3. STANDARDS AND GUIDELINES

3.1 TYPES OF SEARCHES

Types of searches, rather than types of indexes, are specified in Z39.50 because system designers may choose to provide the same functionality in multiple ways. For example, the system may have a single keyword index but allow the user to search author, title, and subject subsets of that index, or the system may have separate author, title, and subject keyword indexes. Browse searches may return all of the terms matching the search in a browseable list, or they may permit the user to browse the entire index, and not just those entries matching the search. The desired functionality is more important than the technical means by which it is achieved.

The system may support the combination of find searches of phrase indexes with other find searches of keyword indexes or search qualifier indexes.

3.1.1 REQUIRED SEARCH TYPES

3.1.1.1 Scan Searches

For scan searches (also known as exact or browse searches), the searcher must enter a search string that begins with the first indexed character in the first indexed subfield and includes all indexed characters in the same order in which they appear in the bibliographic record. The term can be truncated at any point.

The system shall support the following types of scan searches:

- Author
- Title
- Subject

A library shall be able to choose which type of subject entries will be indexed in its catalog, based on the nature of its collection. Subject search options shall include one or more of the following:

- Library of Congress Subjects (LCSH);
- Medical Subject Headings (MeSH);
- Other subject headings, e.g., Library of Congress Children’s Subject Headings, Sears Subject Headings, etc.;
- Combined subject search, i.e., a search of all types of subject headings indexed in the system. Given the contradictions between different subject authority files, linkage to multiple subject authority files in a combined subject index shall not be required.
3.1.1.2 Find Searches

For find searches, sometimes called keyword searches, the searcher may enter any word(s) from an indexed field or subfield. Order of words is important only if system supports use of the positional operator ADJ, or use of parentheses or other conventions to group terms to be processed together, and the searcher uses one of these features.

The system shall support the following types of find searches:

- Author;
- Title;
- Subject;
- General (i.e., a search of all author, title, and subject keywords plus keywords from additional fields).

3.1.1.3 Number Searches

For number searches, the searcher must enter data in the same order in which it appears in the bibliographic record, beginning with the first relevant character in an indexed subfield.

The system shall support searches for standard and other numbers of the following types:

- International Standard Book Number (ISBN);
- International Standard Serial Number (ISSN);
- Library of Congress Control Number (LCCN);
- bibliographic record identifiers (both system-assigned and operator-assigned) as appropriate to the local system.

It should be possible with a single search to retrieve numbers of a given type whether coded as valid, invalid, or obsolete.

3.1.1.4 Search Qualifiers

Indexed data elements that a searcher will ordinarily use to limit find search results—e.g., to titles in a specific language—are generally described as qualifiers. However, a searcher may also wish to search using only these data elements—e.g., to search for all works in the catalog that were published in 1852.

3.1.1.4.1 Qualifiers supported

The system shall support qualification of find searches of either phrase or keyword indexes by find searches of the following data elements.

- single date of publication; dates of publication earlier than, later than, or equal to a specified date of publication; or, a range of publication dates;
- language;
- format.
The system may support qualification by country of publication.

The system may support qualification by field in bibliographic record.

3.1.1.4.2 Multiple qualifiers

The system should support use of multiple qualifiers of the same or different types in a single search, e.g., English and French language, German language and serial format.

3.1.1.4.3 Boolean operators with qualifiers

The system shall support use of the Boolean operators AND, OR, and NOT in conjunction with qualifiers.

3.1.1.4.4 Additional qualifiers

The system should support definition of additional qualifiers.

3.1.1.4.5 Choosing qualifiers from a list

The system should provide the searcher with a list of spelled-out choices for qualifiers other than dates, rather than requiring the searcher to use the codes that appear in the bibliographic record.

3.1.1.4.6 Use of qualifiers as the sole search element

The system may support use of data elements typically used as qualifiers as the only elements in a find search command.

3.1.2 OTHER SEARCH TYPES

The system may support additional types of searches. A library may choose to implement any that are appropriate for its collection.

3.2 MARC21 FIELDS AND SUBFIELDS INDEXED

3.2.1 SPECIFICATION OF FIELDS AND SUBFIELDS INDEXED

It should be possible for a library to specify data elements to be included in each index and, when applicable, each subset of an index.

It should be possible to specify indexing of data from variable length fields at both the field and subfield level. If specification of indexing at the subfield level is not possible, all traditionally indexed subfields in a specified field (e.g., excluding those with numeric designations such as ‡2) shall be indexed.
It should be possible to specify indexing of specific bytes of fixed length fields and fields containing fixed length data elements such as three-character language codes in 041.

3.2.2 REQUIRED FIELDS AND SUBFIELDS IN INDEX ENTRIES

Index entries searched when the searcher enters a specified type of search or a specified qualifier shall be constructed from all of the appropriate MARC21 fields and subfields listed in the Tables at the end of these standards and guidelines. Older MARC format tags in the Tables may also be used as appropriate.

The Required data elements listed in the Tables should be viewed as a minimum set. It may be desirable to index other data elements.

3.3 System Features

3.3.1 STOPWORDS

3.3.1.1 No stopwords in scan searches

There shall be no stopwords in scan or browse searches. (Exception: The system may require the searcher to omit initial articles from title phrase searches.)

The system should not require the searcher to “mask” system-defined operators that are part of scan searches, e.g., by surrounding the operator or the whole search string with quotation marks.

3.3.1.2 Stopwords in find searches

The number of words that must be defined as stopwords in keyword index searches shall not exceed 20. Ideally, it should be confined to the list of system-defined operators.

3.3.1.3 Definition of additional stopwords

It should be possible for a library to define keyword index stopwords in addition to those required by the system.

3.3.1.4 Processing of stopwords in search arguments

If a searcher includes a stopword other than an operator in a keyword search argument, the system should ignore the stopword, i.e., it should process the search as if the stopword were not present. It should also return a message to the searcher indicating that the stopword was ignored.
3.3.2 TRUNCATION

3.3.2.1 Types of truncation implemented

A library should be able to specify which available types of optional (not required) truncation will be implemented in its system.

3.3.2.2 Right-hand truncation of scan searches

Right-hand truncation of scan or browse searches shall be possible. It may be automatic, or the searcher may be required to invoke it. If it is automatic, it should be possible for the searcher to override it.

3.3.2.3 Right-hand truncation of find search terms

Right-hand truncation of any find search term shall be possible. It shall not be automatic.

3.3.2.4 Right-hand truncation of multiple keywords

Right-hand truncation of multiple keywords in a single find search argument shall be possible.

3.3.2.5 Internal truncation

Internal truncation of keyword terms (also known as “masking” or “wild cards”) shall be possible.

3.3.2.6 Internal truncation of multiple keywords in a single search argument

Internal truncation of multiple keywords in a single search argument should be possible.

3.3.2.7 Number of characters represented by a truncation symbol

Specification of the number of characters represented by a truncation symbol should be possible.

3.3.2.8 Multiple types of truncation in a single search argument

Right hand truncation of some keywords and internal truncation of other keywords in a single search argument should be possible.

3.3.2.9 Internal and right-hand truncation of a single keyword

Both internal and right-hand truncation of a single keyword may be possible.

3.3.2.10 Left-hand truncation

Left-hand truncation of keyword search terms may be possible.
3.3.3 OPERATORS

3.3.3.1 Boolean Operators

AND, OR, and NOT shall be available for use

• between two find searches of phrase index strings, or between find searches of a numeric index or keyword index and a phrase index;
• between keywords being find searched in a single index.

3.3.3.2 Positional Operators

ADJ, SAME, WITH, and NEAR should be available for use between keywords.

The ability to specify the maximum number of words in the index between the keywords in a search string that includes SAME, WITH, or NEAR may be available.

3.3.3.3 Default Operator

It should be possible to select a default Boolean or positional operator for find or keyword searches.

3.3.3.4 XOR

The Boolean operator XOR may be available.

3.3.3.5 Multiple Operators

Use of multiple operators in the same search argument and use of parentheses or other conventions to group terms connected by an operator shall be possible.

3.3.3.6 Nesting

Nesting of terms in keyword search arguments should be possible.

3.3.4 NORMALIZATION OF SEARCH STRINGS

The system should normalize search strings using the same routine used for creation of index entries before comparing the search string to index entries, rather than requiring the searcher to normalize data in the search string in order to retrieve correct, complete results.

3.3.5 SETTING QUALIFIERS

It may be possible for the searcher to set qualifiers that remain in effect for each find search performed during a search session or until they are reset. For example, a user might want to restrict results of a long series of subject find searches to German language titles published before 1850 without specifying those qualifiers each time s/he enters a search.
3.3.6 DISPLAY AND SORTING OF INDEX ENTRIES, MOVEMENT AMONG INDEX ENTRIES

3.3.6.1 Types of Displays

The system should present at least two types of displays in response to searches of phrase or heading indexes:

- for scan searches, a browsable display of index entries only, i.e., of all index entries that match the search type, with the closest match to the search string highlighted in some way on the first screen displayed;
- for find searches, a display of brief record-level entries that match the search string (these entries may include additional information about titles associated with index entries, e.g., author index entries plus titles associated with each author, or subject index entries plus associated titles).

The system should present at least one type of index display in response to a numeric search. This display should consist of numbers that match the search string and information about titles associated with those numbers.

The system should present at least one type of index display in response to a find search of a keyword index. This display should consist of information about titles in which the keywords in the search argument appear.

When no hits are returned, the system should display a context sensitive help screen with advice for revising the search.

3.3.6.2 Details in Index Displays

3.3.6.2.1 Number and length of data elements in index displays

The number and length of data elements displayed in index entries of the second type should be sufficient to allow the searcher to identify and distinguish possibly relevant titles.

3.3.6.2.2 Display of number of hits retrieved

In displays of index entries only, the number of hits represented by each entry should be indicated.

3.3.6.3 Sorting of Index Entries

3.3.6.3.1 Number and length of data elements used in sorting

The number and length of data elements used to sort index entries should be sufficient to allow the searcher to distinguish among entries in catalogs of five million or more records.
3.3.6.3.2 Like data elements in a single sequence

All data elements of the same type should be sorted in a single sequence, e.g., all titles should be sorted in a single sequence regardless of the field in which they appear in the bibliographic record. In displays of index entries only, all entries that would normalize the same should be represented by a single index entry with an appropriate hit count.

3.3.6.3.3 Default sorts

3.3.6.3.3.1 Scan search responses

For responses to scan or browse searches, the library should be able to define a default sort plus at least two subsorts for index entries that include information about titles associated with an index entry.

3.3.6.3.3.2 Numeric and keyword search responses

For responses to numeric and keyword searches, the library should be able to define a default sort plus at least one subsort for index entries that include information about titles associated with an index entry.

3.3.6.3.4 Sorting by relevancy ranking

The system may support sorting of find or keyword search results based on relevancy ranking.

3.3.6.3.5 Resorting records

The system should allow the searcher to resort a list of records by specified data elements.

3.3.6.4 Movement From Index Entries to Bibliographic Records and Among Index Entries

3.3.6.4.1 Movement from index entries to bibliographic records

The system should not require the searcher to navigate more than two levels of index displays to reach brief or full displays of individual bibliographic records.

When a search retrieves only one record, the system shall display the record, rather than an index entry.

The system should allow the searcher to retrieve bibliographic records associated with multiple headings in a single transaction.

3.3.6.4.2 Scan searching of keyword indexes

The system may include browseable indexes of author, title, subject, and keyword index entries, through which the searcher can page or scroll.
3.3.6.4.3 Movement among index entries

It **should** be possible to “jump” to a specified point or a specified distance in a long list of index entries, i.e., the user should not be required to page through all entries screen by screen.