)
- database and other continuing indexes
- automatic/algorithmic indexing
- indexes designed for electronic searching (non-displayed indexes)

These guides point to the most important recommendations for these categories of indexes.

The committee that developed this report wishes to thank Jessica Milstead, liaison from the NISO Standards Development Committee and chair of the NISO committee that created the 1984 standard on indexes, for her constant support and assistance, and also Nancy Mulvany of the American Society of Indexers for her valuable contributions as a member of the committee in the early years of the project. We also thank the hundreds of members of information and library associations who contributed feedback and suggestions throughout the work of the committee.

Finally, I would like to thank Patricia Harris, NISO Executive Director, for her support and assistance to the committee throughout the six years that we worked on this project, and especially for her faith in our work when obstacles appeared to be great. I also want to thank our editor Anita DeVivo and typesetter Sue Waterman for their collegial and conscientious assistance.

James D. Anderson, Committee Chair
Rutgers University

Members of the NISO Standards Committee YY:
Barbara Anderson, Knight-Ridder Information, Inc.
Catherine Grissom, U.S. Department of Energy
Barbara Preschel, Public Affairs Information Service
Deborah Swain, IBM and Society for Technical Communication
Hans Wellisch, University of Maryland

0. Introduction

0.1 Summary of major sections
This technical report consists of 9 sections. They are briefly summarized here:

1. Scope of the report describes aspects of index preparation and presentation addressed in the report. This section encompasses principles, rather than detailed procedures, for the presentation of print and electronic indexes compiled by human analysis and by computer algorithm for the retrieval of all types of documents. It includes both displayed indexes, designed for searching by means of human visual inspection, and non-displayed indexes, designed for searching by means of electronic comparison and matching.

2. Definitions lists only major and essential terms with their definitions. These terms and many additional terms are defined in a glossary appended to the report.

3. Function of an index gives an expanded definition of "index" in the context of information retrieval in terms of the minimum functions an index ought to perform.

4. Types of index continues and expands the definition of "index" in terms of the variety of types of index.

5. Design of indexes summarizes the design of indexes in terms of decision options for 18 key aspects or attributes of indexes. For the most part, these guidelines do not favor particular choices or options, but instead urge that decisions on options (a) be based primarily on needs, habits, and preferences of users; (b) that publishers and producers of indexes agree on attribute options prior to the production of an index; and (c) that all special or unusual options or attributes be made clear to index users.

6. Vocabulary recommends sources for and forms of terms used in indexes. These guidelines emphasize the importance of linking alternative terms and forms of terms for the same or similar concepts. They recommend linking terms for related concepts as well. In displayed indexes, the display of vocabulary information should be integrated into the display of the index. In non-displayed indexes, the search interface should provide for the display of vocabulary and term relationships at the time a search statement is created.

7. Headings, entries, and search statements describes a wide variety of syntactic methods and styles for the combination of terms to create index headings and entries in displayed indexes and search statements for non-displayed indexes. The principal recommendation states that such combination is absolutely essential, regardless of the type of index.

8. Display of index arrays lists options and recommendations for the display of indexes or parts of indexes, including arrays of retrieved entries or records from non-displayed indexes.

9. Alphanumeric arrangement contains rules for the arrangement of alphanumeric indexes.

Appended to the technical report are a comprehensive glossary of terms related to indexes and indexing-related references, a bibliography, and a detailed index to the technical report.

0.2 Guides to the technical report
NOTE: Because it is anticipated that these guides may be consulted separately, each guide covers the entire technical report. The result is some repetition among the individual guides.

A. Guide for print indexes to single documents (including back-of-the-book indexes)
Section 1 describes the overall scope of the technical report and section 2 provides definitions for the most important terms. Many other definitions may be found in the glossary at the end of the report.

Section 3 summarizes the functions of all indexes. For print indexes to single documents, item c (distinguishing between major and minor treatments) calls for some method of highlighting headings, subheadings, or locators for major (or, inversely for minor) topics or features (*see also* 7.4.4). Items d, e, g, and h (terminology and access) call for generous provision of cross-references or double entries to accommodate the terminology of the text, the anticipated terminology of users, and guidance to related terms or headings (*see also* section 6). Item f calls for the use of specific, rather than generic, terminology. Item i (combination of terms) calls for pre-coordinated, multi-term headings or main-heading/subheading combinations in order to reduce ambiguity. Item j calls for systematic arrangement of entries, which in most cases will be alphabetical (*see* sections 8 and 9).

Section 4 gives an overview of the variety of index types.

Section 5 guides the design of all types of indexes. For print indexes to single documents, sections 5.1 (subject scope), 5.2 (documentary scope), and 5.3 (domain) are almost always defined by the document to be indexed. A possible exception is an index that limits its subject scope by intentionally excluding certain categories of entries, such as names of individual persons or places. 5.4 discusses the merits of a single index versus separate indexes for special types of topics or features. 5.5 addresses the use of special symbols for special types of text, as in music, choreography, chemistry, or mathematics. Section 5.6 (display media) is not relevant for print indexes.

Section 5.7 suggests that whenever possible, index entries refer to inherent documentary units, such as paragraphs or sections, rather than page numbers. Section 5.8 (indexable matter) is relevant only if certain portions of the document being indexed are ignored or receive varying levels of attention. Most print indexes to single documents are based on human intellectual analysis (5.9).

Section 5.10 (exhaustivity) suggests that the level of detail in indexing — the approximate or average number of index terms per documentary unit (for example, page or paragraph) — be determined in advance. This is especially important when the space allocated for the index is limited. Similarly, 5.11 calls for decisions on the specificity of index terminology, and 5.12 (syntax) calls for a method of combining terms into multi-term headings or main-heading/subheading combinations. Section 5.13 (vocabulary management) calls for a method of cross-referencing or double entry to insure maximum access to particular topics or features.

Sections 5.14 and 5.15 (surrogates and locators) are more relevant for indexes that are separated from the documents they index. Section 5.16 lists the major factors that determine the size of a displayed index, and 5.17 stresses the importance of the arrangement of entries for access. Section 5.18 (search interface) relates only to electronic indexes.

All of section 6 (vocabulary) is directly relevant to printed indexes for single documents, because it discusses the sources and forms of terms and the display of cross-references.

Section 7 surveys a wide variety of syntactical methods for combining terms in both print and electronic indexes. Most print indexes to single documents will use "ad hoc syntax" (section 7.2.1), but indexers may want to consider other types of syntax discussed in section 7.2, such as subject heading syntax (7.2.3) or string indexing based on ad hoc coding (7.2.5.3). Syntactic cross-references (7.2.6) are also important for print indexes to single documents.

Section 7.4 is devoted to locators in displayed indexes, and most of this section is directly relevant to print indexes for single documents. Section 7.5 relates only to non-displayed electronic indexes.

Section 8 covers the display of index entries, and all of it relates to print indexes for single documents except for 8.3 (index display in electronic media).

If index entries are to be displayed in alphanumeric order, section 9 provides recommended rules for arrangement.

B. Guide for database and other continuing indexes

The chief distinguishing attribute of database and other continuing print or electronic indexes is the need to provide continuity in indexing practices and policies, in terminology, and in entry format and display. If a continuing index is displayed in a print format, the preceding Guide A for printed indexes to single documents is generally relevant, except for comments on section 7 (syntax and locators). If a continuing index is based on automatic or algorithmic indexing (rather than on human intellectual analysis), *see also* Guide C; if it is designed for electronic searching, consult Guide D.

Section 1 describes the overall scope of the technical report and section 2 provides definitions for the most important terms. Many other definitions may be found in the glossary at the end of the report.

Section 3 summarizes the functions of all indexes. Database and other continuing indexes should seek to fulfill all of these functions.

Section 4 gives an overview of the variety of index types.

Section 5 guides the design of all types of indexes. Sections 5.1 (subject scope), 5.2 (documentary scope), and 5.3 (domain) are especially important for database and other continuing indexes, because it is through the description of these attributes that users are apprised of the potential usefulness of the database or index. Users should be told, clearly and explicitly, what kinds of topics and features are indexed (what kinds of questions may be asked of the database or index), what categories of documents are indexed, and where these documents or their descriptions come from.

Section 5.4 discusses the merits of a single index versus separate indexes for special types of topics or features. Electronic databases and indexes may

take advantage of both approaches by permitting searches limited to particular fields for special categories of topics or features on the one hand and access to the complete "global" index on the other.

Section 5.5 addresses the use of special symbols for special types of text, as in music, choreography, chemistry, or mathematics.

Section 5.6 (display media) raises the question of the most appropriate medium, such as electronic or print, for the presentation of a database or index.

Section 5.7 (documentary units) asks for consideration of the most appropriate unit for indexing and subsequent retrieval. For example, should users be able to retrieve particular paragraphs or pages, or should retrieval be limited to complete documents, such as whole periodical articles or monographs?

Section 5.8 (indexable matter) asks for decisions regarding the inclusion or exclusion of particular types of documents or segments of documents, or any variation in the level of attention given to various types or segments of documents.

Section 5.9 raises the question of automatic/algorithmic identification of terms versus human intellectual analysis. This technical report applies to both categories of indexing. Guide C below applies specifically to automatic indexing.

Database and other ongoing indexes need to have explicit policies regarding the level of detail in indexing, that is, the average number of terms to be assigned or extracted per documentary unit. Section 5.10 (exhaustivity) discusses this need. Similarly, 5.11 calls for decisions or policies on the specificity of index terminology, which in turn will affect the size of the indexing vocabulary.

Section 5.12 (syntax) discusses the need for database or other continuing indexes to have a regular method of combining terms for retrieval. Print indexes will use some method for creating multi-term headings or main-heading/subheading combinations in advance of publication. If electronic databases are designed for electronic searching, rather than the display of arrays of entries for visual searching, some method for combining terms at the search stage must be provided.

Vocabulary management (5.13) is important for database and continuing indexes because there is a need for continuity of terminology over time.

Sections 5.14 and 5.15 (surrogates and locators) are relevant for database and other continuing indexes that do not include the full text of the documents being indexed. Section 5.16 lists the major factors that determine the size of a displayed index.

Section 5.17 stresses the importance of the display and arrangement of entries when visual searching is the mode of access, whether via print or electronically displayed indexes. Section 5.18 discusses the search interface for database and other indexes that are searched by means of electronic matching or comparison.

Section 6 discusses the vocabulary of indexes, which is especially important for database and other continuing indexes. If the database or index is based on human indexing and presented in a print format, section 6 will guide the selection of terms and the provision of needed cross-references. If the database or index makes use of automatic/algorithmic indexing, vocabulary management is just as important. The preferred terminology described in section 6 should be included among the access options provided, linked to terminology extracted from documents. In print indexes generated algorithmically (for example, KWIC, KWAC, and KWOC indexes), alternative terms (synonymous, equivalent, broader, narrower, related) may be presented by means of cross-references. In electronic indexes, vocabulary management may be provided by means of a thesaurus linked to the display of the index or to the search interface.

Section 7 surveys a wide variety of syntactical methods for combining terms in both print and electronic indexes. Sections 7.1 and 7.2 are devoted to pre-coordinate headings and syntactic methods appropriate for print or other displayed indexes. Section 7.5 is devoted to syntactic methods appropriate for indexes designed for electronic searches. Section 7.3 (weighted terms) discusses methods for differentiating between major and minor topics.

Section 7.4 is devoted to locators in displayed indexes, especially for print indexes.

Section 8 covers the display of index entries in print or electronic media. Section 8.2 focuses on print indexes, while 8.3 discusses electronic indexes.

If index entries are displayed in alphanumeric order, section 9 provides recommended rules for arrangement.

C. Guide for automatic/algorithmic indexing

Automatic indexing uses computer algorithms to select and extract terms from verbal text. Other algorithms may be used to create and format headings for display, to match terms against a thesaurus for linking with synonymous, equivalent, and related terms, to assign weights to terms, and to cluster related terms.

If an automatically generated index provides access to a single document, parts of Guide A

should be relevant. If it is a continuing index, Guide B should also be consulted. And if it is designed for electronic searching (rather than for visual display), Guide D may be helpful as well.

Section 1 describes the overall scope of the technical report and section 2 provides definitions for the most important terms. Many other definitions may be found in the glossary at the end of the report.

Section 3 summarizes the functions of all indexes. Indexes based on automatic indexing should seek to fulfill all of these functions.

Section 4 gives an overview of the variety of index types.

Section 5 guides the design of all types of indexes. See Guide A for a summary of the most important points for indexes to single documents, and Guide B for database or continuing indexes.

The medium through which the index is accessed or displayed (5.6) will determine, in large part, the type of method used for combining terms (5.12, syntax) and for providing access through alternative terms (5.13, vocabulary management).

Decisions regarding indexable matter (5.8) will directly influence the number and nature of index terms extracted. For example, the effectiveness of KWIC, KWOC, KWAC, and other indexes based on small segments of texts, such as titles, is directly related to how well these text segments represent the topics and features of a text. When abstracts or larger segments of text are used, up to and including the complete text, the number of terms extracted becomes very large, and the need for some method for indicating potentially more important terms increases (*see*, for example, 7.3 on weighted terms).

Exhaustivity (5.10), the number of terms extracted per unit of text, tends to be very high in automatic indexing, increasing the need for some method of indicating potentially more relevant terms and for methods of vocabulary management (5.13) for linking synonymous, equivalent, and related terms. Specificity (5.11) will depend on the nature of the terminology of the indexable matter, unless a thesaurus is used to exchange specific terms for more generic terms. In general, these guidelines prefer highly specific terminology.

The capability of combining terms is required for effective searching (5.12, syntax). If the index is displayed for visual searching, terms should be displayed in some type of pre-coordinated headings. For electronic searching, a method for combining terms at the time of search should be provided.

Section 6 discusses the vocabulary of indexes.

When an index is based on automatic indexing, the terminology extracted from texts is usually extremely diverse, making some form of vocabulary management essential. The preferred terminology described in section 6 should be included among the access options provided and should be linked to terminology extracted from documents. In print indexes generated algorithmically (for example, KWIC, KWAC, and KWOC indexes), alternative terms (synonymous, equivalent, broader, narrower, related) may be presented by means of cross-references. In electronic indexes, vocabulary management may be provided by means of a thesaurus linked to the display of the index or to the search interface.

Section 7 surveys a wide variety of syntactical methods for combining terms in both print and electronic indexes. Sections 7.1 and 7.2 are devoted to headings and syntactic methods appropriate for print and other displayed indexes. Section 7.2.2 deals specifically with natural language syntax, often used for the display of indexes based on automatic indexing. Section 7.5 is devoted to syntactic methods appropriate for indexes designed for electronic, rather than visual, searching.

Section 7.3 (weighted terms) discusses methods for differentiating between major and minor topics. Section 7.4 is devoted to locators in displayed indexes, especially for print indexes.

Section 8 covers the display of index entries in print or electronic media. Section 8.2 focuses on print indexes, while 8.3 discusses electronic indexes.

If index entries are displayed in alphanumeric order, section 9 provides recommended rules for arrangement.

D. Guide for indexes designed for electronic searching (non-displayed indexes)

Indexes available in electronic media may be searched by means of algorithmic matching and comparison, rather than through the visual inspection of displayed arrays of index headings.

If an index designed for electronic searching provides access to a single document, parts of Guide A should be relevant. If it is a continuing index, Guide B should also be consulted. If it is based on automatic indexing, Guide C should be helpful as well.

Section 1 describes the overall scope of the technical report and section 2 provides definitions for the most important terms. Many other definitions may be found in the glossary at the end of the report.

Section 3 summarizes the functions of all indexes.

Indexes designed for electronic searching should seek to fulfill all of these functions.

Section 4 gives an overview of the variety of index types.

Section 5 guides the design of all types of indexes. See Guide A for a summary of the most important points for indexes to single documents, Guide B for database or continuing indexes, and Guide C for indexes based on automatic indexing.

Indexes designed for electronic searching will be accessed via electronic media (5.6). This in turn will determine, in large part, the type of method used for combining terms (5.12, syntax) and for providing access through alternative terms (5.13, vocabulary management).

The capability of combining terms is required for effective searching (5.12, syntax). For electronic searching, a method for combining terms at the time of search should be provided by the search interface (5.18). The search interface should also provide access to alternative search terms.

Section 6 discusses the vocabulary of indexes. The preferred terminology described in section 6 should be included among the access options provided through the search interface.

Section 7 surveys a wide variety of syntactical methods for combining terms in both print and electronic indexes. Section 7.3 (weighted terms) discusses methods for differentiating between major and minor topics. Section 7.5 is devoted to syntactic methods appropriate for indexes designed for electronic searching.

Section 8 covers the display of index entries or records. Section 8.3.2 (displays of retrieved records) is directly relevant for indexes designed for electronic searching. If retrieved records are displayed in alphanumeric order, section 9 provides recommended rules for this arrangement.

Guidelines for Indexes and Related Information Retrieval Devices

1. Scope of the technical report

1.1 General statement

This technical report provides guidelines for the content, organization, and presentation of indexes used for the retrieval of documents and parts of documents. It deals with the principles of indexing, regardless of the type of material indexed, the indexing method used (intellectual analysis, machine algorithm, or both), the medium of the index, or the method of presentation for searching. It emphasizes three processes essential for all indexes: comprehensive design, vocabulary management, and the provision of syntax. It includes definitions of indexes and of their parts, attributes, and aspects; a uniform vocabulary; treatment of the nature and variety of indexes; and recommendations regarding the design, organization, and presentation of indexes. It does not suggest guidelines for every detail or technique of indexing. These can be determined for each index on the basis of factors covered in the technical report, including the type of material indexed, the medium of the index, the method of presentation for searching, and the type of user for whom the index is designed.

NOTE: In other contexts, the term "index" is used for other pointing or indicating phenomena or devices, for example, a consumer price index indicates the rise and fall of prices. The construction and display of indexes for purposes other than the retrieval of documents is not covered by this technical report.

1.2 Types of documents

This technical report deals with indexes for single documents and for collections of documents. "Document" is used in the broadest possible sense. (*See* "document" in section 2.)

1.3 Presentation of indexes

This technical report is concerned with basic indexing principles and practices as they affect the presentation of an index, whether the index is a displayed index designed for searching by means of visual inspection or a non-displayed index designed for searching by means of electronic comparison and matching.

The report emphasizes the presentation of an index to human users, rather than the way it is structured or stored electronically. It considers all kinds of indexes for human use, regardless of the medium on which the index is displayed or the method by which the index is presented for searching. The internal representation of computer-readable indexes (inverted files, for example) designed for electronic comparison and matching rather than human visual inspection is not directly addressed. Examples are illustrative, not prescriptive.

1.4 Choice of terms

This technical report covers criteria for the choice and form of terms to be used in headings in displayed indexes, as descriptors in non-displayed indexes, and in the vocabulary management component of indexes. These guidelines permit the use of natural language terms extracted from natural language text, but they call for the display of relationships among terms, whether natural language terms or controlled descriptors or headings, in order to indicate synonymous, equivalent, hierarchical, and associative relationships among concepts represented. In indexes using natural language terms, recommendations on the choice and form of terms should guide the selection of terms to be used as preferred terms around which to gather synonymous, equivalent, and related keywords. (For the compilation of thesauri that may be used to facilitate the display of terms and their relations, *see* ANSI/NISO Z39.19-1993, *Guidelines for the construction, format, and management of monolingual thesauri.*)

1.5 Method of preparation

This technical report is relevant to the preparation of all types of indexes for information retrieval, regardless of whether they are produced on the basis of human intellectual analysis or by automatic or computer-assisted methods, whether they are searched by visual inspection or by electronic algorithm, and whether they are compiled by one indexer or by teams of indexers.

This report does not address indexing software.

2. Definitions

Only the most important terms used in this technical report are listed and defined here. Other terms appear in the glossary at the end of the report.

Within these definitions, terms in italics have

their own definition, entered under the singular noun form, either in this section or in the expanded glossary at the end of the report.

cross-reference. A link between two or more *terms* or *headings* in an *index*. Three types of relationships among terms require cross-references: (a) an equivalence relationship among synonymous or equivalent terms or headings, (b) an associative relationship, indicating an unspecified relationship among terms or headings (called *related terms* or headings), and (c) a hierarchical relationship, indicating a broader/narrower relationship among terms or headings.

descriptor. A *term* chosen as the preferred representation for a *concept* or *feature* in an *index*.

displayed index. An *index* that may be searched by means of visual inspection.

document. A *medium* on or in which a *message* is encoded; thus, the combination of message and medium. The term applies not only to written and printed materials on paper or microforms (for example, books, journals, maps, diagrams), but also to nonprint media (for example, machine-readable records, transparencies, audio recordings, video recordings, film) and, by extension, to natural or humanly made objects intended to convey information. Documents encompass every kind of format and genre, including but not limited to treatises, literary works, patents, technical reports, charts, tables, illustrations, music, artistic works, and multimedia texts.

documentary unit. The *document*, document segment, or collection of documents to which *entries* in an index refer and on which they are based. Examples of verbal documentary units include sentences, paragraphs, pages, complete articles, books, complete serial runs, collections of archival materials, microform sets, and entire library collections. The documentary unit determines the relative size of document, document segment, or collection of documents to which an *index* will point.

entry. The representation of a *documentary unit* in a *displayed index*. It consists of at least a *heading* and a *locator*. More than one *locator* may follow a given heading in a displayed *entry array*, but each locator, in combination with its heading, represents a single entry. An entry may contain a *multi-level heading* and a document *surrogate* in addition to the required locator.

heading. One or more *terms* representing a *topic* or *feature* of a *document* in a *displayed index*; the first element of an index *entry* in a displayed index. A *multi-level heading* consists of a *main heading* followed by a *subheading*, and possibly by a *sub-subheading* and additional headings at successive levels of subordination.

index. A systematic guide designed to indicate *topics* or *features* of *documents* in order to facilitate retrieval of documents or parts of documents. Indexes include the following major components: (a) *terms* representing the topics or features of *documentary units*; (b) a *syntax* for combining terms into *headings* (in *displayed indexes*) or *search statements* (in *non-displayed indexes*) in order to represent compound or complex topics, features, and/or queries; (c) *cross-references* or other linking devices among *synonymous, equivalent, broader, narrower,* and other *related terms*; (d) a procedure for linking headings (in displayed indexes) or search statements (in non-displayed indexes) with particular documentary units or document *surrogates*; and (e) a systematic ordering of headings (in displayed indexes) or a search procedure (in non-displayed indexes).

indexing. The operation of creating an *index* for *information retrieval*. Indexing involves the selection and assignment of *terms* to, or the extraction of terms from, a *documentary unit* in order to indicate *topics, features,* or possible uses of the unit; the combining of terms into *headings* or the tagging of terms for subsequent combination (in *displayed indexes*); the linking of *synonymous, equivalent, broader, narrower,* and other *related terms* or headings; the linking of terms or headings to documentary units or *surrogates;* and the arrangement of headings in a systematic order (in displayed indexes).

locator. The part of an *entry* in a *displayed index* that indicates the location of the *documentary unit* to which the entry refers. Locators range from brief notations, such as page numbers, to full bibliographic citations.

non-displayed index. An *index* that is searched by means of electronic comparison and matching controlled by computer algorithms. The complete index itself is not displayed for searching by means of visual inspection.

term. A word or phrase used to represent a *topic* or *feature* of a *documentary unit* in an *index*.

3. Function of an index

The function of an index is to provide users with an effective and systematic means for locating documentary units (complete documents or parts of documents) that are relevant to information needs or requests. An index should therefore:

a. identify documentary units that treat particular topics or possess particular features.
b. indicate all important topics or features of documentary units in accordance with the level of exhaustivity appropriate for the index.
c. discriminate between major and minor treatments of particular topics or manifestations of particular features.
d. provide access to topics or features using the terminology of prospective users.
e. provide access to topics or features using the terminology of verbal texts being indexed whenever possible.
f. use terminology that is as specific as documentary units warrant and the indexing language permits.
g. provide access through synonymous and equivalent terms.
h. guide users to terms representing related concepts (narrower terms, other related terms, and if possible, broader terms).
i. provide for the combination of terms to facilitate the identification of particular types or aspects of topics or features and to eliminate unwanted types or aspects.
j. provide a means for searching for particular topics or features by means of a systematic arrangement of entries in displayed indexes or, for non-displayed indexes, by means of a clearly documented and displayed method for entering, combining, and modifying terms to create search statements and for reviewing retrieved items.

4. Types of index

Indexes may be categorized by type of object to which headings refer; by type of term used for index headings; by type or extent of indexable matter used to produce the index; by method of arranging entries; by method of term coordination; by type, format, genre, or medium of documents being indexed; by medium of the index; by mode of publication; by periodicity, that is, whether the index is a one-time (closed-end) index or a continuing (open-end) index; and by type of authorship. The following examples illustrate common types of indexes. They are by no means exhaustive.

4.1 Indexes by type of object referred to

a. authors: all types of document creators such as writers, composers, illustrators, translators, editors, choreographers, artists, sculptors, painters, inventors.
b. subjects (topics or features): topics treated in documents and/or features of documentary units (for example, genre, format, methodological approach). Separate indexes are often devoted to special types of topics such as persons, places, or corporate bodies; features, such as genres (for example, poetry, drama); or notations, such as International Standard Book Numbers (ISBN). *See also* 4.2.b, numbers or notations.

4.2 Indexes by type of term used for headings

a. names: proper nouns, such as names of persons, places, corporate bodies.
b. numbers or notations: numerical or coded designations, such as classification notation, patent number, ISBN, date.
c. words and phrases: common words and phrases (as opposed to names or proper nouns).

4.3 Indexes by type or extent of indexable matter on which an index is based

a. full text of documents.
b. abstracts.
c. titles only.
d. first lines only (for example, first lines of poems).
e. citations (reference citations to other documents).

4.4 Indexes by arrangement of entries

a. alphabetical or alphanumeric.
b. classified: Headings arranged on the basis of relations among concepts represented by headings, for example, hierarchy, inclusion, chronology, or other association. Classified indexes are often based on existing classification schemes, such as the Dewey Decimal Classification.
c. alphabetico-classed: Broad headings arranged alphabetically. Narrower headings are grouped under broad headings and arranged alphanumerically or relationally on the basis of hierarchy, inclusion, chronology, or other association.

NOTE: Electronic indexes often have no arrangement that is apparent to the user. However, indexes designed for human scanning, browsing, and examination must have some arrangement, regardless of medium.

4.5 Indexes by method of document analysis

a. human intellectual analysis and identification of

topics and concepts expressed and/or features manifested.

b. computer algorithms designed to identify useful terms, phrases, or features.
c. combination of computer-based and human analysis.

4.6 Indexes by method of term selection

a. assignment of terms to represent topics and features (whether or not the term is in the documentary unit being indexed).
b. extraction of terms from the documentary unit.
c. a combination of assignment and extraction methods.

4.7 Indexes by method of term coordination

a. pre-coordinate combination, such as subject heading indexes, string indexes, chain indexes, keyword indexes (including KWIC, KWOC, KWAC indexes), rotated, and permuted indexes.
b. post-coordinate combination. Includes the use of Boolean operators, proximity measures, and the combination of weighted terms.

4.8 Indexes by type, periodicity, format, genre, or medium of document(s) being indexed

Examples are: books, monographs, periodicals, serials, poetry, fiction, short stories, films, videos, illustrations, pictures, paintings, artifacts, software, computer-readable texts, maps, and sound recordings.

4.9 Indexes by medium of index

a. printed or written.
b. microform.
c. electronic media, including online, CD-ROM.
d. braille.

4.10 Indexes by proximity of documentary units

a. indexes published together with the documentary units to which they refer, including both back-of-the-book indexes and full-text databases.
b. indexes published separately from the documentary units to which they refer.

4.11 Indexes by periodicity of the index

a. one-time, closed-end indexes.
b. continuing, open-end indexes.

4.12 Indexes by authorship

a. authored: An authored index; a separately authored document distinct from the document(s) that is (are) being indexed. It is created

independently by one or more persons through intellectual analysis of text, as distinguished from indexes that are created solely through algorithmic analysis of text carried out electronically.

b. automatically generated.

5. Design of indexes

In the design of indexes, decisions should be made concerning key options and attributes. Careful consideration of all available options and attributes will contribute to a better index, because each option or attribute will influence overall quality and performance of the index. Decisions should be based primarily on needs, habits, and preferences of users. Publishers and producers of indexes should agree on options and attributes prior to the production of an index.

All special or unusual options or attributes should be made clear to users in an introductory statement in print indexes or in on-screen and off-screen documentation for electronic indexes (*see* 8.1, on introductory notes for indexes).

The following key options and attributes are present in most indexes.

5.1 Subject scope

The scope of a subject index to a single document should be the same as the subject scope of the document.

The scope of a subject index covering multiple documents may not necessarily be the same as the subject scope of the document(s) being indexed. This is especially true for "mission-oriented indexes," which focus on mission-related topics or features, ignoring topics and features not within their stated scope. The subject scope should be clearly described both in terms of the subject areas covered (for example: medicine, auto mechanics, food preparation) and in terms of the kinds of topics or features indexed within a subject area. For example, the index may provide access to topics or features such as:

concrete entities: persons (individuals or groups), institutions, artifacts, natural objects

abstract entities: belief systems, disciplines, theories, hypotheses, imaginary entities (for example fictional characters, mythological animals)

attributes and properties

"raw" materials; constituent elements, such as wood, plastic, iron

operations, processes, methodological approaches, events, conditions

places, environments

times and historical periods

5.2 Documentary scope

Indexes are also defined by categories of documents being indexed. The documentary scope or coverage of an index to a single document is obvious. For indexes to multiple documents, such as those provided by indexing and abstracting services or textual databases, it is important to state explicitly the kinds of documents included within the documentary scope of the index with respect to such criteria as:

> medium
> format
> periodicity (monographs, serials)
> audience or level
> language
> nationality (place of publication)
> time (date of publication or date of receipt)
> specific titles (when scope is limited to a stated list of documents)

When a documentary scope is further limited by qualitative selection criteria, these criteria should be described.

5.3 Domain

Domain refers to the "territory" covered in order to locate documents for the purpose of producing an index. The domain for an index to a single document is obvious, but it is not obvious for indexes to multiple documents and should be clearly described with respect to locational or territorial limits and the nature of acceptable sources (primary versus secondary sources).

a. **Locational limits.**

Within a particular subject and documentary scope, index producers can limit domain to a single collection, such as a national or special library, or to several collections of documents, in which case the index may be called a "catalog" or a "union catalog." Similarly, a domain can be limited to documents located in particular places or countries, or it can be universal, attempting to cover documents wherever they are located.

b. **Primary versus secondary sources.**

As a general rule, indexes should be based on primary sources, that is, the actual documents being indexed. When indexes are compiled on the basis of secondary sources (descriptions of documents such as abstracts, reviews, or entries in other indexes, databases, or catalogs) rather than the documents themselves, this practice should be clearly stated and the sources of data described.

5.4 Multiple versus unified indexes

Unified indexes should generally be preferred, but separate indexes may be justified when particular aspects are especially important and should be grouped together rather than dispersed among other entries in a unified index, for example, authors; persons or corporate bodies as subjects; or animal species, products, places, or ingredients; or particular types of documents, such as statutes, legal cases, reviews, maps, illustrations, or advertisements.

Separate indexes may also be desirable when it is awkward to assimilate verbal terms (using natural language words) with non-verbal terms, such as chemical formulae or patent numbers, or terms in different writing systems, such as the Roman alphabet and non-Roman scripts.

Separate indexes for particular subject facets or documentary types are often desirable in electronic indexes to facilitate targeted searches. When such separate indexes are provided, the search system should also allow for global searches across all indexes.

5.5 Codes and symbols

Most indexes within the scope of these guidelines will use the Roman alphabet, punctuation symbols, and Arabic and Roman numerals in accordance with normal English language (or other natural language) usage. Whenever any other symbols are used, for example, for music, choreography, chemistry, mathematics, or non-Roman writing systems, these symbols, the codes that govern their use, and the method for arranging nonalphanumeric symbols in displays should be described.

5.6 Display media

Indexes may be displayed in a wide range of media, including but not limited to print on paper, cards, microforms, or electronic displays linked to online databases or to indexes stored in such media as CD-ROMs or optical disks. Each medium has particular advantages and disadvantages that need to be considered in relation to the needs, habits, and preferences of users. The medium will influence most other options regarding access to the index.

5.7 Documentary units

The size and type of documentary units to which an index refers determine what can be retrieved. For indexes to verbal documents, documentary units can range from lines, statements, paragraphs, pages, sections, articles, chapters, monographs, serials, or series, to entire collections. Analogous units, such as map coordinates, motion picture frames, or quadrants of images, may be used for non-verbal

documents. The smaller the documentary unit, the more direct the referral to a particular topic or feature is likely to be.

Inherent documentary units may be preferable to physical medium units. Numbered or otherwise specified paragraphs or sections of a printed verbal text should be preferred to pages, because paragraphs or sections are more likely to constitute conceptual units. Indexes that refer to inherent documentary units may be used without change when a text appears in a variety of formats. (*See also* 7.4, Locators in displayed indexes.)

5.8 Indexable matter
Indexable matter consists of the portions of documents that are actually analyzed and indexed. Not all portions may be equally important. For example, introductory matter, appendixes, bibliographies, glossaries, illustrations, tables, advertisements, letters, and reviews may or may not need to be indexed, or they may be indexed at different levels of exhaustivity or specificity. Indexing also may be limited to specific portions of text (for example, titles, abstracts, first and/or last paragraphs, or captions). Decisions on appropriate indexable matter should be based on perceived importance to users of documentary units and should be explicitly stated.

5.9 Analysis method
Documents may be indexed through human intellectual analysis, algorithmic machine analysis, or combinations of human and machine analysis. The method of analysis used to produce an index should be stated.

5.10 Exhaustivity
Exhaustivity of indexing is the detail with which topics or features of a documentary unit are analyzed and described. Exhaustivity may be described as the number of unique terms, on average, assigned to or extracted from a documentary unit. It can range from summary indexing in which only a few terms are assigned per documentary unit, to highly exhaustive indexing in which hundreds of terms may be assigned or extracted. (Note that in a displayed index, a single heading often consists of multiple terms.)

Relative exhaustivity also depends on the size of documentary units, so that the same level of exhaustivity can represent quite detailed indexing when the documentary units are small, or summary-level indexing when documentary units are large. *See also* 5.7, Documentary units.

5.11 Specificity
Specificity refers to the closeness of fit between index terms and the topics or features they represent. For example, "pick-up trucks" (a specific term) may be used to represent that type of truck, or "trucks" (a generic term) may be used for all types of truck. "Specific" does not necessarily mean "narrow," because a specific term may be broad or narrow depending on the topic or feature to which it refers.

Specific indexing provides specific terms for all or most topics and features and results in a larger indexing vocabulary than more generic indexing. Generic terms may result in an excessive number of postings in non-displayed indexes or an excessive number of subheadings or locators in displayed indexes. Therefore, indexing terms should be the most specific available, and where a specific term is applied to a concept, a more generic term should not also be applied unless the generic concept is also addressed.

Exhaustivity combines with the specificity of index terminology to determine the depth of indexing.

5.12 Syntax
Index syntax provides the capability and the procedure for combining individual terms to form headings, subheadings, and sub-subheadings in order to furnish context for the initial term in an entry in displayed indexes and for combining individual terms into search statements for searching non-displayed indexes.

Examples of syntax are provided in section 7, Headings, entries, and search statements.

5.13 Vocabulary management
The terminology of an index should match the vocabulary of users whenever possible.

When documentary units consist of verbal texts in the same language as the index, the index should also link the vocabulary of documentary units to the vocabulary of users. Therefore, a large lead-in vocabulary is recommended as an aid to effective retrieval, with cross-references or other linking devices among synonymous and equivalent terms or headings.

An index should also assist users in adjusting the level of specificity of their requests to that of the index and documentary units by providing links between broader and narrower terms. An index can also suggest other avenues of search by linking related or associated terms.

Details of vocabulary management are treated more fully in section 6, Vocabulary.

5.14 Documentary unit surrogation; locators

Unless index terms or headings are attached to or embedded in the full text of a verbal document, indexes must include surrogates that represent or describe the documentary units to which they refer, and locators that point to the location of the documentary units. In many indexes the same representation serves as both surrogate and locator, especially when the full text of the documentary unit is present in the publication, as in back-of-the-book indexes where page, column, or paragraph numbers both represent the documentary unit (as surrogate) and point to its location (as locator).

In other indexes, especially those that point to documents not physically present, the surrogate may consist of an abstract and bibliographic citation. The locator consists of the part of the surrogate that points to the location of the documentary unit.

Some indexes use a series of surrogates and locators, for example, a brief entry number to link a term or heading to a fuller surrogate and locator, which may include a citation, abstract, and subject or feature terms.

Some form of locator is also used to link terms to documentary units or their surrogates in non-displayed indexes, but they are often not displayed to the user. Instead, internal links or pointers are used.

Surrogates, such as citations and abstracts, are widely used in non-displayed indexes. Because the text of these surrogates may often be searched, the use of abbreviations should be avoided, unless an abbreviation (or acronym) has become the preferred term. (*See also* 6.2.2, Spelling.)

Locators are treated more fully in section 7, Headings, entries, and search statements.

5.15 Surrogate display

When indexes use substantial surrogates such as citations, abstracts or annotations, it is helpful to provide a display that brings together in one place all information about a particular documentary unit, so that citation, abstract, or annotation, plus all descriptors or headings may be seen together.

Electronically stored indexes should provide options for a range of surrogate displays ranging from a brief citation through full description, including abstract and all index headings or terms.

5.16 Size of displayed indexes

The size of a displayed index is determined primarily by the number of indexed documents, their size and their documentary units, the exhaustivity of indexing, the specificity of the indexing language, the type of syntax used for the combination of terms, the extent of vocabulary management (number and style of cross-references), the size and style of document surrogates and locators, and the typography and format of the index, including size of pages or other display medium. When an index must fit a predetermined space, all of these aspects should be taken into consideration.

5.17 Index display and arrangement

For displayed indexes, the manner in which entries are arranged and formatted for display is vitally important, because access to the index is often dependent on this display. The arrangement of entries directly affects access, and the clarity of display format and guidance directly affect ease of searching and comprehension.

Displayed indexes may be arranged in alphanumeric, classified, or relational arrays. Classified or relational arrangements are used to bring related entries together, but they usually need their own alphanumeric indexes to facilitate access to relevant sections of the index.

Non-displayed indexes may be complemented with a displayed version of the index for visual browsing or scanning of entries.

These topics are treated more fully in sections 8, Display of index arrays, and 9, Alphanumeric arrangement.

5.18 Search interface

The format and arrangement of a displayed index constitutes its search interface.

For non-displayed indexes, the search interface is a computer program that provides the means for entering search terms or requests, for composing search statements, for exploring alternative terms, and for reviewing surrogates or the actual text of retrieved documentary units. Such search interfaces are relatively new in comparison to displayed indexes, and they are still very much the subject of experimentation and testing.

The search interface is an essential component of a non-displayed index. It is by means of the search interface that essential requirements for any information retrieval index are implemented for non-displayed indexes, including the capability of combining terms to specify desired topics or features and of exploring alternative and related search terminology.

6. Vocabulary

A major recommendation of these guidelines is that an index should provide access to topics or features of documents using the terminology of the documentary unit (when possible) and the terminology of prospective users. Terms or headings assigned to or extracted from documentary units should be linked with alternative synonymous, equivalent, broader, narrower, and other related terms or headings by means of cross-references or duplicate entries (6.8.1.1.) in displayed indexes. Similar linking devices are also required in non-displayed indexes. During electronic searches of non-displayed indexes, synonymous and equivalent terms may be substituted or added automatically to the search query or may be displayed, together with related terms, for selection when the search query is being composed (*See* Section 3, Function of an index, items d-h).

In displayed indexes, a preferred term should be selected for each concept, to which all alternative forms and related terms or headings are linked by means of cross-references *(see also* 6.8.1.1, Cross-references versus double entries).

In non-displayed indexes, preferred terms may serve as the anchors around which alternative and related terms are gathered.

6.1   Sources of vocabulary

The vocabulary for indexes may come from documents to be indexed, index users, human indexers, or compilations of vocabulary, such as thesauri, dictionaries, handbooks, and textbooks. The best source is often the text of the document(s) being indexed. Users of indexes are another valuable source, but it is often difficult or impossible to access their vocabulary directly. When it is possible to collect search terms employed by users, their terms should be incorporated into the index vocabulary. To the extent possible, indexes should link the vocabulary of users to the vocabulary of documents.

Expert indexers may be aware of user vocabulary that is not present in documents being indexed. Their vocabulary expertise should be used to the fullest extent possible.

Compilations of vocabulary (thesauri, subject heading lists, etc.) can also be useful. However, restricting vocabulary to static or closed lists of terms is usually not advisable, because it may lead to unnecessary constraints on access. When controlled vocabularies are used, they should accommodate changing user vocabulary and new terminology.

6.2   Forms of terms

Conventions and customs for the form of index terms have developed for English language indexes, as well as for other natural languages used for indexing. These conventions should be observed in the establishment of preferred terms for the convenience of users, unless there are overriding conventions in a particular discipline, field, or application. In these guidelines, only U.S. English language conventions and customs are cited.

6.2.1   Parts of speech

Nouns, including verbal nouns (gerunds) and noun phrases, are the preferred parts of speech for terms. Adjectives are often used to modify nouns; they are rarely used alone. Prepositional phrases are often used as subheadings to modify main headings or within headings to modify lead terms. Prepositions are also used as role indicators to link terms in string indexes. Adverbs should not be used unless they form an integral part of a term, for example: "very high frequency."

Examples:

courts of inquiry	[prepositional phrase modifying lead term]
dishonorable discharge	[lead term is adjective modifying a noun]
swimming	[verbal noun (gerund)]
alcohol use	[prepositional phrase as
by naval personnel	subheading]
advertising. Japanese cars. Germany effects on sales	
sales. Japanese cars. Germany. effects of advertising	[preposition as role indicator in string indexing]

6.2.2   Spelling

For U.S. indexes, normal U.S. spelling should be used. If there is more than one normal spelling (for example, groundwater, ground-water, ground water), the one used in the document(s) being indexed should be preferred if used consistently. Otherwise, one spelling should be chosen and employed consistently.

Alternative spellings should be linked to the preferred spelling of the term. This is especially important in non-displayed electronic indexes, because even minor variations in spelling (for example, aluminum / aluminium) may lead to the loss of access.

Common contractions, abbreviations, and acronyms should be used as terms or linked to terms. Their spelling should conform to common usage. *See also* 6.2.3, Capitalization, and 6.4, Synonymous and equivalent terms.

6.2.3 Capitalization

All terms, except proper nouns or names, adjectives based on proper nouns, and certain acronyms should be written with lowercase letters. In proper nouns, the first letter of the first word and the first letter of each succeeding word, other than conjunctions, prepositions, and articles, should be capitalized. Acronyms of names of organizations should follow usage of the organization (for example, NATO, Unicef). Other acronyms should follow conventional capitalization (for example, radar, COBOL).

AACR2 mandates only an initial capital letter for English titles of documents, except for proper nouns, names, or adjectives. However it also calls for the capitalization of all major words in the conventional or formal names of important public documents, such as The Articles of Confederation.

Examples:
> Christian education
> COBOL
> Eli Lilly and Co.
> German engineering
> Japanese cars
> Jewish documents
> Marlborough, John Churchill, first Duke
> NATO
> radar
> The tempest.
> Unicef
> United States. Department of Education.
> The wind in the willows.

6.2.4 Singular and plural forms

In English language indexes, it is the convention and custom to use the plural form for terms denoting discrete objects (countables) and the singular form for mass terms and most abstract concepts (noncountables). The plural is used when the question on quantity asks "How many?" The singular is used when the question as to quantity asks "How much?" Indexes should conform to these conventions.

If the singular and plural forms have different meanings, both forms should be used separately if both are needed to represent topics or features of a text. The distinction between them should be clarified with qualifiers:

> memories (reminiscences)
> memory (brain function)
>
> building (process)
> buildings (edifices)

6.2.5 Articles

The use of articles should be avoided in index terms for topics and features.

Initial articles should not be omitted from names of persons, places, or corporate bodies or from titles. Articles should not be transposed. *See also*

6.2.9.2, Corporate body names, 6.2.9.3, Geographical names, 6.2.9.4, Titles of documents, and 6.2.9.5, First lines.

6.2.6 Compound terms

As a general rule, a single term (as opposed to a pre-coordinated or multi-level heading) should represent a single concept. What constitutes a single concept will vary from situation to situation. Frequently two or more terms become "bound" together to express a concept, such as "information science," "birth control," or "form of government." When such compound terms become established, they should be preferred to the alternative of forcing the combination of two separate terms, for example, "science" and "information," or "control" and "birth" or "conception" at the time of searching or when combining terms for headings and entries.

Use of compound terms also helps to avoid "false drops," such as the retrieval of documents on "library schools" when "school libraries" is intended. Similarly, terms like "information" and "science" can occur in many contexts where "information science" is not discussed.

6.2.7 Antonyms and associated terms

When antonyms and other closely associated terms (for example, honors and awards) are combined to form compound terms, the terms not chosen as lead terms should be linked to the compound term by cross-references.

| awards | see | honors and awards |
| evil | see | good and evil |

NOTE: The form and presentation of the cross-references will differ in displayed and non-displayed indexes (*see also* 6.8, Display of vocabulary in indexes).

6.2.8 Word order in multi-word terms

Terms consisting of more than one word, including compound terms, should be used in natural language order without inversion. For example:

> deciduous trees [not: trees, deciduous]

However, access should also be provided by means of substantive terms that are not in the lead position, for example:

trees, deciduous	see	deciduous trees
or:		
trees	see also	deciduous trees

An acceptable alternative to inverted terms is a heading-subheading combination, for example:

> trees
> deciduous
> evergreen

6.2.9 Proper names and titles of documents

Names of persons, corporate bodies, and places should be established, to the extent possible, in accordance with standards used in library practice, because it is advantageous for users to experience a measure of uniformity across information retrieval systems. *The Anglo-American Cataloguing Rules,* 2nd edition (AACR2), provides detailed guidance for the establishment of names. In the few instances where these guidelines diverge from AACR2, differences are noted.

6.2.9.1 Personal names

Personal names should be provided in the form most commonly used, and in as full a form as possible when there is more than one common form. Limiting forenames to initials invites confusion, unless initials are part of the commonly used form of a name (for example, D. H. Lawrence).

When more than one name or form of name is in use, they should be linked as synonymous terms.

Where surnames are in common use, names should be entered under surname, followed by a comma and any given names or initials:

> Lee, Kuan Yew
> Wheatley, Henry B.

Persons identified only by a given name or forename should be entered under that name, followed if necessary by a title of office or other distinguishing epithet:

> Boudicca, Queen of the Iceni
> Leonardo da Vinci
> Ethelred the Unready

Persons normally identified by a title of honor or nobility should be entered under that title, expanded if necessary by their family name:

> Dalai Lama
> Marlborough, John Churchill, first Duke

Compound and multiple surnames, whether hyphenated or not, should be entered under the first part, unless usage favors another practice. For example, Portuguese names are customarily entered under the last part. Cross-references should be established among all possible forms of entry:

> Layzell Ward, Patricia [with cross-reference from: Ward, Patricia Layzell]
> Pérez de Cuéllar, Javier [with cross-reference from: De Cuéllar, Javier Pérez; and from: Cuéllar, Javier Pérez de]
> Trevor-Roper, Hugh [with cross-reference from: Roper, Hugh Trevor]

When two or more persons have the same name, their names constitute homographs and should be distinguished with qualifiers consisting of a fuller form of name or dates where available; otherwise, occupation, title, or nationality should be used:

> Butler, Samuel (1612-1680)
> Butler, Samuel (1835-1902)
> Lawrence, D. H. (David Henry)
> Lawrence, D. H. (Derek Herbert)
> Rickert, Heinrich (philosopher)
> Rickert, Heinrich (politician)

In indexes to biographies, qualifiers may also be used to indicate the relationship between a person and the main subject of the biography, for example: (son), (brother), (mother).

NOTE: AACR2 calls for two different formats for qualifiers, depending on the type of qualifier and the form of name. In contrast, these guidelines call for the use of parentheses for all qualifiers, for the sake of consistency within indexes and to conform to recommendations of the NISO thesaurus standard.

6.2.9.2 Corporate body names

Names of corporate bodies should be entered without transposition in the form most commonly used by the body itself. If more than one form is common, the fuller form should be used. If an abbreviation or acronym is the commonly used form, that form should be used instead of the full form. Initial articles that are part of the commonly used form should neither be omitted nor transposed:

> Der Blaue Adler (association) [not: Blaue Adler, Der (association)]
> The Club (London) [not: Club (London); Club, The (London)]
> Eli Lilly and Co. [not: Lilly, Eli, and Co.]
> The H. W. Wilson Company [not: H. W. Wilson Company; Wilson, H. W., Company, The]
> Helena Rubinstein, Inc. [not: Rubinstein, Helena, Inc.]
> Helene Curtis, Inc. [not: Curtis, Helene, Inc.]
> J. Whitaker & Sons [not: Whitaker, J., & Sons]
> The Library Association (United Kingdom) [not: Library Association (United Kingdom); Library Association, The (United Kingdom)]
> TRON Project [not: The Realtime Operating System Nucleus Project]
> UNESCO [not: United Nations Educational, Scientific, and Cultural Organization]

NOTE: AACR2 calls for the elimination of initial articles from the names of corporate bodies. In contrast, these guidelines follow instead the general AACR2 principle that a corporate body should be entered "under the name by which it is commonly identified." When the name by which a body is commonly identified includes an initial article, it should not be deleted (e.g., The Club). Neither these guidelines nor AACR2 call for the deletion of initial articles from the names of persons or from place names. These guidelines give preference to as much consistency of practice as possible across all types of names.

Unless abbreviations constitute or are part of the commonly used form, names of corporate bodies should not be abbreviated:

> United States. Department of Education [not: DOE]
> United States. Department of Energy [not: DOE]
> University of Nebraska [not: Univ. of Nebraska]
> New York University [not: NYU]

NOTE: In some situations, "U.S." may be considered a commonly used form for "United States."

Corporate bodies that are parts of larger bodies should be entered under their own names unless the name is indistinct or implies subordination. If the name needs the name of a higher body, the lowest level body that can be entered directly under its own name should be used:

> Public Library Association. Audiovisual Committee
> [not: American Library Association. Public Library
> Association. Audiovisual Committee]
> United States. Department of Health and Human Services
> [not: Department of Health and Human Services]

When there are several hierarchical levels, as many intervening bodies as necessary should be included in the name to avoid confusion:

> American Library Association. Resources and Technical
> Services Division. Board of Directors.

Identical names for different bodies constitute homographs and should be distinguished with qualifiers:

> Metropolitan Museum of Art (Cleveland, OH)
> Metropolitan Museum of Art (New York, NY)

Cross-references should link different names for the same body and all possible forms of entry, including inverted forms and forms incorporating names of larger bodies:

> American Library Association. Public Library Association
> see Public Library Association
> Department of Health and Human Services see United States.
> Department of Health and Human Services
> DOE see United States. Department of Education; United
> States. Department of Energy
> Medicine, National Library of see National Library of Medi-
> cine
> The Club (London) see The Club (London) — arranged
> under "Club" [enter reference under "The"]
> The Realtime Operating System Nucleus Project see
> TRON Project [enter reference under both "The
> . . ." and "Realtime . . ."]
> U. S. see United States.
> United States. DOE see United States. Department of Educa-
> tion; United States. Department of Energy
> United Nations Educational, Scientific, and Cultural Orga-
> nization see UNESCO
> Whitaker, J., & Sons see J. Whitaker & Sons
> Wilson, H. W., Company see The H. W. Wilson Company —
> arranged under "H.W. Wilson"

NOTE: The form and presentation of the cross-references will differ in displayed and non-displayed indexes. *See also* 6.8, Display of vocabulary in indexes.

6.2.9.3 Geographical names

Geographical names should be as full as necessary for clarity, with qualifiers to avoid confusion between otherwise identical names:

> Middletown (CT)
> Middletown (OH)
> Middletown (Powys, Wales)

Abbreviations should not be used unless there is a commonly accepted standard, such as U. S. Postal Service abbreviations for states of the United States.

Prefer the English form if there is one in general use. Otherwise use the form in the official language of the country:

> Buenos Aires
> Florence [not: Firenze]

An article or preposition should be retained in a geographical name of which it forms an integral part:

> Des Moines
> Las Vegas
> Los Angeles
> The Dalles
> The Hague

See section 9.4 regarding the arrangement of headings with initial articles.

NOTE: AACR2 calls for the use of older forms of abbreviation for the states of the United States — forms that are no longer in widespread use with the advent of official U.S. Postal Service abbreviations.

6.2.9.4 Titles of documents

To the extent possible within typographic constraints, titles of documents should not be changed or altered. For example, the name of a chemical should not be substituted for a chemical symbol or a numeral replaced with its name.

When it is necessary to correct an error in a title, the correction should be placed in square brackets, prefaced by "i.e.":

> The Paul Anthony Buck lectures [i.e., The Paul Anthony
> Brick lectures]

Titles should not be abbreviated unless very long, and any omissions should be indicated by three dots (ellipsis):

> *Inquiry into the nurturing and elimination of life forms within*
> *marginally controlled ecosystems . . .*

Titles with numerals or other nonalphabetic symbols (e.g., chemical, mathematical, musical, and scientific symbols) should be linked with

equivalent cross-references from names of numerals or names of symbols, if any:

> *1984* see *Nineteen eighty-four*
> *Dollars & sense* see *$$$ & sense*
> *Ten sixty-six and all that* see *1066 and all that*
> *Two thousand and one* see *2001*

If necessary to avoid confusion, qualify the title of a document with a term that will indicate that it is a document:

> *Charlemagne* (play)
> *Genesis* (Anglo-Saxon poem)

If necessary for identification, names of creators, places of publication, dates, or other qualifiers may be used:

> *Ave Maria* (Gounod)
> *Ave Maria* (Schubert)
> *Ave Maria* (Verdi)
> *Natura* (Amsterdam)
> *Natura* (Milan)

An initial article should not be omitted or transposed to the end of the title:

> *Das Kapital* (Marx) [not: *Kapital* (Marx); *Kapital, Das* (Marx)]
> *The tempest* [not: *Tempest; Tempest, The*]

See section 9.4 regarding the arrangement of headings with initial articles.

NOTE: AACR2 calls for the deletion of initial articles from titles. In contrast, these guidelines follow the general AACR2 principle that one should "use the title or form of title in the original language by which a work ... has become known through use in manifestations of the work...." According to this principle, it is inconsistent to delete initial articles in titles such as "Das Kapital."

Prepositions at the beginning of a title should be retained:

> *An die Musik*
> *To the lighthouse*

6.2.9.5 First lines

In first-line indexes, initial articles should be retained in natural order, not transposed. *See* section 9.4 regarding the arrangement of headings with initial articles.

6.2.10 Romanization

Names and words rendered into roman script from another writing system should be based on standard romanization tables unless a well-established English language form exists. Use the ALA-LC romanization tables (*see* section 10, References for full citation).

> Dayan, Mosheh
> Solzhenitsyn, Aleksandr
> *but:*
> Alexander the Great   [not: Alexandros ho Megas]
> Avicenna   [not: Abu Ali ibn Sina]
> Confucius   [not: Kung Fu Tzu]

Cross-references should link alternative forms of romanized names and other terms.

6.3 Homographs

Identical terms that represent different concepts or features can cause confusion and should be differentiated by the addition of a qualifier:

> races (anthropology)
> races (sport)

See also discussion of particular kinds of homographs in 6.2.9.1, Personal names, 6.2.9.2, Corporate body names, 6.2.9.3, Geographic names, and 6.2.9.4, Titles of documents.

6.4 Synonymous and equivalent terms

Research and practice indicate that index users tend not to agree on terms for particular concepts or features. Therefore, it is essential that indexes provide for alternative terms, including abbreviations and acronyms. All terms that may be used for the same topic or feature within the context of an index should be linked so that any such term will lead searchers to the same documents.

Terms including numerals should be linked with equivalent terms having the names of numerals:

> nineteenth century *see* 19th century
> 5 year plans *see* five year plans

Small variations in terms that have little or no impact on alphanumeric arrangement cause few problems in displayed indexes, but such variations can cause terms to be completely lost in indexes that are searched electronically. Therefore, terms with even small variations in spelling or endings (for example, aluminum / aluminium) should be linked in electronic indexes. All such terms with noncontiguous arrangement positions should be linked in displayed indexes.

What constitutes equivalency between terms depends on the level of specificity used in an index. Equivalency is sometimes established between a narrower term that is too specific for the design of an index or thesaurus and a broader term. For example, "chairs" may be a lead-in (nonpreferred) term for "furniture." This relationship is sometimes indicated with a simple "see" or "use" reference and sometimes with a "see under" reference (when a distinction is made between synonymous terms versus narrow-to-broad "equivalent" terms):

> cars *see* automobiles
> convertibles (automobiles) *see under* automobiles

"See under" is also used for a reference from a term to an instance of that term as a subheading under another term:

> editing *see under* books; journals

However, such references would be clearer if the target heading and subheading are included in full (*see also* 6.8.1.3, Location of "see also" cross-references):

> editing *see* (or *see also*) books — editing; journals — editing

When indexes use one term (often called a descriptor) as the preferred term, and link alternative, synonymous, and equivalent terms to it, the unused terms may be usefully displayed as "used for" or "equivalent terms" in a note or display under, or linked to, the preferred term. These "used for" terms help to define the scope of the preferred term (descriptor):

> automobiles
> > *used for* cars, motor vehicles, particular models and types of automobile, such as convertibles, jalopies, sedans, and for particular brands or makes, such as Buicks, Fords, Plymouths, etc.

NOTE: The form and presentation of the cross-references will differ in displayed and non-displayed indexes. *See also* 6.8, Display of vocabulary in indexes.

6.5 Hierarchical relationships among terms

Links between narrower and broader terms are important to guide the narrowing of a search to particular members of a larger set of terms (for example, from "computers" to particular types of computers) or the broadening of a search to all members of a larger set (for example, from "Labrador retrievers" to all species of dog). However, links that point to broader terms may not be appropriate in all types of indexes, particularly when space is limited or when the index includes few additional entries for the species, individual members, or parts of the broader genus or class represented by the broader term.

Examples of hierarchical relationships include:

a. genus/species:
   furniture/chairs
   behavior/aggression
   bears/polar bears

b. discipline/constituent studies:
   geology/petrology

c. class/individual members:
   bridges/Golden Gate Bridge
   standardizing bodies/NISO

d. entity/parts :
   buildings/rooms
   United Nations/UNESCO
   population/immigrants

e. larger and smaller geographic units:
   Europe/France
   New Jersey/Middlesex County/New Brunswick

6.6 Other relationships

Links between terms having relationships other than hierarchical provide searchers with additional options for improving their searches. For example:

a. discipline/objects studied:
   botany/plants
   physical chemistry/molecules

b. theoretical study/application or technology:
   dynamics/mechanical engineering
   state ownership/nationalized industries

c. activity/agent:
   photography/cameras; photographers
   singing/voice; singers

d. activity/thing acted upon:
   angling/fish
   dentistry/teeth

e. activity/product:
   aggression/violence
   cartography/maps

f. closely related topics not generally differentiated in common parlance but differentiated in a particular index:
   boats/ships
   pottery/porcelain

g. related topics separated in a particular index when semantically related nouns and adjectives take different forms:
   law/legal . . .
   women/female . . .

6.7 Changes in terminology

In continuing indexes, care should be taken to link older and newer terms that are synonymous, equivalent, or closely related. The date of the change should be indicated. Examples of changing terminology include:

a. the introduction of a new term as a substitute for an older term:
   wireless [changed in 1950 to] radio
   Negroes [changed in 1968 to] Blacks [changed in 1992 to] African Americans

b. name changes:
   Ceylon [changed in 1972 to] Sri Lanka
   Harris, Jessica [changed in 1979 to] Milstead, Jessica
   Rutgers University. Graduate School of Library Service [changed in 1982 to] Rutgers University. School of Communication, Information, and Library Studies

c. the use of additional terms to express narrower topics previously embraced by a broader term:
   computers *beginning 1990 see also* microcomputers, minicomputers

For alternative formats, *see also* 6.8.3, Scope and history notes.

6.8 Display of vocabulary in indexes

Information about vocabulary and relations among terms or headings (for example, synonymy, equivalence, homography, hierarchy, association) should be presented as an integral part of an index. Searchers should not have to consult separate, unconnected files for vocabulary information. How integration of vocabulary information and index terms or headings is accomplished depends largely on the medium of the index and whether or not indexes are displayed for searching.

6.8.1 Vocabulary information in displayed indexes

In print media, indexes should be displayed as ordered arrays of entries, because it is only through such displays that users can enter the index. Such displays are becoming more common for electronic indexes as well, especially those designed for non-expert users, such as online public access catalogs (OPACs) in libraries and many CD-ROM indexes.

In displayed indexes, whether in print or electronic media, vocabulary information should be integrated into the same sequence of entries that describe documents, using a variety of notes and cross-references.

In electronic displayed indexes, cross-references should provide for the possibility of an immediate link to preferred terms or, upon selection, to any desired related terms (narrower terms, broader terms, or other associated terms).

6.8.1.1 Cross-references versus double entries

In closed-end, one-time indexes (as opposed to open-end, continuing indexes), a duplicate entry under an alternative heading should be made if it would (a) contain the same number or fewer lines than a *see* reference; (b) have three or fewer locators; or (c) have no or only one subheading.

> automobiles 23, 45
> cars 23, 45
> *not:*
> automobiles 23, 45
> cars *see* automobiles

6.8.1.2 Cross-references to multiple terms or headings

When a cross-reference refers to multiple terms or headings, these should be listed by category if the nature of the relationship is indicated (for example, synonymous, equivalent, broader, narrower, related), and within category in alphabetical order, separated by semicolons. If the nature of relationships is not indicated, the referenced terms or headings should be in a single alphabetical sequence.

> sexuality
> *used for equivalent term* sexual nature.
> *see also narrower terms* bisexuality; chastity; heterosexuality; homosexuality; incest; necrophilia; sublimation.
> *see also related terms* gender; sex; sexual identity; sexual problems.
> *see also broader terms* behavior; human nature.

6.8.1.3 Location of "see also" cross-references

"See also" cross-references should normally follow any locators related to a single term or heading from which they refer:

> bears 100, 217, 923 *see also* badgers; koala bears; raccoons

However, because their purpose is not only to suggest additional terms or headings that may be useful, but also to suggest alternative terms or headings, "see also" cross-references should precede subheadings in those indexes in which headings have numerous subheadings. Placing "see also" references before subheadings will prevent these references from being overlooked or found only after perusing unwanted subheadings. In these cases, "see also" cross-references should be clearly distinguished from subheadings. They can, for example, be displayed on lines indented more deeply than subheadings or within boxes:

> economics 144, 195, 229, 363
>
> | *see also* assets; banking; business; commerce; firms; transport; wealth |
> | --- |
>
> bibliographies 208
> mathematical models 160
> statistics 155

Cross-references should be attached to the heading or the subheading from which they refer:

> economics
> statistics 155
> *see also* econometrics

When a cross-reference leads from a subheading under one main heading to the same subheading under another main heading, the reference should include both the main heading and the subheading referred to:

> economics
> statistics 155
> *see also* economic policy — statistics

6.8.2 Vocabulary information in non-displayed indexes

Methods for effectively displaying vocabulary information for non-displayed indexes are not yet well established. The development of such

methods should be encouraged, because vocabulary information is essential for efficient use of indexes of all types.

Vocabulary information, including relations among terms, should be displayed to users in conjunction with search statement preparation, in the same medium that is used to search the index. Users should not have to consult a completely separate vocabulary file, copy terms, and then re-enter them in order to use them for a search.

Users should have the option of automatic or selective addition or replacement of synonymous and equivalent terms.

In the automatic mode, if index terms are limited to preferred terms, a preferred term should replace any synonymous or equivalent terms. When all such terms may be used, then all should be added to the search. The user should be notified of any modification of the search.

In the selective mode, preferred, synonymous, or equivalent terms should be displayed so that they may be selected for replacement or addition.

Users should have the option of seeing displays of other vocabulary information and selecting broader, narrower, or other related terms for use in their search, either in addition to terms already selected, or in place of those terms.

6.8.3 Scope and history notes

A scope note clarifies the scope or application of a term. It should be set off from the term itself by means of typography or layout.

A history note explains changes in usage over time in a continuing index:

"Radio" replaced "wireless" in 1950.

This information may also be presented in the form of a cross-reference:

radio *in pre-1950 volumes see* wireless
wireless *see* radio

When both old and new entries are present in the same index, "*see also*" references should be used:

radio *see also* wireless *for entries before 1950*
wireless *see also* radio *for entries from 1950 onward*

When cross-references refer to newer terms that formerly were subsumed under a broader term, dates should be attached to terms so that users know when such terms were introduced:

computers *see also* microcomputers *from 1977 onward;*
    minicomputers *from 1972 onward*
microcomputers *see also* computers *for entries before 1977*
minicomputers *see also* computers *for entries before 1972*

7. Headings, entries, and search statements

A major recommendation of these guidelines is that an index should make it possible for users to search for multiple topics or features, or aspects of topics or features, in combination. In displayed indexes, this capability is provided by combining terms into headings of one or more levels. In non-displayed indexes, this capability is provided by a search interface that accepts search statements with multiple terms.

The means for combining terms in displayed headings and entries or in search statements is called syntax. A variety of syntactic styles and methods are available. This technical report does not recommend any particular syntactic method for either displayed or non-displayed indexes; it simply states that every index should incorporate some syntactic method so that terms can be combined for the purposes of searching. The type, format, and size of an index and the needs and preferences of users will govern the choice of appropriate method.

In displayed indexes, the combination of terms takes place in advance of the search; pre-combined headings and entries are provided by the index producer. In non-displayed indexes, the combination of terms for searching takes place at the time of the search. The index producer provides the means for combining search terms, but the actual combination of search terms is performed by the user of the index.

Pre-coordinated combinations of terms are also used in non-displayed indexes, but search statements should not be restricted to these pre-coordinated combinations. Pre-coordinated headings can be useful for linking terms that represent relationships present in the topics or features of the documentary unit (*see also* 7.5.4, Links and role indicators).

Examples of the major types of index syntax available appear in this section.

Also, because locators are an integral part of entries in displayed indexes, they are discussed here.

NOTE: This section includes examples of syntax only. The absence of cross-references in these examples in no way implies that the use of any particular syntactic method or style will by itself fulfill the recommendations of this technical report without also providing for some method of linking synonymous, equivalent, broader, narrower, and other related terms. *See* section 6, Vocabulary.

7.1 Entries in displayed indexes

In displayed indexes, an entry consists of at least a heading and the locator that identifies or points to the documentary unit that the heading describes. A heading may consist of more than one index term. In a multi-level heading, the main heading is followed by a subheading, which in turn may be followed by a sub-subheading, and possibly by additional headings at successive levels of subordination.

In displayed indexes, identical headings for subsequent entries are generally not repeated. Nevertheless, each locator represents a single entry. For example, the following display consists of 6 entries.

```
economics  144, 195, 229
    bibliographies 208, 244, 363
```

The actual entries, displayed in full, are:

```
economics      144
economics      195
economics      229
economics — bibliographies      208
economics — bibliographies      244
economics — bibliographies      363
```

In displayed indexes, a heading or subheading should not be followed by more than 5 locators (that is, 6 or more entries under identical headings should be avoided). However, when locators themselves contain distinguishing information, such as bibliographic citations or abstracts, then a larger number of entries can be accommodated under the same heading or subheading.

NOTE: In applying this criterion, a continuous sequence of locators may be considered to constitute a single locator. For example, "economics 144-145, 195-200, 229-230" may be considered to be a heading followed by 3, rather than 10 locators.

Entries with identical headings may be distinguished and made more specific by adding terms or subheadings for context or aspect.

Even when single term entries are unique, the addition of terms or the use of subheadings can provide information to help a user determine whether the documentary unit might be useful. The addition of a term indicating context or aspect can often relieve users of useless pursuits by providing criteria for eliminating irrelevant references without having to consult each one.

7.2 Syntax in displayed indexes

Syntax used in displayed indexes is often called "pre-coordination" syntax because the combination of terms takes place prior to the presentation of the index.

The sections that follow describe examples of pre-coordination syntax.

NOTE: Locators are omitted from all examples. In actual indexes, locators would be placed at the end of each relevant heading, subheading, or string of terms.

7.2.1 Ad hoc syntax

Syntax is often applied on a case-by-case, heading-by-heading basis in closed-end indexes such as book indexes. Individual headings are created as appropriate for the nature of the document(s) being indexed and for the needs of prospective users.

Prepositions should be avoided at the beginning or end of subheadings unless they are needed to avoid ambiguity:

```
clothing
    rationing [not: rationing of]
computers
    compared with abacus
    for management
    in hospitals
    management of
crops
    irrigation [not: irrigation of]
    taxation [not: taxation of]
    sale [not: sale of]
management
    of computers
    use of computers
rationing
    clothing
    food
    fuel
    health care
```

NOTE 1: The above examples are shown as they would appear in a printed index, but other styles of linking headings and subheadings are also possible, depending on index format and medium.
NOTE 2: Ad hoc syntax should be accompanied with cross-references for linking synonymous, equivalent, broader, narrower, and other related terms. *See* section 6, Vocabulary.

7.2.2 Natural language syntax

Some indexes attempt to take advantage of the syntax, or word order, of exiting segments of text (such as titles of documents) in order to provide syntax and context for index headings.

The most common examples are keyword indexes, in which simple concordance algorithms are used to rearrange the title or other text segment under each keyword in the title or segment that is not on a stop list of insignificant words. Cross-references or other forms of vocabulary control or management for linking equivalent, narrower, broader, or other related terms are almost always absent from such indexes.

The chief advantage of natural language syntax

indexes, or concordances, is the speed and low cost of their production. However, many authorities do not consider such term concordances to be true indexes because the added value usually associated with indexing is so limited — nothing more than the automatic rearrangement of terms.

Nevertheless, keyword indexes have been widely used. Their adequacy in performing the functions of an index (section 3) is most directly related to the adequacy of titles or other text segments as representations of the content and features of documentary units (because no other indexing is done, either by indexers or by more sophisticated computer algorithms). Some titles do describe, in summary fashion, the content or features of documents. Others clearly do not.

Natural language syntax indexes would conform to the minimum guidelines of this technical report if the titles or text segments used to represent documentary units were to be augmented, as necessary, to insure adequate representation of the content and features of documentary units, and if cross-references (or other linking methods) for equivalent, narrower, broader, and other related terms were to be integrated into their displays.

The most common formats for natural language syntax keyword indexes are KWIC (KeyWord-In-Context), KWOC (KeyWord-Out-of-Context, and KWAC (KeyWord-Alongside-Context). In the KWIC format, the natural word order of the title or other text segment is preserved on both sides of a keyword; the keyword is arranged in alphabetical order down the center of the page or column with surrounding title or text segment to the right and left. Term pairs, or longer phrases, are preserved on both sides of the keyword, as in "ALTERNATIVE theatre" and "BRITISH theatre" in the first two examples, or "alternative THEATRE" and "British THEATRE" in the last two examples:

ther: the critic and the	ALTERNATIVE theatre. This that
The language crisis in	BRITISH theatre: the drama of
...	
age of crisis in British	THEATRE: the drama of cultural
tic and the alternative	THEATRE. This that and the other:

The KWOC format preserves the traditional format of an index with the lead keyword on the left, followed by the title or other text segment as a subheading. But in this format, word pairs or phrases, such as "alternative theatre" and "British theatre" in these examples, are no longer preserved in conjunction with the lead term. Persons interested in "alternative theater" or "British theater" must scan all subheadings in order to find these

two-word terms:

ALTERNATIVE
　　This that and the other: the critic and the ALTERNA-
　　　　TIVE theatre.
BRITISH
　　The language of crisis in BRITISH theatre: the drama
　　　　of cultural pathology.
...
THEATRE
　　The language of crisis in British THEATRE: the
　　　　drama of cultural pathology.
　　This that and the other: the critic and the alternative
　　　　THEATRE.

In order to preserve word combinations with the lead term, and at the same time to restore the traditional format with lead or main term on the left, the KWAC format was introduced. Unlike the KWIC format, however, term pairs or phrases are preserved only to the right of the keyword, not to the left:

ALTERNATIVE
　　theatre. This that and the other: the critic and the
BRITISH
　　theatre: the drama of cultural pathology. The
　　　　language of crisis in
...
THEATRE
　　the drama of cultural pathology. The language of
　　　　crisis in British
　　This that and the other: the critic and the alternative

NOTE: As with all syntax methods in this section, indexes employing natural language syntax do not conform to the guidelines of this technical report unless they include cross-references or other methods for linking synonymous, equivalent, broader, narrower, and other related terms. *See* section 6, Vocabulary.

7.2.3 Subject heading syntax

Syntax may be provided by using established lists of subject headings. Such lists generally include headings consisting of precombined terms or provisions for combining terms at the time of indexing in accordance with rules or patterns. Combination of terms has been achieved in three ways:

a. linking terms to each other by typographic devices, for example, em-dashes or two hyphens, colons, or full stops (periods).

　　animals — diseases — chemotherapy
　　libraries : New Jersey : New Brunswick

b. modifying the lead term with other words to distinguish between terms that have multiple definitions:

　　plates (engineering)
　　plates (tableware)

c. creating phrase headings:

　　social work with the homeless
　　telephone assistance programs for the poor

These methods may be combined, to create headings such as "plates (engineering) — vibration."

These guidelines advocate natural language word order for terms and headings. Inverted headings, such as "students, foreign," reflect earlier practice, still extant in some lists of subject headings. *See* 6.2.8, Word order in multi-word terms. NOTE: Subject heading syntax should be accompanied with cross-references for linking synonymous, equivalent, broader, narrower, and other related terms. *See* section 6, Vocabulary.

7.2.4 Permuted indexes

Permuted indexes display every possible combination of index terms or descriptors. Such terms are selected by indexers or extracted from text according to various computer algorithms. Because the number of such combinations increases exponentially as the number of words in each heading increases, permuted indexes are usually restricted to headings consisting of no more than two words in combination. The result is a main heading-subheading combination for each word pair.

The following sample headings were generated from the following sets of terms for two different documents; the same set of headings would also be entered under every other keyword in each set as main heading:

Document 1: American Mercury (periodical); editors and editing; Ku Klux Klan; Mencken, H. L.; Methodist Episcopal Church (South); temperance movements; 1910-33.

Document 2: Atlanta (GA); Fowler, Charles Lewis; Georgia; Ku Klux Klan; Lanier University; 1917-22.

Ku Klux Klan
    1910-33
    1917-22
    American Mercury (periodical)
    Atlanta (GA)
    editors and editing
    Fowler, Charles Lewis
    Georgia
    Lanier University
    Mencken, H. L.
    Methodist Episcopal Church (South)
    temperance movements

NOTE: Permuted syntax should be accompanied with cross-references for linking synonymous, equivalent, broader, narrower, and other related terms. *See* section 6, Vocabulary.

7.2.5 String indexing

String indexing uses computer algorithms to combine multiple terms into multiple headings. Each heading has a different term as its lead or main term. The set of terms is treated like a "string" or sequence of terms that is rearranged under each lead term. The terms themselves may be assigned by human indexers, but their manipulation into index entries is governed by computer algorithms. The following sections illustrate common types of string indexing.

7.2.5.1 Rotated terms

The simplest form of string indexing places each term, in turn, in the lead position followed by all other terms in alphanumeric order. Relationships among terms are not indicated. In the following examples, numerals are arranged after letters:

American Mercury (periodical). editors and editing. Ku Klux Klan. Mencken, H. L. Methodist Episcopal Church (South). temperance movements. 1910-33.

editors and editing. American Mercury (periodical). Ku Klux Klan. Mencken, H. L. Methodist Episcopal Church (South). temperance movements. 1910-33.

Ku Klux Klan. American Mercury (periodical). editors and editing. Mencken, H. L. Methodist Episcopal Church (South). temperance movements. 1910-33.

NOTE: Rotated syntax should be accompanied with cross-references for linking synonymous, equivalent, broader, narrower, and other related terms. *See* section 6, Vocabulary.

7.2.5.2 Faceted indexing

Faceted indexing arranges terms in entry strings according to facet relationships. Terms are first placed into facets or tagged with facet indicators. A computer algorithm then uses a predetermined facet order (citation order) to arrange terms in entries.

Faceted indexing that is designed to accommodate broad subject areas uses generic term categories like location, key system or entity, action or effect of action, agent or instrument, viewpoint or aspect, particular instance, document form, and target user. These primary categories are sometimes modified by secondary categories such as part, property, role definer, modifiers, dates, and various connectives, which serve to indicate relationships among terms where these are not self-evident or are ambiguous. The following coded terms will produce the following headings and subheadings:

[location] Germany
[key entity] cars
[modifier] Japanese
[action] sales
[role definer] effects of/on
[agent, instrument] advertising

Germany
    Japanese cars. sales. effects of advertising
cars. Germany
    Japanese cars. sales. effects of advertising
Japanese cars. Germany
    sales. effects of advertising
Sales. Japanese cars. Germany
    effects of advertising
advertising. Japanese cars. Germany
    effects on sales

When faceted indexing is applied to a narrow subject area, facets tend to be tailored to aspects of particular interest in that subject area. In literature, for example, terms may be placed into facets such as specific literatures, performance media, languages, periods, individuals (real), groups/ movements, genres, works, literary techniques, themes/motifs/ figures/characters, influences, sources, processes, methodological approaches, theories, devices/tools, and disciplines. The designated citation order for these facets determines the order of terms in subheadings:

> homosexuality
> > English literature. short story. 1900-1999.
> > > Forster, E. M. "Dr. Woolacott." symbolism.
> > > > treatment of salvation; *homosexuality.*
> > salvation
> > > English literature. short story. 1900-1999.
> > > > Forster, E. M. "Dr. Woolacott." symbolism.
> > > > > treatment of *salvation;* homosexuality.
> > symbolism
> > > English literature. short story. 1900-1999.
> > > > Forster, E. M. "Dr. Woolacott." *symbolism.*
> > > > > treatment of salvation; homosexuality.

NOTE: Faceted syntax should be accompanied with cross-references for linking synonymous, equivalent, broader, narrower, and other related terms. *See* section 6, Vocabulary.

7.2.5.3 Ad hoc coding

Some forms of string indexing require indexers to encode a natural language statement. The statement may be created to describe a document or may already exist as a text segment, such as a title. For example, one such system uses pointed brackets < > to enclose meaningful words or phrases that deserve direct entry; the question mark is used to introduce connectives, usually prepositions; and the symbol @ is used to turn off otherwise automatically generated headings. The following coded statement will result in the headings that follow:

> @effects? of <advertising>? on <sales? of <Japanese <cars>>? in <Germany>>
>
> advertising
> > effects on sales of Japanese cars in Germany
> cars
> > Japanese -. sales in Germany. effects of advertising
> Germany
> > sales of Japanese cars. effects of advertising
> Japanese cars
> > sales in Germany. effects of advertising
> sales
> > of Japanese cars in Germany. effects of advertising

NOTE: Ad hoc coding syntax should be accompanied with cross-references for linking synonymous, equivalent, broader, narrower, and other related terms. *See* section 6, Vocabulary.

7.2.5.4 Chain indexing

Chain indexing is based on the terms and the citation order of facets or aspects in a classification scheme. Chain indexing pre-dates algorithmic-based string indexing, but it shares format characteristics with string indexing. As classification schemes are converted to machine-readable form, algorithms can create, or assist in the creation, of chain indexes.

The chain index produces headings that complement the classification scheme by creating a string (or "chain") of terms from the classification heading but reversing the order in which facets or aspects are cited. The following headings from the Dewey Decimal Classification would produce the following chain index entries:

> 300 social sciences
> 330   economics
> 336   public finance
> 336.2   taxation
>
> economics 330
> public finance : economics 336
> social sciences 300
> taxation : public finance : economics 336.2

NOTE: Chain indexing syntax should be accompanied with cross-references for linking synonymous, equivalent, broader, narrower, and other related terms. *See* section 6, Vocabulary.

7.2.6 Syntactic cross-references

When syntax rules place terms only in secondary positions in headings or entries, cross-references should be used to provide direct access to such terms:

> United States — history — civil war — bibliography
> > [established heading]
> history — United States *see* United States — history
> civil war — United States *see* United States — history — civil war
> bibliography *see also particular topics with "bibliography" as subheading, for example:* United States — history — civil war — bibliography

7.3 Weighted terms

One function of an index is to discriminate between major and minor treatments of particular topics or manifestations of particular features (*see* section 3, Function of an index, item c). One method for achieving this goal is to assign weights to terms or to indicate major and minor terms by means of typography or symbols. Another method is to attach a subheading that indicates minor treatment, such as "also mentioned in" or "passing reference."

A weighting scheme is especially useful in high-exhaustivity indexing when minor or peripheral topics and features are indicated as well as the

major, central topics and features. Weighted terms permit the user who wants only major treatments to eliminate minor treatments, while permitting the user who wants every treatment to find them as well.

In full-text indexing, in which every instance of a term is often noted, term weighting can be used to help point to the more substantial treatments of topics.

In non-displayed indexes, weighted terms can be used as a basis for ranking retrieved records on the basis of estimated relevance.

See also 7.4.4, Methods of emphasizing locators in print indexes; and 7.5.2, Weighted term syntax.

7.4 Locators in displayed indexes

The purpose of a locator is to lead the user to the documentary unit or to a description of the documentary unit to which an index entry refers.

The nature of the locator depends on the medium and type of index and on the type of documentary units to which the index refers. In electronic indexes, index terms or headings may be linked to documentary units or to their surrogates without visible locators.

Locators should refer as directly and succinctly as possible to the documentary units to which index headings refer. (*See also* 5.7, Documentary units.)

7.4.1 Locators for printed documents

Printed books, pamphlets, periodicals, and similar documents normally consist of numbered pages bound into one or more volumes.

Pages are the traditional documentary units for indexes to printed books, pamphlets, and similar documents because pages are usually numbered while inherent textual or conceptual units, such as paragraphs, are usually not numbered. If pages are divided in some way, such as into columns, such smaller units may be used instead of or in addition to pages.

With certain classes of printed material, inherent textual units are often numbered and may therefore be used as locators. For example, parts of plays may be referred to by act, scene, and line number, and parts of books of the Bible by chapter and verse number. If documents have numbered paragraphs, then paragraphs rather than pages should be used as documentary units and paragraph numbers should be used as locators instead of or in addition to page numbers.

When a document consists of a series of uniquely numbered discrete units, such as abstracts, quotations, or case reports, these units' numbers are preferable as locators.

When there is more than one numbered sequence, they should be distinguished typographically:

Livingstone, Ken 1/3, 1/97, 3/94

or

Livingstone, Ken 1:3, 1:97, 3:94

When indexing several issues or volumes of a periodical or serial publication, locators should be based on the numbering of the issues at the time of publication. When documentary units are documents within a collection, for example, articles in a periodical, chapters in a monograph, or letters in an archive, sufficient information should be given to identify the document. For periodical articles, each locator normally consists of: author(s); title of article; title of periodical; volume, issue number, inclusive pagination, and complete date. The content, format, punctuation, and order of elements should conform to the ANSI/NISO Z39.29-1977, *Bibliographic References* (being revised at this writing, draft to be made available through NISO).

NOTE: The international standards ISO 690: 1987 (E) *— Documentation—Bibliographic references—content, form and structure*; and ISO 9115: 1987 (E)*—Documentation—Bibliographic identification (biblid) of contributions in serials and books* prescribe a different order of elements.

Abbreviation of names and titles should be avoided, especially if such names and titles may be searched electronically.

7.4.2 Locators for documents in other media

Documents in other media may, for indexing purposes, be divided into three types:

a. Those consisting of elements that form one or more sequences that are, or may be, continuously numbered and so accessed by the user. Such materials may be treated broadly as in 7.4.1. Examples are a collection of slides, a filmstrip, an audiodisc, or a machine-readable database. Locators would be slide numbers, frame numbers, side and band numbers, and record identifiers respectively.

b. Those consisting of one or more sequences of elements that cannot be distinguished numerically or so accessed by the user. Examples are serially accessed materials such as motion picture film and audio and video recordings. In these cases, relative locators should be devised, such as playing time from a particular point.

c. Those not consisting of sequences, such as maps, plans, charts, pictures, sculptures, and realia. In

some cases specific conventions exist, such as either grid references or coordinates for maps. In other cases, ad hoc locators should be devised.

NOTE: Most machine-readable text files fall into either category a or category b. Locators for such files may also take the form of internal (invisible) links between terms or headings and text segments.

7.4.3 Multiple locators in print indexes to single documents

If a subject is given continuous treatment in a consecutively numbered sequence, reference should be made to the first and last numbered elements only (for example, 3-11). The first and last element should be given in full in order to avoid ambiguity (for example, 20-25, 103-112, 1014-1027, not 20-5, 103-12, 1014-27). However, when locators are extremely long (5 digits or more) or where space is limited, unchanged initial numerals may be elided (for example, 100026-28).

Expressions such as "3ff" or "3 et seq." should not be used because they are confusing to most users and provide incomplete information.

If the treatment of a subject appears in a consecutively numbered sequence but consists of separate treatments (as opposed to a single, continuous treatment), individual locators for each numbered element should be used. For example:

3, 4, 5   [not: 3-5]

7.4.4 Methods of emphasizing locators in print indexes

In a closed-end index, if an entry array includes several locators, the locator leading to the fullest or most significant treatment may be emphasized typographically or by position. In open-end indexes, significant treatments may be indicated in similar ways, such as the use of a special symbol accompanying locators pointing to more comprehensive treatments.

Locators that relate to particular types of matter, such as tables and illustrations, may also be emphasized. Locators to illustrations, for example, may be italicized, enclosed in brackets, or prefixed or suffixed with an "i" or asterisk. Where more than one type of material is indicated, it is preferable to use the same system for all (for example, "i" for illustrations, "m" for maps, "t" for tables).

economics 144, 195, **229**	[major treatment emphasized by typography]
bibliographies 208, **244**, 363	
major university departments	
210-212, 211m, 212t	[map and table indicated]
economics 229; also 144, 195	[major treatment emphasized by location of locator]

See also 7.3 Weighted terms.

7.4.5 Presentation of locators

Locators should be clearly separated from headings by spacing, punctuation, or both, for example, by two spaces, by a comma or colon plus one space, or by a space colon space. The method used should depend on the nature of headings and the kind of punctuation used within headings. For example, headings that may end with commas and dates or other numerals should not use a comma plus space to introduce locators:

Paris, 1989 : 1934, 2045	[not Paris, 1989, 1934, 2045]
vitamin B-12: 13, 15	[not vitamin B-12 13, 15]

The method for presenting locators should be consistent throughout an index.

7.4.6 Presentation of other identifying data

Some indexes add information to citation locators indicating the presence of photographs, tables, and other illustrations or features. These indications, like index headings and subheadings, assist users in deciding whether documentary units are likely to be of value to them. They should be placed after the locator, separated from it by a period.

Doe, John. The indexing of pictures. *Journal of Indexing.* 2: 25-87; 1983. 4 illustrations (1 colored), 1 table, bibliography.

Abbreviations should be avoided because they are often confusing to users. Abbreviations also interfere with effective electronic searching when a print index is made available in electronic form. If abbreviations must be used in order to save space, they should be explained in an introductory note.

7.5 Syntax in non-displayed indexes

In non-displayed indexes, the search statement plays a role similar to that of the index heading in a displayed index, in the sense that the search statement may combine terms representing the topics and features and their aspects that a user is seeking.

Because search terms are combined after any initial indexing is done or in the absence of previous indexing, syntax in non-displayed indexes is often called "post-coordination syntax."

Two common approaches to searching post-coordinate indexes are the use of Boolean operators (AND, OR, NOT) and the use of weighted terms. Post-coordination syntax also includes proximity operators, stemming, and truncation. In addition, such syntactic devices as the use of links and role indicators may influence the application of post-coordinate syntax.

Compared to displayed indexing, non-displayed indexes are new, and additional methods for creating search statements are under development.

7.5.1 Boolean syntax

Boolean syntax combines terms using the operators AND, OR, and NOT. It has become a de facto standard for non-displayed indexes in electronic databases, but the meaning of AND and OR does not correspond to the usual senses of these words. Furthermore, the Boolean search divides a database into two distinct sets, retrieved and not retrieved. Retrieved documents are not ranked in any way on the basis of possible or probable interest. Documents that meet most but not all requirements of a search are not retrieved.

In most syntactic systems, the addition of terms serves to limit the scope of a search. The use of the Boolean OR has the opposite effect. It increases the scope of a search. It is often used to combine synonymous or equivalent terms.

NOTE: Boolean syntax should be accompanied with procedures for displaying synonymous, equivalent, broader, narrower, and other related terms in the indexing vocabulary (*see* section 6, Vocabulary).

7.5.2 Weighted term syntax

Searching by weighted term combination, also called "vector" or "probabilistic" searching, retrieves all documents that are represented by one or more of the search terms. Retrieved documents are then ranked by algorithms that attempt to predict relevance on the basis of such criteria as term frequencies in documents or surrogates, term frequencies in the database as a whole, the location of terms, user profiles, and user relevance judgments.

Both index terms and search terms may be weighted to reflect importance or interest. Such weights can further influence the calculation of "retrieval scores" for the purposes of ranking documents. Instead of dividing a database into two distinct sets (retrieved and not retrieved), weighted combination searching rearranges the entire database along a continuum of estimated degree of interest based on the search statement.

NOTE: Weighted term syntax should be accompanied with procedures for displaying synonymous, equivalent, broader, narrower, and other related terms in the indexing vocabulary (*see* section 6, Vocabulary).

7.5.3 Proximity operators, stemming, and truncation

Both Boolean and weighted term combination syn-

tax may be combined with a wide variety of methods for broadening or limiting the scope of a search statement. These methods include, but are not limited to, the use of: (a) proximity operators, which specify that two or more terms fall within a certain distance relative to each other; (b) stemming, which removes certain suffixes and/or prefixes; and (c) truncation, which permits the use of parts of words. Both stemming and truncation can be used to isolate word roots.

7.5.4 Links and role indicators

Links and role indicators are syntactic devices applied at the indexing stage that are designed to make post-coordinate searching more precise.

Links are used to indicate terms that may be logically linked to represent topics or features of the documentary unit. Linking eliminates false drops, the accidental retrieval of documentary units by the combination of terms that individually describe the documentary unit but that have no logical relationship. For example, the following terms for a documentary unit on the poetry of Thomas Hardy and the novels of E. M. Forster would be linked.

Forster, E. M. — novels
Hardy, Thomas — poetry

If a search is limited to linked terms, then the term combination "Hardy, Thomas — novels" would not retrieve this documentary unit.

Role indicators are used to specify the function of concepts represented by particular terms in particular documentary units, for example:

insulin — therapeutic use
insulin — product

These role indicators would prevent the retrieval of a documentary unit treating the manufacture or marketing of insulin as a product when its therapeutic use is wanted. Role indicators can consist of role terms, as in these examples, or of special notation.

8. Display of index arrays

In print media, individual index entries are displayed in ordered arrays, which provide the means of access to particular headings and entries. Therefore, the method of ordering entries is crucial. In electronic media indexes, entries may be sought by means of electronic matching without regard to index order. However, index displays in electronic media may suggest options for searching and permit browsing and scanning. Such electronic visual index displays also need to be arranged in helpful order. Entries retrieved by means of searching non-displayed electronic indexes are also displayed in

arrays after retrieval. These arrays, too, should be ordered according to useful criteria.

8.1 Introductory notes for indexes

If a displayed index is not straightforward or its conventions are not self-explanatory, an explanatory note should precede the index. Any abbreviations, symbols, or typographical conventions requiring explanation should be included in this note. In the case of separately published indexes, the introductory note should include sufficient bibliographic information (for example, author, title, publisher, place and date of publication or periodical volumes/issues) in order to completely identify the documents indexed (*see also* 5, Design of indexes).

8.2 Index display in print media

Procedures for displaying indexes in print media are well established, while appropriate means for visual displays in electronic media are still in the development stage.

The following sections relate primarily to indexes in print media, but some of the principles discussed are also applicable to visual arrays in electronic media.

8.2.1 Typography

The typography of printed indexes should result in clarity and easy legibility. The size of letters should not be smaller than 6 points.

Different typestyles (for example, bold, italics, or small capitals) may be used to distinguish entries for different types of documents, such as illustrations or titles of works. When an index consists of few main headings and many subheadings, the presentation of main headings in a typeface or style different from subheadings may be useful. Such conventions, when adopted, should be explained in an introductory note. Too much variety, however, may confuse the user.

A variety of other graphic features (boxes, shadings, icons) may be helpful to highlight and distinguish important features of indexes, such as cross-references, key entries, method of arrangement, and search options.

If more than one index is provided for the same document or collection of documents and separate indexes occupy more than two pages, each index should begin at the top of a page or column.

8.2.2 Arrangement of entries

Options for arranging index entries in visual arrays vary with respect to underlying structure and criteria. Structured arrays (as opposed to alphanumeric

arrays) can be helpful in breaking up large sequences into smaller, useful segments. Groupings can be created on the basis of relations among concepts (as in classification schemes) or the meaning or type of concept represented (for example, persons, places, other entities, texts). But structured arrangements can be detrimental to searching when the basis for their arrangement is hidden from and therefore unknown by users. For most index displays, direct and straightforward arrangement on the basis of commonly accepted ordering of alphanumeric characters is preferred, because most users cannot be expected to know less obvious principles for arrangement.

8.2.2.1 Alphanumeric displays

Alphanumeric displays are based on the commonly accepted sorting values of alphabetic letters and numerals. However, there are several options for how alphanumeric arrangement may actually be implemented. Some of these options, with recommendations, are considered in 9, Alphanumeric arrangement.

8.2.2.2 Classified or relational displays

In classified or relational displays, entries are arranged on the basis of relations among concepts represented by headings. Examples include superordination and subordination, class inclusion, chronology, and various types of roles and associations (for example, relations among discipline, action, object or agent of action, material, method, tools, and property). To the extent possible, the basis of arrangement should be made clear by summaries or outlines displayed at the head of the index. In almost all cases, a classified display should be accompanied by an alphabetic or alphanumeric index to the classification captions or headings and/or notation, unless the classified array is very short and can be quickly scanned.

8.2.3 Recurring elements

Indention should be used to avoid the repetition of recurring terms in subsequent headings:

labor
    distribution theory
    earnings
    monopolistic markets
    oligopolistic markets
    perfect competition

rather than

labor : distribution theory
labor : earnings
labor : monopolistic markets
labor : oligopolistic markets
labor : perfect competition

8.2.4 Vertical spacing

At least one blank line should separate major sections of an index, such as sections beginning with different letters in alphabetical indexes. In alphanumeric indexes, a blank line should also separate the non-alphabetical headings (for example, headings beginning with numerals) from the alphabetical sequence.

8.2.5 Entry layout

Entry layout will depend on a variety of factors, such as type of syntax used, length of entries, medium of display, and space available (*see* 5.12, Syntax; and 7.2, Syntax in displayed indexes).

When subheadings and sub-subheadings are used, they may be presented in an indented layout (also called "set-out," "line-by-line," or "entry-a-line"), a "run-in" layout (also called "paragraph style" or "run-on"), or a hybrid of the two styles.

Indented subheadings are preferable to run-in subheadings because users can scan them more quickly and can therefore understand them more easily. However, where economy dictates space-saving measures, run-in subheadings are preferable to shortening the index.

In all layout styles, all items on the same level of subdivision should be indented by the same amount of space (in the indented layout) or delineated by the same punctuation mark, such as a semicolon (in the run-in layout).

8.2.5.1 Indented layout

In the indented layout, each subheading and sub-subheading begins on a new line, progressively indented. All items on the same level of subdivision should be indented by the same amount of space:

```
Aristotle
    debt to Plato 23, 26
    literary criticism in 35, 74, 89-93, 101-197
        on Aeschylus 101-104, 279
        on Aristophanes 195
        on Euripides 104-126, 187, 265-266
        on Homer 103, 190-194, 206
        on Sophocles 127-183, 275-277, 306, 309-310
            Antigone 155
            Oedipus Tyrannus 140-149
    origins of tragedy
        in epic 196
        in revelry 197
```

In the indented layout, a turnover line should be indented more deeply than the indention of a subheading in the same entry. For example:

```
periodicals
    author indexes 276-277
    defined 272
    number of volumes and
        parts 224                    [turnover line]
    titles 152, 224, 331
        abbreviations  225-226, 389
```

8.2.5.2 Run-in layout

The run-in layout should be limited to two levels of heading (for example, main heading and subheadings). If three or more levels are used, the indented layout of subheadings under the main heading should be employed, with the run-in layout being used only for sub-subheadings and further levels of subdivision, as in the hybrid example in 8.2.5.3, Hybrid indented/run-in layout.

Run-in layout (limited to 2 levels):
```
Aristotle 20-22; debt to Plato 23, 26; literary criticism in
    35, 74, 89-93, 101-197; origins of tragedy 196, 197
```

All items on the same level of subdivision should be delineated by the same punctuation mark, such as a semicolon. When there are no locators between headings at two different levels, the two levels should be separated by a colon (*see* "origins of tragedy" in the hybrid example in 8.2.5.3).

8.2.5.3 Hybrid indented/run-in layout

When more than two levels of heading are used (for example, main heading, subheading, sub-subheading), the run-in layout may be combined with the indented layout, the indented layout being used for main and subheadings, the run-in layout being used for sub-subheadings:

```
Aristotle
    debt to Plato 23, 26
    literary criticism in 35, 74, 89-93, 101-197; on Aeschylus
        101-104, 279; on Aristophanes 195; on Euripides
        104-126, 187, 265-266; on Homer 103, 190-194, 206;
        on Sophocles 127-183, 275-277, 306, 309-310
    origins of tragedy: in epic 196; in revelry 197
```

In the hybrid layout, whenever a line "turns over" to the next line, all lines after the first line should be indented more deeply than the deepest subheading indention employed in the index (*see* "literary criticism" in the hybrid example above).

8.2.6 Running headlines

Pages on which an index is printed should bear a running headline or footline. In the case of multiple indexes, these should appear on each page indicating an appropriate title for each index. In the case of separately issued indexes, the words "Index to [title of work]" should be used. A running headline at the top of a page should be differentiated from the guidewords that may appear at the left and right margins of a two-page spread.

8.2.7 Guidewords

Guidewords (also called "catchwords") may be used to indicate the scope of entries on a two-page spread, reproducing all or part of the first and last

heading. Guidewords are positioned at the upper left margin on a verso (left-hand page) and at the upper right margin on a recto (right-hand page).

8.2.8 Columns

A printed index is normally displayed in two columns per page. In large-size documents, it may be set in three or four columns. Indexes in which entries are long (for example an index of first lines or a table of cases in legal works) are better set to full page width.

The width of a column should not be less than 2 inches in order to avoid turnover lines as much as possible.

One line should accommodate an index heading of average length, followed by at least two locators. When an entry or entry array occupies more than one line, consecutive locators should never be divided:

```
periodicals
    titles 152, 224, 331
        abbreviations 225-226
not:
periodicals
    titles 152, 224, 331
        abbreviations 225-
            226
```

Three columns on a page of normal width (5-1/2–6 inches including margins) is not recommended because this may result in many turnover lines, which make the index more difficult to scan, while not saving space.

In a long index, where groups of headings beginning with a new initial letter occupy more than 4 pages, each letter group may begin on a new column or on a new page.

8.2.9 Continuation lines

In typesetting an index into pages or columns, some entries or entry arrays will be continued from the bottom of one column or page to the top of the next column or page.

The continuation of very short parts of entries or of entry arrays from one column or page to the next should be avoided. Examples are one or two locators or the final line of an alphanumeric section of the index. Similarly, the initial line of an entry that begins a new alphanumeric sequence should not fall at the bottom of a column or page.

When an index entry or entry array runs on to a new column or page, the index heading and any subheading and sub-subheading applicable to the run-on entries should be repeated, followed by "(continued)" or the abbreviation "(cont.)" after each level of heading:

On bottom of column or page:
```
thesauri
    adaptation 182
    construction 353, 364
        software 387
```

On top of next column or page:
```
thesauri (cont.)
    construction (cont.)
        standards 374
    defined 381
```

8.3 Index display in electronic media

When electronic indexes provide options for the display of index entries for scanning and browsing, or for the display of retrieved records for review, options for format and fullness of entries or records and for the ordering of arrays should be clearly described on display screens and in documentation.

The design of effective electronic displays is the topic of much research and experimentation. Combinations of pictorial, tabular, graphic, animated, and verbal elements may be useful. It is too soon to suggest guidelines for such displays.

8.3.1 Browsable index displays

Because the viewing area (screen) in electronic media is usually small and constrained and the level of visual resolution is limited compared to print media, it is usually helpful to display entries in stages for scanning and browsing. For example, when entries consist of main headings and subheadings, the initial display may be limited to main headings. When a main heading is selected, for example, by highlighting, then the subheadings linked to that main heading can be displayed. When a subheading is selected, sub-subheadings, locators, document citations, or other surrogates can be displayed.

At all stages of display, captions, headings, or prompts should clearly show users where they are in the index and how they may move around in the index.

8.3.2 Displays of retrieved records

In the display of retrieved records (surrogates or items), users should have options for the fullness of display ranging from brief (for example, title and author only) to full (for example, title, author, full citation, abstract). Users should also have options for the arrangement of retrieved records (for example, reverse chronological order, ranked according to potential relevance, classified by facets or a classification scheme, or ordered alphanumerically by index terms or headings or by citation elements such as authors, titles, publishers, or dates).

8.4 Electronic manuscripts

When indexes designed for print publication are transmitted via electronic media, typographic coding should conform to ANSI/NISO/ISO 12083, *Electronic manuscript preparation and markup*.

9. Alphanumeric arrangement

9.1 Standards

Two de facto standards widely used in libraries and databases in the United States are the American Library Association (ALA) and the Library of Congress (LC) filing rules. The guidelines for alphanumeric arrangement in *The Chicago Manual of Style* are used as a de facto standard by many publishers. These three codes for alphanumeric arrangement are, however, incompatible with each other. The guidelines presented here are based primarily on the *ALA Filing Rules*.

These guidelines recommend that for English language indexes, alphanumeric order be as simple and straightforward as possible, based only on characters that have widely accepted and widely known sorting order: the 26 letters of the English alphabet, the 10 Arabic numerals, and the space. Exceptions, such as special arrangement rules for punctuation or special rules based on the nature of a heading or the entity represented by the heading, are not widely understood and may cause confusion. (*See also* NISO Z39.75-199x, *Alphabetical arrangement of letters and the sorting of letters and other symbols*, in development at this writing.)

9.2 Basic order

The basic order of characters is:

a. spaces, punctuation marks, and symbols other than letters and numerals: All characters in this group have equal sorting value and are placed before any numeral or alphabetic letter. All are treated as if they were a space. Multiple consecutive spaces and their equivalents are to be considered equal to a single space.

b. numerals (0 through 9): Numbers expressed in numerals are arranged according to ascending arithmetical value (*see also* 9.8, Numerals).

c. letters (A through Z): Lowercase and uppercase letters have equal sorting value. Modified letters are treated like their basic equivalents in the English alphabet. For example, ø, ò, ô, ó, ö, all have the same sorting value as the unmodified o.

If nonalphanumeric symbols are prominently featured and must be sorted, a system for their arrangement should be devised and explained, because no standards exist for the arrangement of nonalphanumeric symbols.

Optionally, the ampersand (&) may be arranged as its spelled-out language equivalent (for example: and, et, und, y).

9.3 Word-by-word versus letter-by-letter arrangement

In accordance with the basic order specified in section 9.2, index headings should be arranged by the word-by-word method, in which a space comes before a letter or numeral. This order, which is standard in library catalogs, will cause headings beginning with the same word to be placed next to each other:

N.Y.S.E. constitution and rules	[title]
New, Agnes	
New Brunswick	
new journalism	
new moon	
New, Thomas	
New York	
Newark	
Newfoundland	
news	
news agencies	
news (journalism)	
news-letters *see* newsletters	
news photography	
newsletters	
newspapers	
NEXIS (information retrieval system)	

An alternative arrangement, letter-by-letter, disregards the space and symbols that have the same sorting value as a space. This arrangement may be required for the continuation of an existing index, but is not recommended:

New, Agnes	
Newark	
New Brunswick	
Newfoundland	
new journalism	
new moon	
news	
news agencies	
news (journalism)	
newsletters	
newspapers	
news photography	
New, Thomas	
New York	
NEXIS (information retrieval system)	
N.Y.S.E. constitution and rules	[title]

9.4 Initial articles

Initial articles in the nominative case are ignored for arrangement purposes at the beginning of titles, first lines, topical subject terms, and names of corporate bodies:

Der Blaue Adler (association)
The Club (London)
The H. W. Wilson Company
The Library Association (United Kingdom)
The Movement (English poetry)   [subject term]
The nutcracker (ballet)
Les temps   [title]
An unlikely one will guide me   [first line]
Der Zauberberg   [title]

Initial articles that form an integral part of place name and personal name headings should be considered for arrangement purposes:

El Paso
Le Guin, Ursula K.
Los Angeles
The Dalles

Cross-references should link forms arranged under the part following the initial article with their preferred location under the initial article:

Angeles, Los *see* Los Angeles
Dalles, The *see* The Dalles
Guin, Ursula K. *see* Le Guin, Ursula K.
Paso, El *see* El Paso

NOTE: The arrangement of initial articles may need to be explained in an introductory note.

See also 6.2.9.1, Personal names; 6.2.9.2, Corporate body names; 6.2.9.3, Geographical names; 6.2.9.4, Titles of documents; 6.2.9.5, First lines.

9.5 Subheadings

Subheadings are normally arranged in the same way as headings. Initial prepositions in subheadings should not be ignored:

computers
    compared with abacus
    for management
    in hospitals
    management of

management
    of computers
    use of computers

However, the arrangement of subheadings may be modified by chronological or some other systematic arrangement if such an arrangement is considered helpful to users and can be clearly understood by them:

Music
    history and criticism
        to 400 A.D.
        medieval, 400-1500
        20th century

It is usually better to use numerical headings for chronological topics:

Music
    history and criticism
        to 400 A.D.
        400-1500
        1901-2000

9.6 Headings with the same initial term

Headings beginning with the same term should be arranged in the following sequence:

a. term alone, with or without subheadings.

b. term with qualifier or a term as the lead term of a longer heading. These should be arranged together according to the sorting value of the characters following the lead term:

songs
    bibliography
    history and criticism
    texts
songs, American *see* American songs
Songs and poems   [title]
songs, Cajun *see* Cajun songs
songs (high voice) with piano
Songs just for children   [title]
songs (low voice)
songs (medium voice) with guitar
songs, Zionist *see* Zionist songs
songwriters *see* composers; lyricists

Some indexes arrange headings beginning with the same word on the basis of the nature of the heading or the nature of the item represented (for example, forename or surname; person, corporate body, place name, topic, or title; type of subdivision) rather than the alphanumeric characters in the remainder of the heading. Such arrangements can help to break up long sequences of entries by grouping headings of similar type, but they can also confuse users who do not understand the basis of the arrangement. Whenever headings are arranged on the basis of nonalphanumeric criteria, these criteria should be prominently stated each time they are used.

9.7 Cross-references

A cross-reference introduced by "see" or "see also" or analogous linking terms is not part of a heading and does not affect the position of the heading in an alphabetical sequence. For example:

songwriters *see* composers; lyricists
Songwriters and composers on Broadway   [title]

9.8 Numerals

Headings beginning with numerals should be arranged according to ascending arithmetical value before the alphabetical sequence.

Roman numerals are arranged with their Arabic counterparts (not by their constituent letters). For example, "XIX" is placed before or after "19," depending on the characters (numerals, letters, or spaces) following the numbers:

.75 acres are for sale
3/4 for 3

1:00 a.m.
1.3 acres
2-1/2 minute talk treasury
3 and 30 watchbirds
$6.41 per hen per year
007 James Bond: a report
10% review
XX century cyclopedia and atlas
20 miles down the road
21-8-1968: anno humanitatis
49th parallel
1001 nights
1066 and all that
1984 *see* Nineteen eighty-four

Where numerals occur within headings or sub-headings, they should be arranged according to ascending arithmetical value:

Club 18-30
Club 21
Club 147 fashions
Club one holidays

Numerals as prefixes or infixes in names of chemical compounds in biomedical and chemical texts may be disregarded, unless needed to distinguish homographs:

ethyl
3-ethyl-4-picoline
4-ethyl-*alpha*-picoline
ethylene

9.9 Comprehensive example

The following alphanumeric array is designed to illustrate all of the arrangement situations described in the previous parts of section 9:

.75 acres are for sale
3/4 for 3
1:00 a.m.
1.3 acres
2-1/2 minute talk treasury
3 and 30 watchbirds
$6.41 per hen per year
007 James Bond: a report
10% review
XX century cyclopedia and atlas
20 miles down the road
21-8-1968: anno humanitatis
49th parallel
1001 nights
1066 and all that
1984 *see* Nineteen eighty-four

American songs
Angeles, Los *see* Los Angeles

Der Blaue Adler (association)

Cajun songs
Charles
Charles I, King of England
Charles II, Emperor of Germany
Charles II, King of France
Charles III, King of England
Charles III, King of England   [title]

Charles (airplane)
Charles (AL). Police Department.
Charles, Allen   [surname entry]
Charles and the wise men   [title]
Charles, Duke of York
Charles, Prince of Wales
Charles, Saint
Charles (VA). Municipal Court.
Charles, Virginia   [surname entry]
Charles (yacht)
Club 18-30
Club 21
Club 147 fashions
The Club (London)
Club one holidays
clubs
composers
computers
    compared with abacus
    for management
    in hospitals
    management of

Dalles, The *see* The Dalles
Dalles (The) Public Library *see* The Dalles Public Library

El Paso
El Paso. Police Department
The Extended Simulation Support System *see* TESS
        (computer system)

The H. W. Wilson Company

Le Guin, Ursula K.
libraries
The Library Association (United Kingdom)
Los Angeles
lyricists

management
    of computers
    use of computers
The Movement (English poetry)
music
    Africa
    biography
    cataloging
    dictionaries
    history and criticism
        to 400 A.D.
        400-1500
        1901-2000
        methods
        outlines, syllabi, etc.
    Peru
    United States
Music about the house   [title]
music, African *see* African music
music and architecture
music, baroque *see* baroque music
Music (MS). Park Department
music, Roman *see* Roman music
Music, Valerie   [surname entry]

N.Y.S.E. constitution and rules   [title]
New, Agnes
New Brunswick
new journalism
new moon
New, Thomas
New York
Newark

Newfoundland
news
news agencies
news (journalism)
news-letters *see* newsletters
news photography
newsletters
newspapers
NEXIS (information retrieval system)
The nutcracker (ballet)

Paso, El *see* El Paso
poetry

songs
    bibliography
    history and criticism
    texts
songs, American *see* American songs
Songs and poems   [title]
songs, Cajun *see* Cajun songs
songs (high voice) with piano
Songs just for children   [title]
songs (low voice)
songs (medium voice) with guitar
songs, Zionist *see* Zionist songs
songwriters *see* composers; lyricists
Songwriters and composers on Broadway   [title]

Les Temps   [title]
TESS (computer system)
The Dalles
The Dalles Public Library
The Extended Simulation Support System *see* TESS
        (computer system)

An unlikely one will guide me   [first line]

Der Zauberberg   [title]
Zionist songs

10. References

The following American National Standards Institute/National Information Standards Organization (ANSI/NISO) standards contain provisions related to this technical report. At the time of publication, the editions indicated were valid. Because all standards are subject to revision, users of this technical report are encouraged to investigate the possibility of applying the most recent editions of these standards.

ANSI/NISO Z39.19-1993. *Guidelines for the construction, format and management of monolingual thesauri.* Bethesda: NISO Press, 1994.

ANSI Z39.29-1979. *Bibliographic references.* In revision.

NISO Z39.75-199X. *Alphabetical arrangement of letters and the sorting of numerals and other symbols.* Bethesda: NISO Press. In development.

ANSI/NISO/ISO 12083. *Electronic manuscript preparation and markup.* Bethesda: NISO Press, 1996.

The following International Organization for Standardization (ISO) standards and drafts are cited:

ISO 690: 1987 (E) — *Documentation — Bibliographic references — content, form and structure.*

ISO 9115: 1987 (E) — *Documentation — Bibliographic identification (biblid) of contributions in serials and books.*

ISO 999:1996 — *Information and Documentation — Guidelines for the content, organization and presentation of indexes.*

The following rules for alphanumeric arrangement from the American Library Association and the Library of Congress function as de facto standards in the United States, but they are incompatible with each other. Rules in this technical report are closest to the *ALA Filing Rules.*

American Library Association, Filing Committee. *ALA filing rules.* Chicago: American Library Association, 1980.

Library of Congress, Processing Services. *Library of Congress filing rules.* Prepared by John C. Rather and Susan C. Biebel. Washington: Library of Congress, 1980.

The de facto library standard for the formulation of name headings, both personal and corporate, is the following:

Anglo-American cataloguing rules, 2nd edition, 1988 revision. Prepared by the Joint Steering Committee for Revision of AACR; Michael Gorman and Paul W. Winkler, eds. Chicago: American Library Association, 1988.

For romanization, use: *ALA-LC romanization tables: Transliteration schemes for non-Roman Scripts.* Randall K. Barry, ed. Washington: Library of Congress, 1991.

Many indexers and publishers follow guidelines in *Chicago manual of style.* 14th ed. Chicago: University of Chicago Press, 1993.

11. Bibliography

This technical report assumes basic understanding of indexing and indexes. The following publications will be helpful in providing background information. They are arranged in inverse chronological order.

1996. Foskett, A.C. (Antony Charles). *The subject approach to information.* 5th ed. London: Library Association.

1996. O'Connor, Brian C. *Explorations in indexing and abstracting: pointing, virtue, and power.* Englewood, CO: Libraries Unlimited.

1996. Wellisch, Hans H. *Indexing from A to Z.* 2nd ed., rev. and enl. New York: H.W. Wilson.

1995. Fetters, Linda K. *A guide to indexing software.* 5th ed. Seattle, WA: American Society of Indexers.

1994. *Challenges in indexing electronic text and images.* Raya Fidel, Trudi Bellardo Hahn, Edie M. Rasmussen, and Philip J. Smith, eds. Medford, NJ: Published for the American Society for Information Science by Learned Information, Inc.

1994. Mulvany, Nancy C. *Indexing books.* Chicago: University of Chicago Press.

1992. Bell, Hazel K. *Indexing biographies and other stories of human lives.* London: Society of Indexers.

1991. Bellardo, Trudi. *Subject indexing: an introductory guide.* Washington, DC: Special Libraries Association.

1991. *Index evaluation checklist: a guide for authors, editors, publishers, reviewers, librarians.* Port Aransas, TX: American Society of Indexers.

1991. Lancaster, F. W. *Indexing and abstracting in theory and practice.* Champaign, IL: Graduate School of Library and Information Science, University of Illinois.

1989. *Indexing: the state of our knowledge and the state of our ignorance.* The proceedings of the 20th annual meeting, American Society of Indexers. Bella Hass Weinberg, ed. Medford, NJ: Learned Information.

1989. Salton, Gerard. *Automatic text processing: the transformation, analysis and retrieval of information by computer.* Reading, MA: Addison-Wesley.

1988. Rowley, Jennifer E. *Abstracting and indexing.* 2nd ed. London: Bingley.

1986. Craven, Timothy C. *String indexing.* Orlando, FL: Academic Press.

1985. Soergel, Dagobert. *Organizing information: principles of data base and retrieval systems.* Orlando, FL: Academic Press.

1984. Milstead, Jessica L. *Subject access systems: alternatives in design.* Orlando, FL: Academic Press.

1979. Knight, G. N. *Indexing, the art of.* London: Allen & Unwin.

1978. Borko, Harold; Bernier, Charles L. *Indexing concepts and methods.* Orlando, FL: Academic Press.

1975. *UNISIST: indexing principles.* Paris: UNESCO.

12. Glossary

This glossary provides definitions for terms as they are used in the context of this technical report. Within definitions, terms that have their own definitions, either in this glossary or in section 2 (Definitions) of the report, are in italics. Defined terms are listed in alphabetical order in the singular noun form; however, within other definitions, corresponding terms may appear as plural nouns, adjectives, or other forms.

adjacency operator. *See* proximity operator.

alphanumeric display. An *index* display in which *headings* are arranged in alphanumeric order. Contrast with *relational display.*

arrangement. The ordering of *entries* in an *index* in alphabetical, numerical, or other stated and consistent order. Also called filing, sorting.

array. A displayed sequence of *terms, headings* or *entries. See also* entry array; file.

articulated heading. *See* multi-level heading.

assignment indexing. An *indexing* method by which *terms, descriptors,* or *subject headings* are selected by a human or computer to represent the *topics* or *features* of a *documentary unit.* Assigned terms may or may not occur in the document. *See also* derivative indexing.

associative relationship. A non-hierarchical relationship among *terms* that are conceptually or semantically linked, for example, "cooking" and "food." Terms having an associative relationship are called *"related terms."*

authority file. A set of records of established *descriptors* or *headings* and the *cross-references* to be made to and from them, often citing the authority for the preferred form and the source of variants. Types of authority files include name authority files, subject authority files, and *thesauri. See also* descriptor; vocabulary control; subject heading list.

Boolean operators. The logical operators AND, OR, and NOT, which can be used to combine *terms* for searching in post-coordinate *information retrieval systems. See also* post-coordination.

bound term. *See* compound term.

broader term. A *term* to which another term or multiple terms are semantically subordinate in a *hierarchy*. *See also* narrower term; related term.

catchword. *See* guideword; keyword.

chain indexing. The creation of *multi-level headings* that consist of "chains" of *terms* extracted from a *classification* scheme, arranged in an inverse *citation order* of *facets* to that of the classification scheme itself.

citation order. The order in which *facets* are arranged (cited) in a classified *array*; also the order in which *terms* from facets are placed in a *multi-level heading*.

class. A set whose members share an attribute, characteristic, property, quality, or trait.

classification. The operation of grouping *concepts* or *entities* into *classes* and establishing relations among these classes. *Headings* representing classes are usually arranged in *arrays* that illustrate relations among *classes*, creating a classified *index*, as opposed to an alphanumeric index. *See also* relational display.

classified display. *See* relational display.

closed-end index. An *index* compiled at one time for one or more *documents*. Contrast with *open-end index. See also* monographic index.

compound term. A *term* for a *concept* consisting of more than one word, for example, "compass rose," "first aid," "trade winds"; also a multi-word term representing multiple concepts that are so often considered together that representing them with separate terms would be unwieldy and may result in *false drops*, for example, "science information," "information science." Also called "bound term."

concept. A unit of thought, formed by mentally combining some or all of the characteristics of a concrete or abstract, real or imaginary object, attribute, material, process, operation, event, place, or time. Concepts exist in the mind as abstract entities independent of *terms* used to represent them. *See also* topic; feature.

controlled vocabulary. A subset of the lexicon of a *natural language*. A list of *terms* that may be used for *indexing*, produced by the operation of *vocabulary control*. Controlled vocabularies are usually recorded in *subject heading lists* or *thesauri*.

cross-reference. A linking device between two or more *terms* or *headings* in an *index*. There are three types of relationships among terms that require cross-references: (a) an equivalence relationship among *synonymous* or *equivalent terms* or headings, (b) an *associative relationship,* indicating an unspecified relationship among terms or headings (called *related terms* or headings), and (c) a *hierarchical* relationship, indicating a broader/narrower relationship among terms or headings. *See also* broader term; narrower term.

depth of indexing. The result of the combined effects of *exhaustivity* and *specificity* in an *index*.

derivative indexing. An *indexing* method by which words occurring in the title or *text* of a *documentary unit* are extracted by a human or computer to serve as indexing *terms*. Also called extractive indexing. *See also* assignment indexing.

descriptor. A *term* chosen as the preferred representation for a *concept* or *feature* in an *index. See also* subject heading.

difference. *See* modifier.

display. *See* alphanumeric display; relational display.

displayed index. An *index* that may be searched by means of visual inspection. *See also* non-displayed index.

document. A *medium* on or in which a *message* is encoded; thus, the combination of message and medium. The term applies not only to written and printed materials on paper or microforms (for example, books, journals, maps, diagrams), but also to nonprint media (for example, machine-readable records, transparencies, audio recordings, video recordings, film) and, by extension, to natural or humanly made objects intended to convey *information*. Documents encompass every kind of format and genre, including but not limited to treatises, literary works, patents, technical reports, charts, tables, illustrations, music, artistic works, and multimedia texts.

document-oriented indexing. *See* entity-oriented indexing.

documentary unit. The *document*, document

segment, or collection of documents to which *entries* in an index refer and on which they are based. Examples of verbal documentary units include sentences, paragraphs, pages, complete articles, books, complete serial runs, collections of archival materials, microform sets, and entire library collections. The documentary unit determines the relative size of document, document segment, or collection of documents to which an *index* will point.

domain. The territory covered in order to locate *documents* for the purpose of producing a bibliography, database, or *index* to multiple documents, ranging, for example, from a single collection to one or more countries or the entire world. When the domain is limited to a single collection, the resulting index is usually called a catalog.

entity-oriented indexing. *Indexing* based entirely or primarily on the *topics* and *features* of *documentary units* rather than on the anticipated needs and requests of users. Also called document-oriented indexing. *See also* request-oriented indexing.

entry. The representation of a *documentary unit* in a *displayed index*. It consists of at least a *heading* and a *locator*. More than one *locator* may follow a given heading in a displayed *entry array*, but each locator, in combination with its heading, represents a single entry. An entry may contain a *multi-level heading* and a document *surrogate* in addition to the required locator.

entry-a-line layout. *See* indented layout.

entry array. A sequence of *entries* sharing the same *heading* in a *displayed index*.

entry term. The first *term* in an *entry*, to which direct access is provided. *See also* heading; lead-in term; lead term.

entry vocabulary. All *terms* by which access may be gained to an *index*, including both those that lead to *documentary units* and those from which *cross-references* lead to other terms that are used in their place. *See also* lead-in term.

equivalent term. A *term* that is used for another term in the context of an *index*. Includes *synonyms*, but also *broader* or *related terms* when they are used for another term.

exhaustivity. The average number of *terms* assigned

to a *documentary unit* in a particular *index* or retrieval system.

extractive indexing. *See* derivative indexing.

facet. Fundamental or important aspects of a *topic*, such as Ranganathan's "personality, material, energy [i.e., activity, operation, process], space, and time." In literature, for example, facets may represent such aspects as language, nationality, genre, period, theme, writer, etc.

faceted indexing. The assignment of *terms* to *facet* categories and the ordering of terms within *headings* in accordance with a *citation order* of facets.

false drop. An irrelevant reference retrieved when terms match, but meanings are unrelated because of *homography* ("fans" retrieving references to rotating blades when devotees are intended), or a confusing combination of terms ("school" and "library" retrieving references on library schools, when school libraries are intended).

feature. An aspect of a *document* other than a *topic*. Features include such aspects as authorship, style, methodology, quality, usefulness, level of complexity, language, format, publication date, etc.

file. A sequence or *array* of two or more *entries* or *records* in an *index* or an *information retrieval system*. In the context of computers, the term is now also used for a computer-readable text.

filing. *See* arrangement.

free-text term. A *natural language term* appearing in *documents* or their descriptions that may be used in searching. *See also* keyword.

generic posting. The assignment of a *broader term* instead of a specific *term*, for example, using "furniture" to index a *documentary unit* on sofas. *See also* equivalent term; up-posting.

guideword. A *term* or *heading* placed at the top left of a verso (left-hand) page and the top right of a recto (right-hand) page to indicate the scope of *entries* on a two-page spread in a printed *index*. Also called "catchword" and "scope headline."

heading. One or more *terms* representing a *topic* or *feature* of a *document* in a *displayed index*; the first element of an index *entry* in a displayed index. *See*

also main heading; multi-level heading; subheading; sub-subheading; subject heading.

hierarchy. A system of ranked *terms* in which a superordinate or higher term is broader in semantic scope than a subordinate or lower term. Hierarchical *arrays* display *narrower terms* under *broader terms*.

homograph. *Terms* that have the same spelling, but different meanings, such as "race (anthropology)," "race (sports)." Homographs should be distinguished by *qualifiers*.

identifier. A proper name (or its abbreviation) of a person, institution, place, object, operation, or process, optionally treated as a type of *term* distinct from *descriptor*. Identifiers may be held in a separate *file*, such as an *authority file*, and their form may be controlled (for example, the name of an international organization having different names in various languages, only one of which is selected as an authorized term or descriptor).

indented layout. The display of *multi-level headings* in an *entry array* in which each new *subheading* and *sub-subheading* begins on a new line, progressively shifted to the right under a *main heading*. Also known as "entry-a-line," "line-by-line" or "set-out layout."

index. A systematic guide designed to indicate *topics* or *features* of *documents* in order to facilitate retrieval of documents or parts of documents. Indexes include the following major components: (a) *terms* representing the topics or features of *documentary units*; (b) a *syntax* for combining terms into *headings* (in *displayed indexes*) or *search statements* (in *non-displayed indexes*) in order to represent compound or complex topics, features, and/or queries; (c) *cross-references* or other linking devices among *synonymous, equivalent, broader, narrower*, and other *related terms*; (d) a procedure for linking headings (in displayed indexes) or search statements (in non-displayed indexes) with particular documentary units or document *surrogates;* and (e) a systematic ordering of headings (in displayed indexes) or a search procedure (in non-displayed indexes).

indexable matter. The portions of *documents* that are actually analyzed and *indexed. See* section 5.8. For topics and features to be indexed, *see* section 5.1, Subject scope.

indexing. The operation of creating an *index* for

information retrieval. Indexing involves (a) the selection and assignment of *terms* to, or the extraction of terms from, a *documentary unit* in order to indicate *topics, features*, or possible uses of the unit; (b) the combination of terms into *headings* or the tagging of terms for subsequent combination (in *displayed indexes*); (c) the linking of *synonymous, equivalent, broader, narrower*, and other *related terms* or headings; (d) the linking of terms or headings to documentary units or *surrogates;* and (e) the *arrangement* of headings in a systematic order (in displayed indexes).

indexing language. In a broad sense, any vocabulary, including uncontrolled vocabulary, used for *indexing* and the rules of *syntax* for its application. In a narrower sense, a *controlled vocabulary* or *classification* system and the rules of syntax for its application. An indexing language is used for the representation of *topics* and *features* of a *documentary unit* and for the retrieval of *documentary units* from an *information retrieval system*.

information. Refers both to an entity (for example, a *message* recorded in a *text* and represented in a *document*) and to the process of informing or becoming informed. What constitutes an informative message and successful information (as process) is subjective. Preferably, terms like "message," "text," or "document" should be used when referring to potentially informative entities.

information retrieval system. A set of operations and the associated equipment, procedures, algorithms, and documentation by which *documentary units* are *indexed* and the resulting *records* are stored and displayed, so that selected records (and/or the documentary units they represent) can be retrieved.

keyword. A word occurring in the *natural language* of a *document* or its *surrogate* that is considered significant for *indexing* and retrieval. Any word not on a *stop list* contained in a verbal segment of a document or assigned to a *documentary unit*, such as, title, abstract, *subject headings*. Used as *lead terms* in keyword indexes such as *KWIC (KeyWord-In-Context), KWOC (KeyWord-Out-of-Context)*, and *KWAC (KeyWord-Alongside-Context)* indexes. *See also* free-text term.

KWAC (KeyWord-Alongside-Context) index. An *index* in which each significant word in a string of *text* serves as *lead term* or access point, followed by the portion of the string that follows

the word, then by the portion of the string that precedes the word.

KWIC (KeyWord-In-Context) index. An *index* in which each significant word in a string of *text* serves as *lead term* or access point, by being graphically emphasized and surrounded by the rest of the string. The lead terms or access points are arranged in a column in the middle of the *entries* rather than at the left.

KWOC (KeyWord-Out-of-Context) index. An *index* in which each significant word in a string of *text* serves as *lead term* or access point, followed by the complete string. Multi-word terms that include the lead term are not preserved, because the lead term is always followed by the first word of the string.

lead-in term. A *term* that is not used as an *index* term but which leads to one or more index terms (preferred terms or *descriptors*) by means of a *cross-reference*.

lead term. The first *term* in a *heading*. Distinguished from *"lead-in term." See also* entry term.

line-by-line layout. *See* indented layout.

link. In *indexing* for *post-coordinate* electronic searching, a *syntactic* device used to indicate *terms* that may be logically combined to represent *topics* or *features* of a *documentary unit* in order to prevent *false drops. See also* role indicator.

literary warrant. Justification for the representation of a *concept* in an *indexing language* or for the selection of a preferred term or *descriptor* because of its frequent occurrence in *documents*.

locator. The part of an *entry* in a *displayed index* that indicates the location of the *documentary unit* to which the entry refers. Locators range from brief notations, such as page numbers, to full bibliographic citations.

main heading. The first heading in a *multi-level heading*, which is followed by a *subheading*.

medium, media (pl.). The physical entity on or in which a *message* is recorded. A medium and a message recorded in or on it constitute a *document*.

message. *Concepts* conveyed by the *text* of a *document*.

mission-oriented index. An *index* that focuses on particular *topics* or *features* of *documents* that relate to a particular mission or objective, ignoring unrelated topics or features. *See also* request-oriented indexing.

modifier. In a *compound term*, one or more components that serve to narrow the extension of a focus and specify one of its subclasses. Also known as "difference." *See also* subheading.

monographic index. An *index* compiled for a single *document. See also* closed-end index; serial index.

multi-level heading. A *heading* consisting of a *main heading* that is modified by a *subheading*. The subheading may in turn be modified by a *sub-subheading* and possibly by additional headings at successive levels of subordination.

narrower term. A *term* that is subordinate to another term in a *hierarchy. See also* broader term; related term.

natural language. A language used by human beings for verbal communication. Words extracted from natural language *texts* for *indexing* purposes are often called *keywords. See also* free-text term.

near-synonym. *See* quasi-synonym.

non-displayed index. An *index* that is searched by means of electronic comparison and matching controlled by computer algorithms. The complete index itself is not displayed for searching by means of visual inspection.

nonpreferred term. One of two or more *synonyms*, lexical variants or *equivalent terms* that serves as a *lead-in term*. A nonpreferred term should be linked to a preferred term *(descriptor)* or *heading* by means of a *cross-reference* or other linking device.

open-end index. An ongoing *index* compiled at set intervals or continuously updated. *See also* serial index; closed-end index.

paragraph layout. *See* run-in layout.

permuted index. The representation of *terms* in *headings* in every possible combination or permutation. *See also* rotated index.

post-coordination. The combination of *terms* at the

time of a search for a compound *concept*, for example, "cataloging" + "periodicals" for the concept "cataloging of periodicals." *See also* pre-coordination.

postings. The number of *documentary units* to which a *term* or *heading* is assigned.

pre-coordination. The formulation of a multi-term *heading* or a *multi-level heading* to express a compound *concept* in a *displayed index*, for example, "cataloging of periodicals" or "cataloging — periodicals." Pre-coordination differs from the establishment of *compound terms* as *descriptors*, for example, "birth control" (a compound term) vs. "birth control — education — United States" (pre-coordinated terms).

preferred term. *See* descriptor.

probabilistic indexing. The use of weights either through computer algorithm or human estimation to indicate the estimated probability that a *term* will lead to the retrieval of a relevant *documentary unit*. It may be implemented by: (a) The assignment of weights to terms associated with documentary units to reflect the probability that a documentary unit described by a particular term will be considered useful; and (b) the assignment of weights to *terms* to reflect their relative importance in the representation of a search request. Document term weights are often based on relative term frequency or location. The use of weighted terms permits the ranking of retrieved documentary units on the basis of their predicted usefulness.

proximity operator. A search operator that specifies that two or more search *terms* fall within a stated distance relative to each other (for example, adjacent, not separated by more than 1, 2 or more words, within the same sentence or paragraph or record).

qualifier. A word or phrase added to a *term* to distinguish among *homographs* or to clarify the meaning of a term, for example, "races (anthropology)," "races (sports)." A qualifier is considered to be part of a term or *heading*; all qualifiers except epithets (for example, king, saint) should be put in parentheses. *See also* modifier; scope note.

quasi-synonym. A *term* whose meaning is not exactly *synonymous* with that of another term, yet which may nevertheless be treated as its equivalent in a particular *index*. *See also* equivalent term; generic posting; synonym.

record. The description or representation of a *documentary unit* in an *index*. The record usually consists of such elements as author, title, abstract, *terms* or *descriptors*, and location or the entire text of a document.

related term. A *term* that is semantically but not *hierarchically* linked to another term by means of a *cross-reference*, for example, "cooking *see also* food." *See also* associative relationship; broader term; narrower term.

relational display. An *index* display based on relations among *concepts* represented by *headings*. Relations include superordination and subordination, class inclusion, chronology, and various types of *roles* and associations (for example, relations among discipline, action, object or agent of action, material, method, tools, and property). Also called "classified display." Contrast to *alphanumeric display*. *See also* classification.

request-oriented indexing. *Indexing* that is based primarily on anticipated requests or searches and only secondarily on the *topics* or *features* of *documentary units*. *See also* entity-oriented indexing; mission-oriented index.

role. A type of action by which the *topic* represented by a *term* relates to a topic represented by another term in an index *heading* or string of *descriptors*, for example, application, comparison, influence, operation, process. A role does not indicate either a *hierarchical* or an *associative relationship*.

role indicator. A word, phrase, abbreviation or symbol identifying the *role* of a *topic* represented by a *term*.

romanization. The conversion of a non-roman script into the roman alphabet by means of *transcription* or *transliteration* or a combination of the two methods.

rotated index. An *index* in which each *term* serves as a *lead term* or *main heading* in an *entry*, and all other terms constitute a *subheading*; non-lead terms may be displayed in alphanumeric order or their original order may be maintained, as in a *KWAC index*. *See also* permuted index.

run-in layout. The display of *multi-level headings* in an *entry array* in which *subheadings* and *sub-subheadings* are arranged in a single paragraph indented

under a *main heading*. Also known as "paragraph" or "run-on layout." In this layout, sub-subheadings can only be displayed by the use of special punctuation or in a hybrid layout, mixing indented subheadings and run-in sub-subheadings. *See also* indented layout.

run-on layout. *See* run-in layout.

scope headline. *See* guideword.

scope note. An explanation, definition, or clarification of a *term*. A scope note is not part of a term. *See also* qualifier.

search statement. One or more *terms* or phrases submitted to an electronic *non-displayed index* for the purpose of locating *entries* or *records* of interest. Terms may be combined in accordance with *syntax* rules, such as *Boolean* logic. Terms may also be *truncated* or combined with various delimiters, such as *proximity operators*.

"see also" reference. A linking device between two or more *terms* or *headings*, for the purpose of suggesting additional *broader, narrower,* or other *related terms* or headings.

"see" reference. A linking device between an unused or *nonpreferred term* or *heading* and the *synonymous* or equivalent *descriptor* or heading to be used in its place; in *non-displayed indexes*, synonymous and *equivalent terms* may be linked so that all may be included in a search, rather than designating one of the linked terms as a "preferred" term and the others as "unused terms."

serial index. An *index* compiled for a serial (newspaper, periodical, etc.). *See also* open-end index; monographic index.

set-out layout. *See* indented layout.

sorting. *See* arrangement.

specificity. The closeness of fit between the meaning of an indexing *term* and the *topic* or *feature* of a *documentary unit* to which it refers. "Specific" does not mean "narrow." A specific term may be broad or narrow depending on the topic or feature it refers to and its relationship to *broader* or *narrower terms.*

stemming. The removal of suffixes and/or prefixes

from *terms* in automatic *indexing* or in electronic searching. *See also* truncation.

stop list. A list of words considered to be of no value for retrieval. It may consist primarily of function words — articles, conjunctions, and prepositions — but may also include words that occur very frequently in a domain.

string indexing. The creation of *multi-level headings,* or "strings" of *terms,* from individual index terms by computer algorithm. Index terms may be coded, sometimes by *facet* or *role.* A string indexing algorithm puts each important term in the lead position (as *main heading*) and arranges other terms as *subheadings.*

sub-subheading. A modifying *heading* subordinated to a *subheading* in a *multi-level heading. See also* pre-coordination.

subheading. A modifying *heading* subordinated to a *main heading* in a *multi-level heading. See also* pre-coordination.

subject. *See* concept; feature; topic.

subject heading. A *term* or combination of terms used to indicate the summarized overall *topic* of a *documentary unit. Pre-coordination* of terms representing multiple and related *topics* or *features* is a characteristic of subject headings that distinguishes them from *descriptors,* which tend to represent individual *concepts* or features. Subject headings are used in *displayed indexes* and library catalogs, whereas descriptors are designed for *post-coordination* in *non-displayed indexes* in electronic *information retrieval systems.*

subject heading list. An alphabetical list of *subject headings* with *cross-references* from *nonpreferred terms* or *headings* and linking devices between *related terms* and headings. These lists often include separate sequences of standardized *subheadings* that may be combined with subject headings. Rules for applying subheadings usually accompany such lists.

surrogate. A representation of a *documentary unit* in an *index* or *information retrieval system,* such as a citation or citation plus abstract.

synonym. A *term* having a different form, but exactly or very nearly the same meaning as another

term. *See also* equivalent term; quasi-synonym.

syntax. The combination and modification of *terms* to form *headings* and *subheadings* in *multilevel headings* or to form *search statements* for *nondisplayed indexes*; also, the rules for such combination.

term. A word or phrase used to represent a *topic* or *feature* of a *documentary unit* in an *index*.

text. Any organized and meaningful pattern of symbols manifested in a *document*. A text may be verbal (a representation of speech by a writing system); visual, as in the visual arts; musical, as represented in musical notation; performance, as represented in choreography notation; aural, as in sound recordings; etc. Many disciplines, such as chemistry and mathematics, have special symbols to represent texts. *See also* message.

thesaurus. A collection of vocabulary with linking devices among *synonymous, equivalent, broader, narrower,* and other *related terms*; from the Greek for treasure. An indexing thesaurus is a *controlled vocabulary* in which equivalence, *hierarchical,* and *associative relationships* among *terms* are displayed and identified by relationship indicators, which should be employed reciprocally. Its purposes are to promote consistency in the *indexing* of *documents,* predominantly for *post-coordinate information retrieval systems,* and to facilitate searching by linking *lead-in terms* with *descriptors*. A search thesaurus displays vocabulary and term relationships for the purpose of facilitating the retrieval of documents in *free text* searching or from multiple databases in which different controlled vocabularies are used.

topic. An entity, attribute, material, process, operation, event, place, or time period, etc., treated in a *document*. *See also* concept; feature.

transcription. The process of recording the phonological and/or morphological elements of a language in terms of a particular writing system. *See also* romanization.

transliteration. The process of recording the graphic symbols of one writing system in terms of corresponding graphic symbols of another writing system.

truncation. The removal of letters from *search* terms in order to increase the number of terms that will be matched in electronic searches. *See also* stemming.

turnover line. A line indented beneath another to accommodate words that cannot be fit on the preceding line.

uncontrolled vocabulary. *Terms* derived by extraction or selection of significant words or phrases, usually from full *text,* titles, or abstracts. May also refer to search terms freely chosen by a searcher. *See also* free text; keyword.

unit of analysis. *See* documentary unit.

up-posting. The automatic assignment of *broader terms* in addition to the specific *term* by which a *documentary unit* is indexed. *See also* generic posting; specificity.

"used for" term. An unused *term* that is considered *synonymous* or *equivalent* to a preferred term, to which it refers by means of a *see reference* in a *displayed index,* or some analogous link in *nondisplayed indexes. See also* lead-in term; nonpreferred term.

vector. In indexing, a sequence of values, each of which reflects the weight assigned to a *term* associated with a *documentary unit. See also* probabilistic indexing; weighting.

visual index. *See* displayed index.

vocabulary control. The process of organizing a list of *terms:* (a) to indicate which of two or more *equivalent terms* is authorized for use; and (b) to indicate *hierarchical* and *associative relationships* among terms in the context of a *thesaurus* or *subject heading list. See also* authority file; vocabulary tracking and management.

vocabulary tracking and management. The process of tracking, mapping, organizing, and displaying a vocabulary to facilitate *indexing* and/or searching. The results are often displayed in a "search" or "end-user" *thesaurus* and/or integrated with the display of an *index*. Vocabulary tracking and management is similar to *vocabulary control,* except that instead of limiting or controlling the use of vocabulary, it describes and displays vocabulary that has been or may be used.

weighting. The assignment or algorithmic calculation of weights for *index* or search *terms* in order to permit ranking of retrieved items on the basis of predicted relevance. Weighting algorithms are often based on term frequency in *documentary units,* sometimes in relation to frequency across collections. Previous relevance judgments can also be used to increase or decrease weights for terms that lead to relevant or irrelevant documents.

INDEX TO THE TECHNICAL REPORT

In accordance with recommendations in this technical report, the entries in this index refer to numbered sections rather than to page numbers. Consequently, this index may be used in both print and electronic formats. Because the *Foreword* and *Preface* are not numbered, these titles are used for references to these preliminary sections. The index was compiled using NEPHIS: the Nested Phrase Indexing System developed by Timothy C. Craven (*String indexing,* Orlando, FL: Academic Press, 1986). NEPHIS is an example of ad hoc coding syntax for string indexing (*see* section 7.2.5.3).

browsing

displays *(continued)*
    synonymous terms : 6.4
    term relationships : 6.8, 6.8.1, 6.8.2
    "used for" terms : 6.4
    vertical spacing : 8.2.4
    vocabulary : 6.8.1, 6.8.2
document analysis
    methods : 4.5, 5.9
document-oriented indexing
    see: entity-oriented indexing
documentary scope : 5.2
documentary units : 5.7
    definition : 2, 12
    surrogates : 5.14, 7.4.6
documents

> *see also types of documents:* Audio recordings; charts; database records; filmstrips; illustrations; manuscripts; maps; motion pictures; periodicals; photographs; realia; serial documents; slides; video recordings

    capitalization of titles : 6.2.3
    collections. locators : 7.4.1
    definition : 2, 12
    first lines : 6.2.9.5
    non-print. locators : 7.4.2.
    printed. locators : 7.4.1
    serial. locators : 7.4.1
    sources for indexes : 5.3
    titles : 6.2.9, 6.2.9.4; initial articles : 9.4
    types : 1.2, 4.8
domain : 5.3
    definition : 12
double entries
    versus cross-references : 6.8.1.1

electronic manuscripts
    standards for coding : 8.4
electronic media
    display of indexes, browsing : 8.2, 8.3, 8.3.1
electronic search indexes
    display of term relationships, vocabulary : 6.8.2
    guide : 0.2D
entity-oriented indexing
    definition : 12
entity / part relationships : 6.5d
entries
    continuation lines : 8.2.9
    definition : 2, 12
    displayed indexes : 7.1
    double. versus cross-references : 6.8.1.1
    layout : 8.2.5
    subheadings : 8.2.5, 8.2.9
    syntax : 5.12
    turn-over lines : 8.2.5.1, 8.2.5.3
entry-a-line layout
    see: indented layout
entry arrays
    continuation lines : 8.2.9
    definition : 12
entry terms
    definition : 12
entry vocabulary
    definition : 12
equivalence relationships : 6.6f
equivalent terms : 3g
    cross-references. display : 6.4
    definition : 12

errors
    in titles of documents : 6.2.9.4
exhaustivity of indexing : 3b, 5.10
    definition : 12
extractive indexing
    see: derivative indexing

faceted indexing : 7.2.5.2
    definition : 12
facets
    definition : 12
    generic : 7.2.5.2
    literature : 7.2.5.2
false drops
    definition : 12
    in searching : 6.2.6
family names
    see: surnames
features
    definition : 12
files
    authority files. definition : 12
    definition : 12
filing
    see: arrangement
filing rules
    see: alphanumeric arrangement
films
    locators : 7.4.2b
filmstrips
    locators : 7.4.2a
first lines : 6.2.9.5
    initial articles in alphanumeric arrangement : 9.4
footlines
    running : 8.2.6
forenames
    as lead terms : 6.2.9.1
forms of terms : 1.4, 6.2
free-text terms
    definition : 12
full-text indexes : 4.10
functions of indexes : 3

generic posting
    definition : 12
genus / species relationships : 6.5a
geographical names : 6.2.9.3
    initial articles in alphanumeric arrangement : 9.4
geographical relationships : 6.5e
graphics in indexes : 8.2.1
guidelines
    see: standards and guidelines
guidewords : 8.2.6, 8.2.7
    definition : 12

heading-subheading combinations
    linking devices, punctuation : 7.2.3a
headings : 7.1

> *see also:* entries; subheadings; terms

    articulated. *see:* multi-level headings
    continuation lines : 8.2.9
    definition : 2, 12
    initial articles in alphanumeric arrangement : 9.4
    inverted: 6.2.8, 7.2.3
    main. definition : 12
    multi-level : 7.2; definition : 12
    phrase : 7.2.3c

headings

run-on layout
    see: run-in layout
running footlines, headlines : 8.2.6

scope
    of indexes. documentary : 5.2; subject : 5.1
    of this technical report : 1, 1.1
scope headlines
    see: guidewords
scope notes : 6.8.3
    definition : 12
sculpture
    locators : 7.4.2c
search interfaces : 3j, 5.18
search statements
    definition : 12
    post-coordinate combination of terms : 7
    role in non-displayed indexes : 7.5
    syntax : 5.12, 7
searching
    Boolean : 7.5.1
    false drops : 6.2.6
    links among terms : 7.5.4
    methods : 1.5
    probabilistic : 7.5.2
    proximity operators : 7.5.3
    role indicators : 7.5.4
    stemming : 7.5.3
    truncation : 7.5.3
    vector : 7.5.2
    weighted term : 7.5.2
"see also" references
    see also: cross-references
    definition : 12
    location : 6.8.1.3
"see" references
    see also: cross-references
    definition : 12
    versus "see under" references : 6.4
"see under" references
    see also: cross-references
    versus "see" references : 6.4
selection of terms : 1.4, 4.6, 6.1
semantic relationships : 6.6g
serial documents
    locators : 7.4.1
serial indexes : 4.8
    definition : 12
set-out layout
    see: indented layout
single indexes
    versus multiple indexes : 5.4
singular forms
    versus plural forms of terms : 6.2.4
size of displayed indexes : 5.16
slides
    locators : 7.4.2a
software
    for indexing, not covered in this report : 1.5
sorting
    see: arrangement
sources
    of documents : 5.3; vocabulary : 3d-e, 6.1
spaces
    in alphanumeric arrangement : 9.2

spacing
    vertical, in display of indexes : 8.2.4
species / genus relationships : 6.5a
specificity : 3f, 5.11
    definition : 12
    links among levels : 5.13
spelling : 6.2.2
    cross-references : 6.4
standards and guidelines for
    alphanumeric arrangement : 9.1
    bibliographic citations : 7.4.1
    coding electronic manuscripts : 8.4
    corporate body names : 6.2.9.2
    geographical names : 6.2.9.3
    indexing : 10
    names : 6.2.9
    personal names : 6.2.9.1
    titles of documents : 6.2.9.4
Standards Committee YY
    see: Z39.4 Committee
Standards Development Committee (NISO)
    charge to Z39.4 Committee : *foreword*
stemming
    definition : 12
    in searching : 7.5.3
stop list
    definition : 12
string indexing : 7.2.5
    definition : 12
structured arrangement
    headings with same initial term : 9.6
style
    see: layout (of entry arrays)
sub-subheadings
    definition : 12
subheadings : 8.2.5, 8.2.9
    alphanumeric arrangement : 9.5
    chronological arrangement : 9.5
    continuation lines : 8.2.9
    definition : 12
    hybrid indented/run-in layout : 8.2.5.3
    indented layout : 8.2.5, 8.2.5.1
    prepositions : 7.2.1
    run-in layout : 8.2.5, 8.2.5.2
subject heading lists
    definition : 12
subject headings : 7.2.3
    definition : 12
subject scope : 5.1
subjects
    see: concepts; features; topics
subordinate corporate bodies
    names : 6.2.9.2
summary of this technical report : 0.1
surnames : 6.2.9.1
surrogates : 5.14, 7.4.6
    definition : 12
    display : 5.15
symbols
    alphanumeric arrangement : 9.2
    in titles of documents : 6.2.9.4
    textual : 5.5
synonymous terms : 3g, 6.4
synonyms
    definition : 12
syntactic cross-references : 7.2.6