

AN ARCHIVE TAPE PROCESSING SYSTEM
FOR THE TRIANGLE RESEARCH LIBRARIES NETWORK

by

Jeanne Clifford Sawyer

AN ARCHIVE TAPE PROCESSING SYSTEM FOR THE TRIANGLE RESEARCH
LIBRARIES NETWORK

by

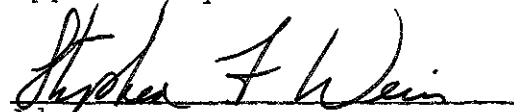
Jeanne Clifford Sawyer

A Thesis submitted to the faculty of
The University of North Carolina at
Chapel Hill in partial fulfillment of
the requirements for the degree of
Master of Science in the Department
of Computer Science.

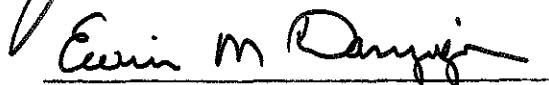
Chapel Hill

1982

Approved by:


Adviser


Reader


Reader

JEANNE CLIFFORD SAWYER. An Archive Tape Processing System for the Triangle Research Libraries Network (Under the direction of Stephen Weiss.)

The purpose of the archive tape processing system is to process tapes received through the OCLC-MARC Subscription Service, thereby constructing and maintaining a master database of bibliographic records for the libraries of each of the three institutions, Duke University, North Carolina State University at Raleigh and the University of North Carolina at Chapel Hill. This thesis discusses the problems addressed by the TRLN tape processing system and the manner in which they were resolved.

TABLE OF CONTENTS	PAGE
Introduction	1
The Task	5
The OCLC Cataloging System	7
Problems with OCLC and Archive Tapes	9
I. Diversity of requirements within an institution.	10
II. Error detection	12
III. Error correction	13
IV. Collection of statistics	14
The TRLN Archive Tape Processing System	16
I. Diversity of requirements within an institution.	17
II. Error detection	18
III. Error correction	20
IV. Collection of statistics	23
Conclusion.	25
Glossary	27

INTRODUCTION

Libraries have traditionally maintained catalogs of the materials in their collections to aid users in finding specific items. Various points of access such as author name, title, or subject are generally provided. Catalogs have been created in a variety of formats including printed book catalogs and the familiar card catalog, the most common manually created catalog because of its relative ease of updating.

The task of constructing the card catalog, i.e., creating catalog records, producing the cards and filing them into the existing catalog, however, has always been one of the most labor intensive, and therefore expensive, aspects of library service. The value of sharing cataloging records has long been recognized as an important means of reducing the cost of providing them. When libraries accept the same cataloging standards, the same bibliographic record should be adequate for all libraries that own the item represented by that record. Since in practice, the Library of Congress (LC) determines standards that most libraries in

NOTE: A glossary of special terms is appended. The first occurrence of these terms in the text is marked with an asterisk(*) .

the United States follow, most libraries are able to use cataloging records produced by LC with little, if any, modification. Thus, libraries can save substantial amounts of time and money by using LC records. The same theory applies to some extent to bibliographic records created by libraries for items not cataloged by LC. However, usually some modifications to these records are required by other libraries before they can be used, mainly because of differences in local interpretation of the standards. Even when the records must be carefully checked, it is still more efficient to modify a catalog record produced by someone else than to create the entire record. The National Union Catalog is an example of a source for such records in the form of catalog cards reprinted in book format. After the cataloger determines what data should be included in his catalog, whether by copying or modifying an existing record or by creating a record, the master card must be typed, duplicates of the master prepared for all access points, headings typed onto the duplicates, and the set of cards filed.

Although some individual libraries began trying to reduce the cost of cataloging by using computers, it was not until OCLC, Inc., originally the Ohio College Library Center, began providing online cataloging service in 1971 that many libraries began to realize the potential of automating the production of catalog records. While the traditional method of preparing and filing cards is still

frequently used, the services of OCLC and other bibliographic utilities are increasingly being employed, not only to produce catalog cards and share cataloging data, but also to produce catalog records in machine-readable format.

Although OCLC allows the accumulation of records in machine-readable form, most libraries have only taken advantage of the capability for sharing cataloging data and producing cards, which are then filed manually into existing card catalogs. In general, the possibilities for more flexible searching and for faster and easier catalog maintenance offered by computers manipulating machine-readable records are just beginning to be explored. The research libraries at North Carolina State University at Raleigh (NCSU), Duke University in Durham and the University of North Carolina at Chapel Hill (UNC-CH) recognize these possibilities and are developing an online catalog that will make the machine-readable bibliographic records created through OCLC available for their patrons and staff.

These libraries have a long history of cooperative activities. A joint committee, established in 1976, called the Triangle University Libraries Cooperation Committee (TULCC) meets regularly to review ongoing activities and to recommend new areas of cooperation to the library directors. The advantages of an online catalog were readily apparent to members of this committee who must deal with many issues requiring timely access to all three catalogs. Some of these areas include close coordination of book and material

selection and establishment of direct borrowing privileges for patrons of all three institutions. In many respects, the collections of the three libraries can be regarded as a single resource.

In order to meet these needs for mutual access to the three collections, funding was sought to develop an online catalog network which would be able to meet the needs of library patrons as well as staff. A computer located in each library would support the online catalog for that library. These catalogs would be operated in a distributed network of the three compatible computers so that users located at any of the three libraries would be able to consult any of the three catalogs. If desired, the network could later be expanded to include other libraries in the area or could be replicated in other areas. A two year grant for \$554,000 was received in 1979 under Title II-C of the Higher Education Act of 1965 (Strengthening Research Library Resources) to begin development of the Triangle Research Libraries Network (TRLN).

THE TASK

The first phase of development of the online network required constructing local master databases(*) of bibliographic records for each of the three libraries. These bibliographic master databases will provide the primary source of data for most library functions, including the public catalog. To serve this purpose, each bibliographic record in the master database must represent a different title held by the library, or, if multiple records are used for different locations, it must be possible to associate those records that represent the same title. It must be possible to show the relationship of physical pieces (e.g., volumes in the case of books) to the title represented by a bibliographic record and to show where the pieces are located. The content of the records, both bibliographic and holdings data(*), change frequently, so a flexible means of updating the database is required. Since the records themselves are extremely complex, adequate validation of records and a means of correcting errors without engendering more problems are very important.

The TRLN libraries decided to construct the databases from machine-readable records created through the OCLC system. Although this placed certain constraints on the

methods that TRLN could use for constructing the databases, there are several important advantages in using OCLC. The OCLC system is in full operation, so there is no doubt that the libraries can accumulate bibliographic records by using it. Indeed, all three libraries were already committed to using OCLC for cataloging and had been creating machine-readable cataloging records in the form of archive tapes(*). In addition, the problems that TRLN would have to solve are known, and, although the solutions may be somewhat awkward, there are solutions. Most importantly, having access to the OCLC resource database for shared cataloging is so significant that to try to develop an independent cataloging system would be foolhardy. Thus the method chosen by the TRLN libraries for initially constructing and maintaining the master databases is a system to process archive tapes created through the OCLC system. This thesis discusses the problems in using the OCLC cataloging system to acquire machine-readable records for individual library catalogs and the methods used by TRLN to solve them.

THE OCLC CATALOGING SYSTEM

OCLC, located in Ohio, provides an online resource database of over six million bibliographic records for its member libraries to use. If a record is not already in the database, an individual library may add it, thus making the new record available to other OCLC member libraries. All member libraries thus participate in constructing the resource database. However, once a record has been entered into the resource database and other member libraries have used it to create their own local copies, only OCLC staff can update the record. A record input by a library can be corrected by that library if no other member has used it, but generally errors discovered by member libraries are reported by written notification to OCLC, a policy intended to provide quality control over the database by causing all changes to be reviewed by a central staff.

While only OCLC can modify the OCLC database, records in that database may be edited by local library staff to create local records that meet their own specifications. The original OCLC master record is not modified when a library creates a local record. Instead, the local version of the record is put onto catalog cards and onto magnetic tapes called archive tapes, which can then be purchased from

OCLC. The OCLC system does not retain any local versions of records, so if a library later needs to change a local record, the OCLC master record must be completely re-edited to match local specifications unless the library has developed a local system for editing records. When a library does edit a master record to produce a local record, the entire record is copied onto the library's archive tape; no means for changing only portions of local records is provided by OCLC. Thus, the archive tapes purchased by the library simply provide a trace of all cataloging transactions in the form of complete catalog records, as created by the member library. The archive tape may have many copies of a record, each intended to replace its predecessor.

The following is an example of how an item is cataloged through OCLC:

1. A cataloger at a terminal in Library X retrieves a record from the OCLC resource database that matches the item to be cataloged.
2. Something in the OCLC record needs to be changed to match local policy, perhaps the form of an author's name. The cataloger indicates all changes on the OCLC screen.
3. The cataloger "produces" the record, causing cards to be printed and the record to be written on the library's archive tape. The only change to the OCLC master record is the addition of a holdings symbol to show that Library X has the item.
4. The cards arrive at the library. Inspection shows a typographical error. The archive tape arrives also, but without a tape processing system, it provides no useful information.

5. The cycle is repeated until there are no known errors in the record. Once no errors are apparent, the process is complete and the cards can be filed. The amount of work involved each time the cycle is repeated does not decrease since the cataloger must recreate the entire record, rather than only fixing the errors. Note that the first time the record is created, the cataloger has the book in hand. By the time errors are discovered and the cataloger re-edits the record, the item has already been shelved and thus is not readily available.

The machine-readable bibliographic records conform to a standard format, MARC(*), which dictates both the record organization and contents. The data portion of the record consists of two sections: a fixed length section of coded data and a variable length section of a variable number of variable length fields. The entire fixed length section and most of the variable length section are strictly defined. However, there are a few variable fields available for local use. Each variable field has a numeric tag associated with it to identify the nature of the field. For example, the title has a tag (245) which distinguishes it from authors (100) or subject headings (650). Although there are a large number of fields, only a few actually occur in any one record. The cataloger selects the data and assigns the appropriate fixed length codes and field tags. The OCLC system then constructs the control portion of the record which includes some general information about the total record, such as the total length, and the record directory, which gives the tag, starting location and length for each variable field.

Problems with OCLC and Archive Tapes

Before they can be used by a local library, the machine-readable cataloging records on the archive tapes must first be processed to change the collection of archive records into a local master database of bibliographic records. Unfortunately, although the basic concept of what is required is quite simple, large research libraries such as those in the Research Triangle must solve several problems with using the OCLC system in the manner expected by OCLC, i.e., the most recently created archive record(*) becomes the new local master record(*) for an item, completely replacing any previously existing local master record. Since the archive records are on the archive tape in order of creation, recognizing new master records is very straightforward. This is the primary task any archive tape processing system must accomplish; there are a number of such systems available. However, none adequately address the problems caused by creating local master records in this way. Four of these problems addressed by TRLN are discussed below.

I. Diversity of requirements within an institution.

One problem is caused by the OCLC system assumption that all holding libraries within an institution are able to use exactly the same bibliographic data(*) and that holdings statements for the entire institution can be constructed.

It is neither possible nor desirable for separate, specialized collections within an institution, such as rare books collections, to have to conform to the needs of the main collection in complex research libraries. Separate shelflists(*), which are created independently of each other and of the main library, are generally maintained by such special collections. Holdings of these collections are not reflected in other shelflists in the institution; therefore, holdings statements for the entire institution cannot be constructed. Although this may not be an ideal situation in an automated environment where the computer could selectively retrieve records for a given collection, if separate card shelflists have traditionally been maintained it is generally not practical to combine them retrospectively. Special collections also have special requirements for bibliographic data, e.g., more detailed information about an item than is generally required, or specialized subject headings may be needed. As a result, there may be more than one correct version of a record, i.e., complementary records(*), maintained in the local master database.

In addition to requiring separate versions of a record for independent collections, it is occasionally necessary to keep more than one version of a record within a cataloging center(*). OCLC has established standards as to what conditions justify creating separate records for two items. For example, British and American editions of a work justify

separate records, but reprintings of the same edition do not. Creation of OCLC master records must conform to these standards. However, special collections sometimes need finer distinctions than the OCLC standards allow. Thus, for example, although different printings of the same edition do not ordinarily require separate records, if the different printings were of an original edition of a significant work such as a Dickens novel, then separate records would be justified for a rare books collection. In this case, multiple versions of a record within a cataloging center are required. These are called multi-use records(*) .

II. Error detection.

Although proof-reading is the only method for detecting errors in the actual data, automated methods can be used to detect some tagging and coding errors. Without automated validation, such as can be provided by an archive tape processing system, it is impossible to detect errors in the non-printing portions of the record. Although OCLC does some validation of tags, no consistency checks are made.

Adequate error detection and correction is especially significant because the problems in records stem not only from ordinary mistakes but also from changes in the MARC format over the years. For example, many fields now require indicators, called filing indicators(*), to show the number of characters to be disregarded in sorting, as for initial articles. Formerly, these indicators were not allowed; thus

older records do not contain them. It is necessary to correct out-of-date records so that the entire database is brought up to current standards.

III. Error correction.

Since the OCLC system always copies the entire record onto the archive tape, updating the master requires completely replacing the record. All changes to records, no matter how minor, force the entire record, including both the bibliographic data and the holdings information, to be completely re-created. Whenever a record must be corrected, the same process that created the original local record must be repeated.

OCLC does not provide any means for humans to see the entire local record, so reconstructing the local version of the data is difficult and time-consuming. The bibliographic record that appears printed on catalog cards does not include all of the data that must be included in the local machine-readable record. Thus the cataloger cannot see the complete record as it was originally created, and must attempt to reconstruct the information from what partial information is on the card and from memory. Data which are not printed on the cards include information in highly coded form, e.g., the language and country of publication, and coded cataloging information such as whether a name is personal or corporate or whether an initial article is present in a title. These data are not important for

maintaining a card catalog, but are vital for an accurate catalog in an automated environment. Generally the item represented by the record is no longer in hand and the complete original local record cannot be read, so the reconstruction involves some guesswork. For example, it may be impossible to determine from data printed on the catalog card if an item is a federal, state or local government publication or even if the item is a government publication at all. Trained catalogers are required to correct even minor typographical errors. In addition to the time required to reconstruct the cataloging data, large amounts of staff time are required to key changes and re-key local variations to the OCLC record. As a result, it is virtually impossible for a large library to have all desired or needed changes actually reflected in a local database by using the OCLC system. In many cases, the changes are made manually in the card files, but are simply never made in the machine-readable files.

IV. Collection of statistics.

A wide variety of statistics are collected in a large library in order to monitor collection growth and cataloging activity. For collection growth, statistics are kept of the number of titles and physical volumes (a single title may consist of multiple volumes) in the library as a whole as well as within each departmental library. Breakdowns are maintained for each copy as to whether it was new to the

institution as a whole, new to the department, or an additional copy at a location that already had a copy. Separate totals are maintained by type of material, i.e., books, phonodisks, microfilm, etc. Cataloging activity statistics involve keeping track of the source of cataloging copy (the Library of Congress, other OCLC members or the local library), the type of cataloging being done (new items, recataloging an item that already existed in the collection, transferring items from one location to another, etc.), and the production and type of cataloging done by individual catalogers.

The effort to collect these statistics is tedious, cumbersome and inaccurate. Incoming archive records, when compared to previously existing local master records, include most of the data necessary to produce these statistics. Thus, an archive tape processing system can logically address the problem of compiling statistics.

THE TRLN ARCHIVE TAPE PROCESSING SYSTEM

The TRLN Archive Tape Processing System is a batch system, using OCLC archive tapes and the old master database as input and producing an updated master database of bibliographic records as the primary output. A variety of printed reports, mainly of exceptions and statistics, are also produced. The system consists of six programs and three subroutines with a total of 10,700 lines of source code. It is written in ANSI '74 COBOL. The operating environment for the system is the Univac 90/80 located at the Administrative Data Processing Department of UNC-CH.

Archive tapes containing approximately 13,500 records are received every four weeks for the three libraries. Of these, approximately 18% represent updates to previously existing records. The local master databases include a total of 59,500 records for Duke, 278,225 records for NCSU and 183,900 records for UNC-CH. (Duke joined OCLC recently and NCSU has been engaged in an extensive project to convert existing records to machine-readable form).

Design of the tape processing system began with several constraints already present. 1. The libraries were committed to continuing their membership in OCLC or another

bibliographic utility(*). Specifically, since the archive tapes are received from OCLC, the local system had to deal with OCLC system characteristics. 2. Local master records had to be maintained in accordance with national standards for bibliographic records, i.e., it was necessary to maintain the full MARC record with all formats (monographs, serials, maps, sound recordings, etc.) included. 3. Maintaining a separate master database for each of the TRLN libraries was required.

Solutions for the four major problems discussed in the previous section are addressed by the TRLN Archive Tape Processing System. Each will be considered in turn.

I. Diversity of requirements within an institution.

The archive tape processing system handles the problem of different requirements within an institution by maintaining physically separate local master records for each cataloging center or separately administered library(*). It did not seem productive in a batch-mode system using magnetic tape storage to attempt to eliminate data redundancy by merging records where so many differences would still have to be maintained. Similarly, complete, separate multi-use records are maintained where a cataloging center has determined the need for fine distinctions among records.

The tape processing system automatically sets a

cataloging center code as part of the record identification key based on the holding library code. These codes, assigned by OCLC, are already entered in the record for other purposes, i.e., to indicate precise locations of items and to order cards for the appropriate catalogs. Thus, cataloging staff are not required to input special codes to signal to the tape processing system that complementary records for different cataloging centers should be created. However, an additional transaction code is necessary to allow a cataloging center to delete its record from the local master database.

Unlike complementary records for different cataloging centers, the system cannot determine automatically when multi-use records should be maintained. Therefore, an explicit multi-use number is necessary to indicate that an incoming record is a new multi-use record rather than an update of an already existing local master record. Multi-use numbers are assigned sequentially with a separate sequence for each cataloging center's group of multi-use records. The first version of a record has an implicit multi-use number. When a second version of a record is required, the next multi-use number must be explicitly entered. Thus, for the majority of records, which are the first version of a record, no special action is required from cataloging staff.

II. Error detection.

The archive tape processing system includes extensive validation routines to check for errors in the catalog records. Of course, errors in the non-coded cataloging data cannot be detected automatically, but codes that are not filled in at all and inconsistencies in tagging can usually be detected by the system.

The archive tape processing system detects two types of errors which can be identified but cannot be corrected automatically: 1) Processing errors that prevent the system from determining the appropriate action to take with a specific archive record and 2) Logic errors in the cataloging data that should be brought to the attention of cataloging staff but which do not prevent proper processing of the record. There are two types of logic errors: bibliographic errors and warnings.

Processing Errors. Processing errors are caused either by OCLC or local transaction code errors or by problems with the record key (which consists of the OCLC number, cataloging center code and multi-use number). They include such problems as attempting to cancel a record that does not exist in the master file or using an invalid multi-use number. Generally records with processing errors must be corrected by re-editing the OCLC master record. Under TRLN procedures, cataloging staff is notified of the occurrence of processing errors through printed exception reports.

Logic Errors. Logic errors are grouped into two categories according to the severity of the error. The more severe category, called "bibliographic errors", consists of such errors as having no title or no call number on a record. Records that have bibliographic errors are added to the local master databases because the system must be able to process subsequent archive records, but are flagged since the error is so severe as to render the record useless to library users. These records are not printed in COM catalogs or other listings. Cataloging staff is notified of the occurrence of errors in bibliographic records by exception reports; their correction must always be made through OCLC.

The second category of logic error is not severe enough to prevent use of the record. Therefore, this category causes "warnings" to be generated in the exception reports although the record is added to the master databases and is used in all products derived from the master databases. Warnings are caused by such inconsistencies and errors as leaving coded information blank or indicating multiple dates while supplying only a single date. Records which cause warnings must also be corrected by re-editing the OCLC master record.

III. Error correction.

The optimal method for maintaining local records would be an editing system that would allow cataloging staff to

examine local master records and make changes in an online environment. The TRLN libraries anticipate developing such a system, but it was felt that a simpler system that could be used immediately would also be worthwhile. Therefore, a system of local transaction codes was developed to allow holdings data to be treated separately from bibliographic data.

The system developed requires that each type of data always be considered in its entirety, i.e., if a change to any bibliographic data element is required, all bibliographic fields must be re-edited. It was decided not to develop a system of transaction codes that would allow editing of specific fields since such a system operating in batch mode would be extremely complex. Since the online system for editing local records will ultimately eliminate the need for the transaction code system, the goal of the transaction code system is merely to provide some relief to cataloging staff responsible for maintaining the databases as promptly but inexpensively as possible.

The system works as follows. Mnemonic codes are entered by the cataloger into one of the local use fields when a record in the OCLC system is being edited. The local codes are interpreted by the archive tape processing system to produce a specific action. They are used only when changing a record that already exists in the local master database. For a specific record, if only the bibliographic data should be replaced, a prescribed transaction code must

occur in the archive record. If only the holdings data should be replaced, a different transaction code must occur. Use of these codes is optional since records can continue to be completely re-edited. If the entire archive record, including both holdings and bibliographic data, either should be added to the master database or should entirely replace an existing record, no transaction code is necessary.

In addition to easing the difficulties of correcting errors manually, the tape processing system automatically fixes problems in individual records where the appropriate correction can be determined with reasonable certainty, i.e., if the effort involved in manually correcting the computer's errors is less than the effort of manually correcting the original errors. In most cases, for example, the system supplies missing filing indicators. These indicators are supplied in accordance with a language code that also appears in the record by using a table of language codes, initial articles and the associated filing indicators. Since the language code does not always match the actual languages in the fields, all records which have automatically supplied filing indicators are printed for manual review of the indicators. For example, an analysis in English of Les Miserables could simply be titled Les Miserables. However, since the text is in English, the record would be coded as English. In this case, since "Les" is not an initial article in English, the system would

incorrectly supply zero as the filing indicator.

IV. Collection of statistics.

Collection size and cataloging activity statistics are determined by the nature of the item being cataloged or recataloged and its relationship to the existing collections. Which counts should be incremented and the amount of the increment are determined by comparing the holdings in the archive record with any previously existing record for the item. Cataloging staff must provide information about the nature of the cataloging transaction in order for proper statistics to be gathered. For example, recataloging items previously in the collection, transferring an item from one holding location to another, and cataloging for an item with more than one title bound together each requires the addition of relationship information.

The indicators associated with the holdings data field have been designated by OCLC as available for local use, and, since this field is the basis for gathering statistics, these indicators are appropriate to use for providing the necessary relationship information.

The archive tape processing system will compare the holdings statement of the incoming archive record to the holdings statement in the previously existing master record, if necessary, and increment the appropriate totals based on the comparison and the indicator values. For example, when

a new record is added to the master database with no special indicators, the number of new titles is incremented by one. If another copy is added, the new archive record's holdings statement will show both copies, but comparison of the holdings statements of the new archive record and the old master record will indicate that one copy was already counted and that the total for added copies should be incremented by one.+

+Although functional requirements for the statistical reporting module of the archive tape processing system are complete, the module has not yet been implemented.

CONCLUSION

The TRLN libraries have made substantial progress towards effectively using computers to share bibliographic information, maintain catalog records that represent the specific library's holdings and selectively retrieve bibliographic records in response to individual queries. By using the OCLC Cataloging System, the libraries share bibliographic data and create local catalog records in machine-readable form. The TRLN Archive Tape Processing System accomplishes its primary goal of building and maintaining master databases of bibliographic records for the TRLN libraries as well as the secondary goals of alleviating some of the difficulties caused by the OCLC system. As long as the only means of receiving local machine-readable cataloging records while participating in OCLC is through the OCLC-MARC Subscription Service, the TRLN Archive Tape Processing System will remain a reasonable approach to creating and maintaining local master databases. Although the retrieval aspects of the system have thus far only been considered at a general level, the TRLN libraries have prepared the basis for detailed planning, design and implementation of an online retrieval system and public

catalog by participating in OCLC and by developing the Archive Tape Processing System.

GLOSSARY

Archive record or archive tape record: a record that appears on the archive tape which is received through the OCLC-MARC Subscription Service. The records are created by use of the PRODUCE or UPDATE functions in the OCLC Cataloging System. An archive record may either be new to the local master database or a correction to or deletion of a previously existing record.

Archive tape: magnetic tapes received from OCLC that contain the local versions of bibliographic records created by a library.

Bibliographic data: as applied to the use of local transaction codes, all information included in the OCLC-MARC record except the holdings data.

Bibliographic utility: an organization that makes a resource database of bibliographic records available for individual libraries to use in creating their local bibliographic records. Bibliographic utilities

currently in use in the United States include OCLC, Inc. and the Research Libraries Information Network (RLIN).

Cataloging center: a holding location or group of holding locations (or libraries) within each separately administered library that requires its own versions of cataloging records. A cataloging center is generally required when a holding location(s) does its cataloging independently. Complementary records are always maintained for each cataloging center, even if the bibliographic data is actually identical to another cataloging center's bibliographic data. The scope of a cataloging center can be defined by each separately administered library. For example, at UNC-CH, the Rare Books Collection is a cataloging center within the Academic Affairs Library.

Complementary records: different records with the same OCLC number that are maintained in the local master database to accommodate possible differences in practice between cataloging centers. Thus, if two cataloging centers have cataloged an item, two records will be maintained in the master database. In contrast, if two locations within a cataloging center have an item, only one record will be maintained in the master database, but it will show that both locations have the item. The bibliographic data in complementary records may in fact

be identical in some cases, but there is no way for the system to combine such records. If the number of them becomes unmanageable, the possibility of combining them can be considered. (cf: multi-use records)

Filing indicators: part of the coded information associated with certain fields to show the number of characters at the beginning of the field that should be disregarded in sorting. Such indicators must be used in order to avoid use of initial articles (e.g., a, an, the) for sorting when the language of the data being sorted may not be known.

Holdings data: as applied to the use of local transaction codes, the information entered into a specific field consisting of a holdings location followed by an enumeration of the copies, volumes, etc. held at that location. This pattern of a holding location and holdings statement is repeated for as many holding locations as required.

MARC format (MACHine Readable Cataloging): a machine-readable communications format for bibliographic data that was developed by the Library of Congress and has been accepted as a standard by the library community.

Master database or local master database: the complete, permanent machine-readable collection of master records

that is maintained for each institution.

Master record or existing local record: the presumably complete and correct permanent version of a record maintained in the master database for each institution. Once in the master database, a master record may be changed by re-editing the appropriate record in the OCLC database. This creates a new archive record which is then used to replace all or part of the master record for each cataloging center that catalogs an item (cf. Complementary records).

Multi-use records: local records representing items that are so closely related that OCLC does not distinguish among them. A system of numbers is used to indicate when multi-use records should be maintained and to distinguish between multi-use records. These numbers are assigned sequentially by cataloging staff as needed.

Separately administered libraries: libraries within an institution that have separate administrations such as UNC's Academic Affairs and Health Affairs libraries or Duke's Perkins Library System and Law Library. Separately administered libraries belong to OCLC under the same membership but have different institutional symbols in the OCLC union catalog. Separately administered libraries can be treated completely separately or as different cataloging centers.

Shelflist: the call-number ordered file of bibliographic records traditionally used by libraries for the master version of the holdings data. In addition to inventory functions, the shelflist is used to ensure that the same call number is not assigned to different items.