DOCUMENT RESUME

ED 205 180 IR 009 483

TITLE Requirements Statement for the Name Authority File Service.
INSTITUTION Council on Library Resources, Inc., Washington, D.C.
PUB DATE Apr 81
NOTE 97p.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Cataloging; Information Networks; Information Systems; *Library Networks; *Library Standards; National Programs; *Online Systems; Program Proposals; Quality Control; Systems Development
IDENTIFIERS *Name Authority Files

ABSTRACT

This publication specifies a set of general and technical requirements that must be met by the proposed Name Authority File Service in order to create an integrated, consistent name authority file for a nationwide library and information services network. Contributions from selected institutions operating online to a single computer-based system will provide the input to the file, which will be made available to the nation's libraries in a variety of formats. Chapters identify the purpose of the system, its functions, its administration and maintenance, its user-oriented system outputs (printed, computer-output microform, and machine-readable), and its user-oriented system inputs. The initial loading of data, interfaces with other systems, and system design constraints are described. Appendices provide a display of the LC/MARC communications format specified for use in the Name Authority File Service and the proposals for file quality control. (FM)

Requirements Statement for the Name Authority File Service

Task Force on a Name Authority File Service

Bibliographic Service Development Program

Council on Library Resources, Inc.
One Dupont Circle, NW, #620
Washington, D.C. 20036

April 1981

PREFACE

This document specifies requirements that must be met by the Name Authority File Service currently being developed for use by the nation's libraries. Planning for that service was prompted by a September 1979 meeting at which representatives of shared cataloging systems and major libraries met to discuss authority issues.

The following report, "An Integrated Consistent Authority File Service for Nationwide Use," is based on discussions at that meeting and was published in the Library of Congress Information Bulletin. It is presented here as background information for the requirements statement that follows.

From: Library of Congress Information Bulletin 39(28):244-248, July 11, 1980

APPENDIX

An Integrated Consistent Authority File Service for Nationwide Use

Bibliographic Service Development Program
Council on Library Resources, Inc.

[**Editor's Note:** The following report is an internal working document of the Council on Library Resources' Bibliographic Service Development Program. In the interest of giving wide dissemination to the ideas contained in the document, the Library of Congress has agreed to publish the report in the LC *Information Bulletin*. The Council would welcome any comments on the report and these should be addressed to C. Lee Jones, Program Officer, Council on Library Resources, Inc., One Dupont Circle, N.W., Washington, D.C. 20036.]

In recent months, many of those persons who have been involved in establishing pieces of the projected national library and information service network have become convinced that an important, indeed critical, component will be an integrated, consistent authority file. An authority file contains the words, terms, and cross references—the access points—that impose consistency on files of bibliographic data. Consistent files with clearly established access points are, of course, much easier for users to search. The authority file that these network planners have in mind would contain records contributed by a variety of institutions and created in accordance with established rules and procedures. It would be available for use nationwide. Building such a file would reduce the aggregate costs of authority work by the sharing of authority records; in addition, it would assist in the creation of more consistent bibliographic records, which, in itself, would improve current shared cataloging efforts.

Authority work is the most costly part of the cataloging process. While much cataloging information is now successfully shared, the means for sharing the most expensive product, the authority work, has not been established. This paper describes a project to establish, build, and maintain an authority file service.

The following pages contain a description of the present environment and basic concepts involved in implementing an authority file service. The rationale for the proposed approach is explained and the general requirements of the system are identified. A suggestion is made concerning the administration of such a service. Finally, a proposal for implementing the project is presented.

For purposes of clarity, definitions of terms used in this paper appear in an appendix.

Background

For most of their history, libraries generally have constructed their catalogs using two basic authority control files—one maintained at the Library of Congress (LC) and separate ones maintained within each library. When a library accepts the authority or access points in an LC bibliographic product, it relies upon the authority work carried out by LC. But when such a library has to catalog an item and create access points for which LC authority data cannot be found, it must establish an authority mechanism of its own.

Shared cataloging systems such as OCLC, Inc., RLIN (Research Libraries Information Network, which is operated by the Research Libraries Group), and WLN (Washington Library Network) provide an opportunity to build an integrated, consistent authority file system. However, until recently, discussions relating to such a development had been diffuse and not encouraging. In September 1979, a meeting of representatives from the major shared cataloging systems (OCLC, RLIN, and WLN), the Library of Congress, the National Library of Medicine, and the National Agricultural Library (the last three hereafter referred to as the national libraries), and the Council on Library Resources (CLR) took place. At that time a change in attitude, even a change in commitment, became apparent. Each institution was represented by a senior policy officer (with one exception), a senior computing expert, and a senior bibliographic policy officer. They agreed, in general, on four points:

(1) It is possible and desirable to establish and share an integrated, consistent authority file for nationwide use.

(2) It is possible and desirable to develop a set of procedures for building and maintaining such an authority file.

(3) It is possible to develop a set of general and specific requirements for an authority file service.

(4) It is possible to develop a set of design elements for authority control systems.

The September meeting yielded other points of agreement as well. First, LC should accept responsibility for the management of procedures for building and maintaining the file. Second, some agency, not necessarily LC, must physically maintain the authority file and provide access to it. Third,

some agency, probably not LC, but possibly CLR's Bibliographic Service Development Program (BSDP), should begin to stimulate the processes required to create an authority file service that will be available nationwide.

In October 1979, CLR's BSDP Program Committee went on record as supporting a concept embodied in a proposal, submitted by the Research Libraries Group (RLG) and the Washington Library Network, that would provide these two institutions with the capacity to build and maintain an integrated authority file. The Program Committee also encouraged CLR staff to continue work toward an integrated authority file.

The Present Environment

Authority work is carried on in many institutions but historically there has been no strategy to coordinate these efforts. It will be useful to note some of the major authority activities now under way.

The Library of Congress continues to produce high quality bibliographic records that are under authority control and are used in nearly every shared cataloging effort in this country and, to some extent, abroad. The authority control system at the Library of Congress is largely a manual system, although now the authority records produced are converted to machine-readable form. LC began sharing its machine-readable authority records in 1977, when the subject heading file was first distributed through the MARC Distribution Service. Name authority records have been distributed since 1978, and LC's machine-readable name authority file is available for on-line searching at LC. Only part of the total LC authority file is in machine-readable form, however. Many thousands of retrospective name authority records are still maintained in card form.

The other two national libraries, the National Library of Medicine (NLM) and the National Agricultural Library (NAL), each have authority files separate from that at the Library of Congress and from each other. In addition to the traditional authority conventions of the library community, NAL perceives certain obligations to conform to the authority conventions followed by the abstracting and indexing community. There have, however, been discussions among the three national libraries on developing joint authority files.

The major bibliographic utilities have approached the issue of authority control in different ways. OCLC does not now provide authority control over records in its bibliographic data base, although it does provide for its subscribers search only access to name authority records distributed by LC.

The Research Libraries Group's RLIN is committed to establishing an authority control system that will link the authority and bibliographic records in its files. Work on that system is under way with completion targeted for late 1980. A sophisticated off-line authority control system has been used by the New York Public Library (NYPL) to produce its own book catalogs for several years. This file will become part of the RLIN authority file.

The Washington Library Network is unique among U.S. networks in that its authority control system is on-line, and provides machine checks of new headings contributed to the data base. There is a manual verification process for all headings identified by machine as new entries into the WLN authority file. (The University of Toronto Automated Library System is a Canadian network with authority control characteristics that are similar to WLN.) As with other authority control systems, WLN's continues to evolve. Because both the NYPL and WLN authority systems preceded the creation of the LC MARC communications format for authority records, their authority formats are not fully compatible with that of LC.

In addition to the bibliographic utilities, there are some institutional systems that incorporate authority control, most notably those of the University of Chicago and the National Library of Medicine. There are also bibliographic services provided commercially that incorporate some degree of authority control over records. Again, to the extent that LC bibliographic records are used, these systems have an implicit control of access points. When records not generated by LC are used, access points may or may not be controlled, depending upon the system.

A common thread in the present bibliographic environment is the LC MARC bibliographic record and the consistency it gives other files because of its integrity. However, at present, none of the agencies with authority systems share authority data in an on-line mode with any other agency.

The Concept

The concept is a simple one: build and maintain an integrated, consistent authority file using contributions from selected institutions operating on-line to a single computer-based system. The contributing institutions and others will use the file for local (institution specific) processing requirements although the authority file itself will not carry institution specific data. The file will be made available to the nation's libraries in a variety of formats: for example, printed, microform, and machine-readable. In order to implement this new service, certain technical, procedural, and administrative issues must be resolved.

The institutions initially selected to contribute to the file will be large, general, research-oriented libraries (academic, public, special) because the capture of the authority work performed by a carefully chosen set of institutions will probably satisfy a large portion of the nation's authority requirements. Contributing institutions will include those reflecting the needs of users of public and state libraries as well as of academic and research communities.

Based upon the September discussions and upon the anticipated capabilities of the RLIN system, it is expected that the authority file will be built and maintained in an on-line mode at RLIN with LC's participation. The RLG/WLN authority control project now under way should provide the background and capabilities to permit RLIN to assume these file management responsibilities. This same project, incorporating as it does the authority work and experience of WLN, may serve to include at a very early stage the needs of public and state libraries.

Though there are both name and subject authority records, this project will be limited to name authorities, including uniform titles and series. The issue of subject authority records will be addressed at a later date.

The LC name authority file will be the base upon which the project is established. There are several other high-quality authority files that may be merged with the LC file including those of NYPL, WLN, Stanford, and Chicago. The prospects for merging these files are being examined by LC, NYPL, and RLG.

LC, in cooperation with the contributing institutions, will establish procedures required for the input and maintenance of authority records. LC will provide necessary training required at each institution to implement the input and maintenance procedures. Since manual authentication of new records is costly and time-consuming (witness the CONSER experience), a method of quality control using machine verification will be explored with sampling techniques employed to test results. Because the objective is to build a single, consistent authority file, and since there are bound to be disagreements between contributing institutions, some mechanism must be found to resolve problems. LC, as the quality control agent, might reasonably assume that role.

RLIN, as the authority file system manager, will be responsible for all aspects of technical performance, such as the systems and computer programs for input, machine validation, and distribution, and will work with the contributing institutions and others to meet requirements for display formats and record access. The authority file which results from this project will be designed so that it can be used by libraries and networks for a variety of library processes such as cataloging and reference.

Once the project is under way, it will be possible to expand purposefully the number of institutions participating in file development and maintenance, either directly or through one of the on-line participants. As examples, LC now has cooperative arrangements with Northwestern University, The Texas State Library, and the Government Printing Office, which are assisting in the establishment of authorities in specific areas.

So far, we have focused on an authority file service. As time and experience permit, it may be possible to expand the work to include more sophisticated procedures leading toward a nationwide authority control system. Such a system would, in effect, link authority records to the specific bibliographic records in which the established terms are used. These procedures need not preclude the development of individual or local authority control systems nor prejudice their continuing usefulness once a nationwide system was in place. The problems and benefits of linking multiple authority control concept are described in more detail in the following section.

General Requirements

If the developments of an authority file service for nationwide use is to proceed in an orderly fashion, it is necessary to specify its functional requirements. The work may logically be separated into phases. As noted earlier, establishing and maintaining an integrated, consistent authority file is the initial concern of the current project and constitutes phase one of the work required for the total system. The following requirements one through four relate to this first phase. Requirement five relates to the use of the authority file and constitutes phase two. The final requirement explicitly relates the authority file to bibliographic files, a relationship that would result in an authority control system. Because, on a nationwide basis, this would require sophisticated file-relationships between an authority file physically housed in one system and bibliographic data bases distributed among many systems, it is considered a separate (and distant) phase three. Before phase three, individual systems may use the authority file in conjunction with their own system capabilities to link authority and bibliographic records. Requirements five and six are less well understood at this time in a nationwide context.

(1) *Collect authority data.* The integrated authority file is the instrument in which authority data from a variety of contributing sources are collected. Computer hardware and software, bibliographic standards

governing the intellectual content of records, and procedures for adding records to the file must be developed. These elements will permit the creation of the authority records and collection of these records into a single, consistent, logical file housed at a single site. The procedures must ensure conformance to standards pertaining to the bibliographic data and to its presentation (format). Though substantial machine checks are envisioned, sampling techniques must be devised to monitor the verification and validation of additions to the authority file for quality control purposes. By definition, sampling does not provide a 100-percent confidence level in the quality of records in the file. But, if a satisfactory confidence level can be achieved by sampling techniques it is unlikely that comprehensive manual verification will be required. Samples will likely be drawn from records submitted by each contributing institution in order to monitor record quality relative to source of contribution.

(2) *Maintain the authority file.* Authority files are dynamic; they grow in number and change in content as authority work progresses. Thus they are best maintained on-line. Technical facilities and bibliographic procedures for creating, correcting, replacing, and updating records and establishing the relationships that exist between records are all required. Adequate data security mechanisms are also necessary to protect data in the file, that is, to prevent unauthorized sources from adding or modifying records. Machine editing and search capabilities should support the use of the authority file, using sophisticated techniques for automatic alerting at the terminal of errors and record inconsistencies, and automatic checking of new records against the existing authority records in the file. In later project phases, these procedures could also be used to validate headings in bibliographic records.

(3) *Provide on-line access to authority data.* Institutions will have on-line access to the authority file through a variety of mechanisms. It is anticipated that contributing institutions, which will be permitted to add and maintain records, will also use the file for cataloging and reference activities. Other institutions may have on-line access for cataloging and reference purposes provided they are members of RLG, are linked in some other way to RLIN, or have access to any other shared cataloging system that provides access to the authority file. Access to the authority file will be via standard communication protocols, such as the library applications level protocol (NCLIS/NBS) for communication between computers, currently being refined by ANSC Z39 (American National Standards Committee Z39).

(4) *Provide off-line access to authority data.* Because not all libraries, vendors, and utilities are likely to have on-line access to the authority file, the file will also be made available in other forms such as printed and microform formats, as well as in machine-readable forms such as magnetic tape. Therefore, the technical and bibliographic specifications and procedures to produce these products are required.

The following general requirements go beyond the building of an authority file and deal with the validation of access points in bibliographic records and authority control. These requirements are viewed as future enhancements of the authority file service.

(5) *Verification/validation of entries in bibliographic data.* Entries in bibliographic records will be checked against or derived from the authority file to insure that they are established as headings in that file. They may also require access to at least a portion of the bibliographic record(s) in which the heading has previously been used.

(6) *Establish links between authority records and bibliographic records.* The development of a capability to create and maintain links between authority records and the bibliographic records in which these headings (access points) are used will provide the capacity for consistent access to bibliographic records, the production of a variety of catalogs, and the on-line cataloging process. The system would assure that only established forms from the authority file are used in bibliographic records. This requirement is included as a long-range requirement and may be difficult to implement on a nationwide basis.

Administration

Because this project will involve several organizations each contributing in different ways, a well-defined management structure is essential. The management plans should reflect the cooperative nature of the undertaking and stress effective coordination of effort and continuing consultation among participants. Decisions in two basic areas will be required: those relating to the technical development and maintenance of the file (that is, the hardware, computer programs, screen specifications, and product specifications), and those concerning the bibliographic policy, standards, and procedures, (that is, participants, training, institutional interrelationships, rules, and guidelines).

As the projected manager, RLG must necessarily be accountable to participants for technical management of the authority file since it will physically house and maintain the file and since it is developing internal capacities that allow for the acceptance of this external responsibility. Though RLG will, by virtue of its role in managing the file, have immediate access to it, provision will be made to assure that all other

shared cataloging services and non-RLG institutions have timely, possibly on-line, access to the file. Because of the Library of Congress's position as principal contributor to the file and its current role in analyzing and disseminating information on bibliographic standards, LC has agreed to assume management responsibility for bibliographic policy, standards, and procedures. This alignment of responsibility is consistent with views of participants at the September 1979 authority issue meeting.

Once the authority file system becomes operational, representatives of the contributing institutions, again expected to be a small but diverse group of large, general, research-oriented institutions (academic, public, and special libraries), will form an advisory group to the project. These institutions will take an active part in advising on the continuing growth and maintenance of the system. RLG and LC will coordinate the input from these participants, act as arbiters in decision making, and have responsibility for implementation in their respective areas.

While this project is focused on the needs of American libraries, Canadian observers may well be invited to participate. Insofar as possible, the system should be developed so that it can take its place among other national authority systems as efforts to link bibliographic activities world-wide move forward.

Implementation Guidelines

As part of its Bibliographic Service Development Program, CLR will appoint a task force to assist in the continuing planning and review required to establish the authority file service. This task force, which is distinct from the advisory group mentioned above, will cease to exist once the system becomes operational and the Advisory Committee begins to function.

The task force will be charged with several tasks:

(1) Review this concept paper with special attention to the general requirements;

(2) Develop necessary specific requirements for an integrated, consistent name authority file service expanding particularly upon general requirements 1-4 above; consultants may be employed to assist the task force and CLR staff;

(3) Define the specific tasks required to implement the project.

While these tasks will lead to the establishment of the file, there are other issues that need attention. The task force will address, with LC and the file manager, such factors as selection of contributing institutions, training, forms of access, frequency of updates, screen displays, and products. How will the file be distributed? Will there be charges for access? How will the file manager support operating costs? How will institutions not associated with a shared cataloging activity gain access to the file? How will other libraries contributed to the authority file?

Finally, the task force will have to focus on the issue of administration/governance. Since this is only one element in a much larger milieu, how will this issue be resolved for the authority file service?

Completion of this agenda should result in the desired product—an operating, growing, integrated, consistent name authority file service for U.S. libraries and those who may join them.

Appendix

Definitions

For purposes of clarity it will be useful to define a number of terms as they will be used in this project. The terms are arranged in a logical rather than alphabetical sequence.

Catalog: A set of bibliographic records under control of authority files which describe a set of resources contained in collections, libraries, networks, and so forth. It is the instrument by which bibliographic control is maintained and by which the relationship between individual bibliographic records can be indicated, for example, the works of a single author or works on a particular subject. The catalog may include other types of records as well, such as cross-references and on-order information.

Bibliographic Control: The functions necessary to generate and organize cataloging records of library materials for effective retrieval by name, subject, and so on. Access points such as names and subjects must be consistent within a data base. Authority control is the particular function that provides that consistency.

Authority Control: Establishment of logical links between the authority and bibliographic files, that is, between individual authority records and all bibliographic records in which the authority (heading or access point) is used.

Authority File: A set of records that identify the established forms for headings (names, subjects, and so on) or access points. An authority file includes established forms of headings used in individual institutions, groups of related institutions, or networks of related and/or unrelated institutions. Authority files include cross-references from variants to the established forms for headings and links from earlier to later forms; they may link broader and narrower terms and related forms.

Authority Work: The functions necessary to establish, maintain, and use authority files.

Consistency (Authority File Context): Each heading (entry) in the authority file is created and carried as a unique element of the authority file and is therefore consistent (not in conflict) with other records in the authority.

System: An assembly of components united by some form of regulated interaction to form an organized whole. A system can be designed to perform any function, for example, to build an authority file.

File: A collection of related records treated as a unit.

Physical File: A collection of related records resident in a common physical environment. All of the file resides in one location.

Logical File: A collection of related records independent of their physical environment. Portions of the same logical file may be located in different physical locations.

ISSN 0041-7904 *Key title:* Library of Congress information bulletin

TABLE OF CONTENTS

CHAPTER I

OVERVIEW AND PURPOSE

The purpose of this document is to specify the set of requirements that must be met by the Name Authority File Service.

The purposes of the Name Authority File Service are:

- o to collect and maintain authority data for names, titles, and series
- o to record and maintain the relationships between and among headings for names, titles, and series; these relationships will include relationships
  - between and among established headings
  - between variant heading and established form(s)
- o to ensure integrity of heading forms
- o to provide query access to authority data

In addition, the Name Authority File Service must incorporate data definitions identical to or compatible with those specified in Authorities: A MARC Format[1] and its addenda, and must accept data in and produce data for distribution in this format.

---

[1] United States. Library of Congress. MARC Development Office. Authorities: A MARC Format; Specifications for Magnetic Tapes Containing Authority Records. Preliminary edition. Washington: Library of Congress, 1976. 67 p. ISBN 0-8444-0219-2.
_____ Addendum, no. 1 -

The Name Authority File Service must be able to be used by a large number of institutions, although some functions (e.g., ADD AUTHORITY RECORD) will be restricted to a limited number of institutions designated as contributing sources. Contributing sources will be responsible for the creation and maintenance of data in the file.

CHAPTER 2

FUNCTIONS

The purpose of this chapter is to outline the functions defined for the Name Authority File Service and to provide a brief description of each.

The setting in which these functions will generally be performed is technical processing, specifically cataloging or catalog maintenance, although some functions will also be available to non-technical processing personnel. The functions described here will

- permit addition of an authority record
- permit changing an existing authority record
- permit deletion of an existing authority record
- ensure integrity of heading forms and other authority data
- enable access to authority data
- facilitate maintenance activities on authority data

2.1 ADD AUTHORITY RECORD

The function Add Authority Record

- accept input of authority records from a contributing source

- o accept only authority records in agreed upon formats equivalent to or compatible with content and content designation specified in Authorities: A MARC Format and its Addenda[1]

- o accommodate authority records using characters defined to the LC/MARC character set, as specified in Appendix I -- LC/MARC Character Set

- o record and maintain the data in an authority record so that to the extent required the data supplied by different contributing sources may be identified

- o accept input of authority records subject to system security requirements

Operators at terminals will add authority records through online, interactive use of the system. The system should also accommodate computer-to-computer transmission of authority records (from external sources to the host system, from the host system to external sources).

2.2 CHANGE AUTHORITY RECORD

The function Change Authority Record must:

- o accept change of authority records from more than one contributing source

- o accept only changes to authority records in agreed upon formats equivalent to or compatible with

[1] See footnote 1, page 1

content designation specified in Authorities: A MARC Format and its Addenda[2]

- o accommodate changes to authority records using characters defined to the LC/MARC character set, as specified in Appendix I - LC/MARC Character Set
- o record and maintain the changed data in an authority record so that to the extent required the data changed by different contributing sources may be identified
- o accept change of an authority record subject to system security requirements and the conditions specified below

Operators at terminals will change authority records through online, interactive use of the system. The system should also accommodate computer-to-computer transmission of changes (from external sources to the host system, from the host system to external sources).

Conditions for Changing an Authority Record

A. Adding New Information to an Existing Authority Record

An existing authority record may be changed by the addition of new information under conditions that would include:

1. See from references (LC/MARC Authorities tag 4nn) may be added by any contributing source at any

[2] See footnote 1, page 1.

time, provided that the addition does not conflict with data already in the file.

2. see also from references (5nn) may be added by any contributing source at any time, provided that the addition does not conflict with data already in the file.

3. cataloger's notes (668, 670, 675, 676) may be added by any contributing source at any time.

4. any other data (i.e., data not covered by (1), (2), or (3)) may be added only after all contributing sources have had the opportunity to review and comment upon the proposed addition. (cf. subsection C below.)

B. Changing Information In or Deleting Information From an Authority Record

An existing authority record may be changed by the changing of or deletion of existing information under conditions that would include:

1. all data <u>except</u> <u>cataloger's notes</u> (670, 675, 676) may be changed or deleted <u>only after</u> all contributing sources have had the opportunity to review and comment upon the proposed change or deletion. (cf. subsection C. below.)

cataloger's notes (668, 670, 676 ...) may be changed or deleted by any contributing source at any time.

Such changes or deletions must be ... with rules agreed to by contributing sources

C. Update Review by Contributing Sources

The host system must support a machine-readable "routing" or "mail" facility that identifies records for which update review is requested.

When a change requires update review, the record will be flagged for review for several days, with the understanding that:

1. if there is no objection to the proposed addition, change or deletion, the addition change or deletion will automatically be made by the system after the appropriate review period has elapsed.

2. if any contributing source objects to the proposed addition, change or deletion, the proposal and objections will be forwarded for Conflict Resolution. (cf. subsection D, below.)

This flagging and review is referred to as the "proposed change cycle."

D. Additional, Special Criteria for Some Changes by the Library of Congress[3]

As the conflict resolution agency, the Library of Congress will have some special, but limited, prerogatives for certain types of changes:

1. The Library of Congress will be permitted to change some content designations without using the

[3]See Chapter 3 -- Administration and Maintenance for discussion of the role of the Library of Congress as Conflict Resolution Age...

"proposed change cycle." The Library of Congress will be able to:

a. correct tag use within 1nn, 4nn, or 5nn (that is, change a tag from 100 to 110, but not change a tag from 400 to 500).

b. correct tag use of cataloger's notes (670, 675, 676).

c. correct indicator use.

d. correct subfielding and use of subfield.

2. The Library of Congress will be permitted to change use of punctuation in established headings 1nn) or references (4nn or 5nn) without using the "proposed change cycle."

3. The Library of Congress will be permitted to change certain information about references in the Library of Congress Name Authority Data used as the Base File in the Name Authority File Service.

In November 1980, the Library of Congress "flipped" its Name Authority File Data. This "flip" changed any record that contained a see reference (4nn) that was the established form under AACR-2. This "flip" resulted in the following:

- the AACR [illegible] (4nn) became the established heading (1nn);

- the pre-AACR-2 established heading (1nn) became a see reference (4nn)

As part of the processing that performed the "flip" the Library of Congress Name Authority Data was scanned to identify references that had not been evaluated previously by LC staff to assess their conformity to AACR-2. Records containing references determined to be not evaluated were flagged as "not evaluated."

When the Base File for the Name Authority File Service is created with the Library of Congress Name Authority Data there may remain some records flagged to show that references were "not evaluated."[4] Using the Name Authority File Service, the Library of Congress may review these references, and upon evaluation LC will be permitted to remove the "not evaluated" flag without going through the "proposed change cycle."

Any such content designation, punctuation, or "not evaluated" reference changes made by the Library of Congress will have the effect of identifying the changed record for redistribution.[5]

Other contributing sources that wish to make content designation, punctuation, or "not evaluated" reference changes must use the "proposed change cycle" (described above at subsection C). Upon its review of "proposed changes," the Library of Congress may elect to make immediately those changes that are proposed for

[4] It is possible that at the time the Base File is created there will be no records in the Library of Congress Name Authority Data that are flagged as "not evaluated."

[5] Records would be flagged for redistribution so that products supplied by the system would be current.

content designation, punctuation, or "not evaluated" references which it is permitted to make without using the "proposed change cycle."

For all other types of changes -- those described at subsections A and B, above -- the Library of Congress will follow the same routines and procedures as other contributing sources.

2.3 DELETE AUTHORITY RECORD

The function Delete Authority Record must:

- o accept deletion of authority records from more than one contributing source
- o accept deletion of an authority record subject to system security requirements and the conditions specified below.

Operators at terminals will delete authority records through online, interactive use of the system. The system should also accommodate computer-to-computer transmission of deletions (from external sources to the host system, from the host system to external sources).

Conditions for Deleting an Authority Record

A [illegible] authority record [illegible] [illegible] [illegible] been flagged for review through [illegible] proposed change [illegible]. The record for which deletion has been proposed will be flagged for review for several days, with the understanding that

1. if there is no objection to the proposed deletion of the record, the system will automatically delete the record after the appropriate review period has elapsed.

2. if any contributing source objects to the proposed deletion of the record, the proposal and objections will be forwarded for Conflict Resolution (cf. Chapter 3 -- Administration and Maintenance -- in this document).

2.4 EDIT AUTHORITY DATA

The function Edit Authority Data provides machine[6] editing and quality control of data received by the system through the functions:

2.1 Add Authority Record

2.2 Change Authority Record

The host system should provide as much quality control through system software as can be demonstrated to be useful and practical. System-supported editing and quality control should be provided online for online, interactive use of both of the above functions.

It may, however, be appropriate or necessary to perform some system-supported editing or quality control through scheduled batch processing.

[6]For administrative quality control see Chapter 3 -- Administration and Maintenance.

Machine editing and quality control of authority data must be such that

- o duplicate established headings cannot be added to the file
- o headings may not be changed in or deleted from one record unless related records are also altered
- o records may not be deleted from the file unless related records are updated with necessary changes or deletions

Some specific suggestions for machine editing and quality control of data in the Name Authority File are found in Appendix II -- Proposals for Quality Control for the Name Authority File Service.

2.5 QUERY AUTHORITY DATA

The function Query Authority Data must support query capabilities for headings in authority data and for other, non-heading authority data. Query capabilities for these two general types of authority data are discussed below at:

A. Query of Heading Data

B. Query of Other, Non-Heading, Authority Data

The function Query Authority Data must support operations (AND, OR, NOT) in search statements and search criteria in these statements.

Operation at terminal [illegible] Name Authority Data will [illegible] use of the system.

The system should also accommodate computer-to-computer transmission of queries (from external sources to the host system, from the host system to external sources).

Some mechanisms for scheduled, batch searching should also be provided. These mechanisms should process standing search requests and/or search requests entered by operators at terminals saved for scheduled batch processing.

A. Query of Heading Data

The following types of search criteria must be supported for query of heading data:

1. Query using right truncated value of a heading as criteria.

2. Query using words or words from a heading as criteria.[7]

   a. Query using any word or words from a heading as criteria.

   b. Query using any right truncated word or words from a heading as criteria.

   c. Query using a combination of words(s) and right truncated word(s) from a heading as criteria.

[7] Word searches should not incorporate [illegible] "[illegible]is" [illegible] words however may be designated as "[illegible]ela," o[illegible] "[illegible] words, [illegible]vailable for searches at a lower priority online or available for [illegible] batch searching.

In addition, the query capabilities for heading data must provide the following types of access[8] to authority data by the ability to search for:

1. Any Heading in the Name Authority File
2. Headings by type
   a. Personal Names (LC/MARC Authorities tag, 100, 400, 500)
   b. Corporate Name (110, 410, 510)
   c. Conference Names (111, 411, 511)
   d. Uniform Titles (130, 430, 530)
   e. Geographic Names (151, 451, 551)
   f. Topical subjects (550)[9]
3. Title Information
   a. Uniform Titles (130, 430, 530)
   b. Titles in name title headings (subfields "t", "n" and "p" in any heading)
   c. Any title (combination of 3 a and 3 b above)

[8] These types of [illegible] the first priority.

[9] The file may contain see [illegible] supply subjects (550), in agreement with [illegible] practice to supply such references under certain circumstanc

4. Headings by role in authority data

   a. Any Established Heading (1nn)
   b. Any See From Reference (4nn)
   c. Any See Also From Reference (5nn)

5. Headings by Role in Authority Data and by Type

   a. Established Headings

      - Personal Names (100)
      - Corporate Names (110)
      - Conference Names (111)
      - Uniform Titles (130)
      - Geographic Names (151)

   b. See From References

      - Personal Names (400)
      - Corporate Names (410)
      - Conference Names (411)
      - Uniform Titles (430)
      - Geographic Names (451)

   c. See Also From References[10]

      - Personal Names (500)
      - Corporate Names (510)
      - Conference Names (511)
      - Uniform Titles (530)
      - Geographic Names (551)
      - Topical Subjects (550)

[10] See footnote 5 page 11

6. Headings by Other Characteristics

   a. Personal Name Surname and Forename(s) Initial(s)

   b. Names by Form of Name Indicator values, e.g., capability to specify a search for personal names that are forenames, that are single surnames, that are multiple surnames, or that are family names

B. Query of Other, Non-Heading Authority Data

The following types of search criteria must be supported for query of other, non-heading authority data:

1. Query using right truncated value as criteria

2. Query using word or words as criteria[11]

   a. Query using any word or words as criteria

   b. Query using any right truncated word or words as criteria

   Query using a [illegible] and right truncated word(s) as criteria

   [illegible] using complete [illegible] as criteria

The types of search criteria supported for query of other, non-heading data will be those appropriate to the specific type of data.

A review of non-heading authority data suggests that the fields or data elements in Table 1 are logical candidates for access.

TABLE 1. Fields or Data Elements Which are Logical Candidates for Access

Identifier	Name
LC/MARC Authorities tag 001	Authority Record Control Number*
008, pos. 9	Authority/Reference Record Descriptor**
008, pos. 14	Heading Use: Main or Added Entry Aspect of Heading**
008, pos. 15	Heading Use: Subject aspect of Heading**
008, pos. 16	Heading Use: Series Aspects of Heading***
008, pos. 28	Government Agency**
010	Library of Congress Card Number*
035	Local System Number*
043	Geographic Area Code***
045	Chronological Coverage Code***
052	Geographic Classification Code***
[not determined]	Flag to identify records that are in "proposed change cycle"
[not determined]	National standard authority record control number (under development)

* Type of search: complete value

** Type of search: complete value; consider as secondary or limiting search

*** Consider providing access if use of fields is implemented

CHAPTER 3

ADMINISTRATION AND MAINTENANCE

The purpose of this chapter is to describe the procedures, methods and responsibilities of the host computer system for the contributing sources to the Name Authority File Service.

The host computer system, Research Libraries Information Network (RLIN), will be responsible for physical maintenance of the data, the associated applications and system software, and the system hardware. For additional discussion of host system responsibilities, see chapters 2 (Functions) and 8 (System Design Constraints).

The contributing sources will be responsible for maintaining data content of records. The procedures and methods used for this maintenance will reflect and encourage a sense of shared responsibility by all contributing sources for the content and quality of the data. When conflicts occur about the content of or the quality of the data, they will be resolved on a timely basis. A regular cycle for conflict resolution will be provided.

To assist contributing sources in resolving conflicts, the Library of Congress will be designated as the conflict resolution agency.

To assist all contributing sources in data maintenance and in conflict resolution, an electronic "mail" or "routing" service will be used. This service will provide a vehicle for:

- one contributing source to query another about data the latter had created

- one contributing source to communicate to all other sources proposed changes to existing records or deletions of existing records

- any contributing source to comment upon proposed changes or deletions, and to review comments of other contributing sources

CHAPTER 4

USER ORIENTED SYSTEM OUTPUTS

The purpose of this chapter is to describe the user oriented system outputs for the Name Authority File Service.

The following user oriented system outputs have been identified and are described in the subsections that follow:

4.1 Displays

4.2 Printed Products

4.3 Computer-Output-Microform Products

4.4 Machine-Readable Products

4.1 DISPLAYS

Displays will be supplied by the system in response to these functions:

2.1 Added Authority Record[1]

2.2 Change Authority Record[1]

2.3 Delete Authority Record[1]

2.4 Query Authority Data

[1]For additional discussion of displays for these functions, see Chapter 5 -- User Oriented System Inputs.

In response to these functions, at least two general types of displays must be provided:

(1) List of headings, records, etc., satisfying the request (search) criteria

(2) For authority data:

(a) Full record display
- with content designation
- without content designation

(b) Brief record display

(c) Tailored displays (e.g., user may specify the data to be displayed)

Any display should include information sufficient for the user to be aware of work progress -- for example, a display following a query should convey to the user what information was input as the request (e.g., search criteria, display specifications).

Any display provided in response to a given function should be appropriate to that function -- for example, a request to Change Authority Data should return a full record display of authority data with content designation.

List displays should arrange information in a meaningful and logical order (lists of headings, for example, will be in alphabetic order by heading.) Content designated displays should arrange data in agreed upon order by tag. Displays not using content designation may incorporate the use of labels (e.g., see from references in authority data may be identified by labels such as VARIANT NAME.)

At the time of initiation of a request for which the system will return a display, the user must also be able to specify the desired type of display. System default displays must also be provided, so that the user needs to request a certain display only if it is different from the system-supplied default.

Displays will be returned to operators at terminals; these terminals may be cathode-ray-tube (CRT) terminals or hard copy terminals. Displays must incorporate use of the LC/MARC character set (cf. Appendix I).

4.2 PRINTED PRODUCTS

Printed products will be supplied by the system through printers attached to cathode ray tube (CRT) terminals or printing (hard copy) terminals, and at remote printers as part of scheduled batch processing.

Printed products produced at terminals will typically be limited to printing of displays.

Printed products produced through scheduled batch processing would include reports resulting from the performance of scheduled batch processing activities (e.g., performing query through scheduled batch processing) and management reports, system performance statistics, etc.

4.3 COMPUTER-OUTPUT-MICROFORM (COM) PRODUCTS

The system will support the production of computer-output-microform (COM) products of authority data. Such products will conform to existing national and international standards with respect to header information, reduction ratios, etc.

The principal COM product (probably to be issued as microfiche) is viewed as an expansion of the current Library of Congress product, LC Name Authorities. This Library of Congress COM fiche product contains Library of Congress authority data; the COM fiche product from the Name Authority File Service would include authority data supplied by all contributing sources to the Name Authority File.

Other COM products may also be produced periodically by the system, based upon schedules determined by users. COM products may also be created upon demand.

4.4 MACHINE-READABLE PRODUCTS

Two types of machine-readable products will be supplied by the system:

1. machine-readable records in appropriate Library of Congress Authorities communication format.[2] Such records will be produced in the required medium: tape, disk, etc., or for direct computer-to-computer communication.)

2. machine-readable data in format(s) to support the production of computer-output-microform (COM) products.

Machine-readable products will be supplied periodically by the system, based upon schedules determined by users. Machine-readable products may also be created on demand.

---

[2] See footnote 1, page 1.

CHAPTER 5

USER ORIENTED SYSTEM INPUTS

The purpose of this chapter is to describe the user oriented system inputs for the Name Authority File Service.

The following user oriented system inputs have been identified and are described in the subsections that follow:

5.1 Displays

5.2 Workforms

5.3 Other

5.1 DISPLAYS

Formatted displays for cathode-ray-tube (CRT) terminals will be required for functions which the operator may use to add, change, or delete data, or to add or delete records.

Such displays should incorporate, whenever possible, prompts and tabbing to assist the operator in data addition, change or deletion.

The host system will provide displays sufficient to meet the functional requirements of the Name Authority File Service.

5.2 WORKFORMS

Formatted paper worksheets will be required for functions for which coding is required or desired away from terminal. The host system will supply formatted worksheet displays to CRT terminals (cf. 5.1).

5.3 OTHER

Under certain circumstances, the system will need to provide other types of user oriented system inputs in addition to displays and workforms. These other, user oriented system inputs will typically be produced as reports as part of system-supplied quality control (for example, if some machine editing is performed as scheduled batch processing, reports produced that identify errors, problems, etc., would be used by operators to identify records and data requiring attention).

CHAPTER 6

INITIAL CREATION (LOADING) OF DATA

The Base File for the Name Authority File Service will be the Library of Congress name, title and series authority records that are for established headings that conform to the Anglo-American Cataloging Rules, second edition, and that are available in machine-readable form when the Name Authority File and the Name Authority File Service are implemented.

CHAPTER 7

INTERFACES WITH OTHER SYSTEMS

The Name Authority File Service will be designed so that it may utilize data from other systems through computer-to-computer interconnection, following national and international standards for communications protocols and for data content and content designation (e.g., Library of Congress MARC communications formats).

CHAPTER 8

SYSTEM DESIGN CONSTRAINTS

The purpose of this chapter is to identify and describe briefly the performance objectives and operating requirements for the Name Authority File Service.

The following topics are discussed in sections below:

8.1 Performance Objectives for Functions

8.2 Data Currency Requirements

8.3 Data Security Requirements

8.4 System Backup

8.5 System Software and Hardware

8.6 Training

8.7 System Growth Requirements

8.1 PERFORMANCE OBJECTIVES FOR FUNCTIONS

All functions -- online, interactive or scheduled batch -- must be provided in a timely fashion.

The host system must provide an average response time of less than 2-3 seconds for online, interactive functions; response time will be measured from the initiation of a function to the beginning of system response. Initial system response to complex

searches, such as some word searches, may be data satisfying the search criteria or indication that such response will be delayed. (Whenever the system is aware that the response satisfying the request will be delayed because of system load or the nature of the request, the system should notify the user of this condition.)

System response to a request for scheduled batch processing should occur within 24 hours, unless the requestor indicates that a longer time before response is acceptable.

8.2 DATA CURRENCY REQUIREMENTS

The Name Authority File and the Name Authority File Service must be available through the host system for online, interactive use maximum available time, seven days a week, except for scheduled and necessary system maintenance.

The results of any online, interactive function that updates data in any file (that is, that adds a new record, or changes or deletes an existing record) must be immediately available to any user at any terminal by all access points.

8.3 DATA SECURITY REQUIREMENTS

Data security for all files and data will be provided such that only authorized personnel from authorized institutions using authorized commands (language) may activate system functions. The system will also ensure that restrictions on the provision of functions are observed.

At least the following data and file security measures must be provided:

1. Proper operation and backup, including restart/recovery, for the online system.

2. The addition of new records, and the change or deletion of existing records should be restricted to institutions designated as contributing sources; certain functions will be restricted using appropriate file security techniques. Other institutions may have access only to read the files.

3. An existing record should be "locked" when changes are being made to it. This will prevent more than one contributing source from updating the same record at the same time.

4. Changes to existing records should be monitored if not restricted:

    a. Certain fields or elements within fields should not be altered. For example, information may be added to a field but not deleted -- the LC/MARC tag 040 -- Cataloging Source is, for example, such a field.

    b. Certain fields should not be deleted -- the LC/MARC tag 008 is, for example, such a field.

8.4 SYSTEM BACK UP

The host system supporting the Name Authority File Service, Research Libraries Information Network (RLIN), must provide

proper operation and back up, including restart/recovery procedures.

8.5 SOFTWARE AND HARDWARE

Communications software: The whole system should support query access through telephone dialup, one or more of the value-added communication networks, and leased communication lines.

Communications hardware: The whole system may require users of the system to obtain specific hardware and lease communication lines for communication with the system in full-face block transmission mode.

Query only access should be supported from any ASCII computer terminal in line-by-line transmission mode (CRT or hard copy). The whole system may limit full-face block mode to a designated terminal or terminals.

8.6 TRAINING

The Library of Congress will be responsible for providing a training program for contributing sources.

As part of the training program, contributing [illegible] receive documentation describing applications available, methods of system use, and conventions to be followed by contributing sources

As required, the staff of the [illegible] Libraries Information Network (RLIN), will participate in the training program. RLIN staff will provide system training programs for institutions that are not contributing sources but are users of the Name Authority File on the host system.

8.7 SYSTEM GROWTH REQUIREMENTS

The name Authority File Service must be designed for the purposes outlined in this document and to meet the functions outlined herein. Use of all functions is expected to increase over time; the system design must be such that future expansion may be accommodated.

APPENDIX I

LC/MARC CHARACTER SET

The character set in this appendix is the extended set defined for use in the LC/MARC communications format and specified for use in Name Authority File Service displays (cf. section 4.1 of the basic document). It is reprinted here from MARC Formats for Bibliographic Data Washington: Library of Congress, Automated Systems Office, 1980.)

APPENDIX III.B.INTRO. CHARACTER SET - INTRODUCTION

MARC tapes follow the standard character set, American Standard Code for Information Interchange (ASCII), adopted by the American National Standards Institute for information interchange on magnetic tape. This standard consists of a 7-bit code; however, the MARC character set has been expanded to an 8-bit code for a 9-channel tape and a 6-bit code for a 7-channel tape. The basic MARC code set is 8-bit.

A. EIGHT-BIT CODE

Some of the standard ASCII characters, such as braces, are not part of this character set; however, no characters have been substituted for these code positions. Other characters such as diacritical marks will be left in their standard (unused) positions and duplicated in another portion of the code set reserved for special characters and diacriticals.

B. SIX-BIT CODE

This code set is derived by removing the sixth and eighth bits from the 8-bit code set. The standard 6-bit set includes lowercase alphabetic characters, numerals, most punctuation marks, and other special characters. Three charactees in the standard set are designated as nonlocking shift codes of which only two (73_8 and 75_8) are used at present. The presence of one of these codes indicates that the next character is in either one of two nonstandard 6-bit code sets. Code 73_8 precedes characters in non-standard set I, and code 75_8 precedes characters in non-standard set II.

C. DIACRITICAL MARKS

Diacritical marks intended to go either over or under an alphabetic character will precede the character to which they belong.

D. EXTENDED LIBRARY CHARACTER SET

Although the present ASCII code configuration could accommodate some additional graphics, there are too few unused positions to provide enough codes for all future needs, i.e., complete character sets for Greek, Arabic, Hebrew, etc. For this reason, Greek, subscript, and superscript characters have been placed in separate character sets. These sets will be indicated by a locking escape sequence, consisting of the escape character, ESC (18_{16} 33_8), followed by a single lowercase alphabetic character. The following escape sequences will be used:

ESCs=Standard 8-bit set

ESCg=Greek set

ESCb=Subscript set

ESCp=Superscript set

All records begin in the standard set. When an escape is made to another character set, all characters following the escape sequence will be interpreted as being part of the variant character set until another escape sequence is reached or the end of the record is reached. The same escape sequence and conventions used in the 8-bit code will be used in the 6-bit code. The 6-bit shift character 73_8 will be used to reach the ESC character.

APPENDIX III.B.1. CHARACTER SET - PART 1: EXPANDED 8-BIT ASCII FOR MARC RECORDS

Dec.	Hex	Binary	Graphic	Name and/or Function
0	00	0000 0000		Null
1	01	0000 0001		Start of heading
2	02	0000 0010		Start of text
3	03	0000 0011		End of text
4	04	0000 0100		End of transmission
5	05	0000 0101		Enquiry
6	06	0000 0110		Acknowledge
7	07	0000 0111		Bell
8	08	0000 1000		Backspace
9	09	0000 1001		Horizontal tabulation
10	0A	0000 1010		Line feed
11	0B	0000 1011		Vertical tabulation
12	0C	0000 1100		Form feed
13	0D	0000 1101		Carriage return
14	0E	0000 1110		Shift out
15	0F	0000 1111		Shift in
16	10	0001 0000		Data link escape
17	11	0001 0001		Device control 1
18	12	0001 0010		Device control 2
19	13	0001 0011		Device control 3
20	14	0001 0100		Device control 4
21	15	0001 0101		Negative acknowledge
22	16	0001 0110		Synchronous idle
23	17	0001 0111		End of Transmission block
24	18	0001 1000		Cancel
25	19	0001 1001		End of medium
26	1A	0001 1010		Substitute
27	1B	0001 1011		Escape
28	1C	0001 1100		End of file
29	1D	0001 1101		End of record
30	1E	0001 1110		Field terminator
31	1F	0001 1111	‡	Double dagger (delimiter)
32	20	0010 0000		Space
33	21	0010 0001	!	Exclamation Point
34	22	0010 0010	"	Quotation marks
35	23	0010 0011	#	Number sign
36	24	0010 0100	$	Dollar sign
37	25	0010 0101	%	Percent sign
38	26	0010 0110	&	Ampersand
39	27	0010 0111	'	Apostrophe

APPENDIX III.B.1. CHARACTER SET - PART 1: EXPANDED 8-BIT ASCII FOR MARC RECORDS

40	28	0010 1000	(	Opening Parenthesis
41	29	0010 1001	)	Closing parenthesis
42	2A	0010 1010	*	Asterisk
43	2B	0010 1011	+	Plus
44	2C	0010 1100	,	Comma
45	2D	0010 1101	-	Hyphen (minus)
46	2E	0010 1110	.	Period (decimal point)
47	2F	0010 1111	/	Slash
48	30	0011 0000	0	
49	31	0011 0001	1	
50	32	0011 0010	2	
51	33	0011 0011	3	
52	34	0011 0100	4	
53	35	0011 0101	5	
54	36	0011 0110	6	
55	37	0011 0111	7	
56	38	0011 1000	8	
57	39	0011 1001	9	
58	3A	0011 1010	:	Colon
59	3B	0011 1011	;	Semicolon
60	3C	0011 1100	<	Less than
61	3D	0011 1101	=	Equals
62	3E	0011 1110	>	Greater than
63	3F	0011 1111	?	Question mark
64	40	0100 0000	@	Commercial at sign
65	41	0100 0001	A	
66	42	0100 0010	B	
67	43	0100 0011	C	
68	44	0100 0100	D	
69	45	0100 0101	E	
70	46	0100 0110	F	
71	47	0100 0111	G	
72	48	0100 1000	H	
73	49	0100 1001	I	
74	4A	0100 1010	J	
75	4B	0100 1011	K	
76	4C	0100 1100	L	
77	4D	0100 1101	M	
78	4E	0100 1110	N	
79	4F	0100 1111	O	
80	50	0101 0000	P	
81	51	0101 0001	Q	

APPENDIX III.B.1. CHARACTER SET - PART 1: EXPANDED 8-BIT ASCII FOR MARC RECORDS

82	52	0101 0010	R	
83	53	0101 0011	S	
84	54	0101 0100	T	
85	55	0101 0101	U	
86	56	0101 0110	V	
87	57	0101 0111	W	
88	58	0101 1000	X	
89	59	0101 1001	Y	
90	5A	0101 1010	Z	
91	5B	0101 1011	[	Opening bracket
92	5C	0101 1100	\	Reverse slash
93	5D	0101 1101	]	Closing bracket
94	5E	0101 1110		
95	5F	0101 1111		
96	60	0110 0000		
97	61	0110 0001	a	
98	62	0110 0010	b	
99	63	0110 0011	c	
100	64	0110 0100	d	
101	65	0110 0101	e	
102	66	0110 0110	f	
103	67	0110 0111	g	
104	68	0110 1000	h	
105	69	0110 1001	i	
106	6A	0110 1010	j	
107	6B	0110 1011	k	
108	6C	0110 1100	l	
109	6D	0110 1101	m	
110	6E	0110 1110	n	
111	6F	0110 1111	o	
112	70	0111 0000	p	
113	71	0111 0001	q	
114	72	0111 0010	r	
115	73	0111 0011	s	
116	74	0111 0100	t	
117	75	0111 0101	u	
118	76	0111 0110	v	
119	77	0111 0111	w	
120	78	0111 1000	x	
121	79	0111 1001	y	
122	7A	0111 1010	z	
123	7B	0111 1011		

APPENDIX III.B.1. CHARACTER SET - PART 1: EXPANDED 8-BIT ASCII FOR MARC RECORDS

124	7C	0111 1100					
125	7D	0111 1101					
126	7E	0111 1110					
127	7F	0111 1111	Delete				
128	80	1000 0000					
129	81	1000 0001					
130	82	1000 0010					
131	83	1000 0011					
132	84	1000 0100					
133	85	1000 0101					
134	86	1000 0110					
135	87	1000 0111					
136	88	1000 1000					
137	89	1000 1001					
138	8A	1000 1010					
139	8B	1000 1011					
140	8C	1000 1100					
141	8D	1000 1101					
142	8E	1000 1110					
143	8F	1000 1111					
144	90	1001 0000					
145	91	1001 0001					
146	92	1001 0010					
147	93	1001 0011					
148	94	1001 0100					
149	95	1001 0101					
150	96	1001 0110					
151	97	1001 0111					
152	98	1001 1000					
153	99	1001 1001					
154	9A	1001 1010					
155	9B	1001 1011					
156	9C	1001 1100					
157	9D	1001 1101					
158	9E	1001 1110					
159	9F	1001 1111					
160	A0	1010 0000					
161	A1	1010 0001	Ł	Polish L - uppercase			
162	A2	1010 0010	Ø	Scandinavian O - uppercase			
163	A3	1010 0011	Đ	D with crossbar - upppercase			
164	A4	1010 0100	Þ	Icelandic thorn -			

APPENDIX III.B.1. CHARACTER SET - PART 1: EXPANDED 8-BIT ASCII FOR MARC RECORDS

				uppercase
165	A5	1010 0101	Æ	Digraph AE - uppercase
166	A6	1010 0110	Œ	Digraph OE - uppercase
167	A7	1010 0111	′	Miagkiy znak
168	A8	1010 1000	·	Dot in middle of line
169	A9	1010 1001	♭	Musical flat
170	AA	1010 1010	®	Subscript patent mark
171	AB	1010 1011	±	Plus or minus
172	AC	1010 1100	Ơ	O-hook - uppercase
173	AD	1010 1101	Ư	U-hook - uppercase
174	AE	1010 1110	ʻ	Alif
175	AF	1010 1111		
176	B0	1011 0000	ʼ	Ayn
177	B1	1011 0001	ł	Polish l - lowercase
178	B2	1011 0010	ø	Scandinavian o - lowercase
179	B3	1011 0011	đ	D with crossbar - lowercase
180	B4	1011 0100	þ	Icelandic thorn - lowercase
181	B5	1011 0101	æ	Digraph ae - lowercase
182	B6	1011 0110	œ	Digraph oe - lowercase
183	B7	1011 0111	″	Tverdyi znak
184	B8	1011 1000	ı	Turkish i - lowercase
185	B9	1011 1001	£	British pound
186	BA	1011 1010	ð	Eth
187	BB	1011 1011		
188	BC	1011 1100	ơ	o-hook - lowercase
189	BD	1011 1101	ư	u-hook - lowercase
190	BE	1011 1110		
191	BF	1011 1111		
192	C0	1100 0000		
193	C1	1100 0001		
194	C2	1100 0010		
195	C3	1100 0011		
196	C4	1100 0100		
197	C5	1100 0101		
198	C6	1100 0110		
199	C7	1100 0111		
200	C8	1100 1000		
201	C9	1100 1001		
202	CA	1100 1010		
203	CB	1100 1011		
204	CC	1100 1100		

APPENDIX III.B.1. CHARACTER SET - PART 1: EXPANDED 8-BIT ASCII FOR MARC RECORDS

205	CD	1100 1101		
206	CE	1100 1110		
207	CF	1100 1111		
208	D0	1101 0000		
209	D1	1101 0001		
210	D2	1101 0010		
211	D3	1101 0011		
212	D4	1101 0100		
213	D5	1101 0101		
214	D6	1101 0110		
215	D7	1101 0111		
216	D8	1101 1000		
217	D9	1101 1001		
218	DA	1101 1010		
219	DB	1101 1011		
220	DC	1101 1100		
221	DD	1101 1101		
222	DE	1101 1110		
223	DF	1101 1111		
224	E0	1110 0000	?	Pseudo question
225	E1	1110 0001	`	Grave
226	E2	1110 0010	´	Acute
227	E3	1110 0011	^	Circumflex
228	E4	1110 0100	~	Tilde
229	E5	1110 0101	¯	Macron
230	E6	1110 0110	˘	Breve
231	E7	1110 0111	˙	Superior dot
232	E8	1110 1000	¨	Umlaut or Dieresis
233	E9	1110 1001	ˇ	Hacek
234	EA	1110 1010	°	Angstrom or circle above
235	EB	1110 1011	⁀	Ligature, first half
236	EC	1110 1100	⁀	Ligature, second half
237	ED	1110 1101	'	High comma, off center
238	EE	1110 1110	˝	Double acute
239	EF	1110 1111	˘	Candrabindu
240	F0	1111 0000	¸	Cedilla
241	F1	1111 0001	˛	Right hook
242	F2	1111 0010	.	Dot below character
243	F3	1111 0011	..	Double dot below character
244	F4	1111 0100	°	Circle below character
245	F5	1111 0101	=	Double underscore
246	F6	1111 0110	_	Underscore

247	F7	1111 0111	̧	Left hook/tail
248	F8	1111 1000	̧	Right cedilla
249	F9	1111 1001	̆	Upadhmaniya
250	FA	1111 1010	̃	Double tilde, first half
251	FB	1111 1011	̃	Double tilde, second half
252	FC	1111 1100		
253	FD	1111 1101		
254	FE	1111 1110	'	High comma, centered
255	FF	1111 1111		

GREEK, SUBSCRIPT, AND SUPERSCRIPT CHARACTER (8-BIT)

A. GREEK

The following Greek characters are preceded by the escape sequence ESCg ($1B_{16}$ 67_{16}). This is a locking escape and must be ended by the escape sequence ESCs ($1B_{16}$ 73_{16}) to return to the standard set.

Dec.	Hex	Binary	Graphic	Name
97	61	0110 0001	α	Alpha
98	62	0110 0010	β	Beta
99	63	0110 0011	γ	Gamma

B. SUBSCRIPTS

The following subscript characters are preceded by the escape sequence ESCb ($1B_{16}$ 62_{16}). This is a locking escape and must be ended by the escape sequence ESCs ($1B_{16}$ 73_{16}) to return to the standard set.

Dec.	Hex	Binary	Graphic	Name
48	30	0011 0000	₀	Subscript 0, etc.
49	31	0011 0001	₁	
50	32	0011 0010	₂	
51	33	0011 0011	₃	
52	34	0011 0100	₄	
53	35	0011 0101	₅	
54	36	0011 0110	₆	
55	37	0011 0111	₇	
56	38	0011 1000	₈	
57	39	0011 1001	₉	
40	28	0010 1000	₍	Subscript left parenthesis
41	29	0010 1001	₎	Subscript right parenthesis
43	2B	0010 1011	₊	Subscript plus
45	2D	0010 1101	₋	Subscript minus

C. SUPERSCRIPTS

APPENDIX III.B.1. CHARACTER SET - PART 1: EXPANDED 8-BIT ASCII FOR MARC RECORDS

Superscript characters have the same values as the subscript character set. However, the escape sequence to enter the superscript set is ESCp ($1B_{16}$ 70_{16}). The escape sequence to return to the standard set is ESCs ($1B_{16}$ 73_{16}).

APPENDIX III.B.2. CHARACTER SET - PART 2: 6-BIT CHARACTER SET

STANDARD 6-BIT CHARACTER SET DERIVED FROM 8-BIT ASCII CHARACTER SET

Dec.	Octal	Binary	Graphic	Name and/or Function
0	00	000 000		Space
1	01	000 001	!	Exclamation Point
2	02	000 010	"	Quotation marks
3	03	000 011	#	Number sign
4	04	000 100	$	Dollar sign
5	05	000 101	%	Percent sign
6	06	000 110	&	Ampersand
7	07	000 111	'	Apostrophe
8	08	001 000	(	Opening parenthesis
9	09	001 001	)	Closing parenthesis
10	12	001 010	*	Asterisk
11	13	001 011	+	Plus
12	14	001 100	,	Comma
13	15	001 101	-	Hyphen (minus)
14	16	001 110	.	Period (decimal point)
15	17	001 111	/	Slash
16	20	010 000	0	
17	21	010 001	1	
18	22	010 010	2	
19	23	010 011	3	
20	24	010 100	4	
21	25	010 101	5	
22	26	010 110	6	
23	27	010 111	7	
24	30	011 000	8	
25	31	011 001	9	
26	32	011 010	:	Colon
27	33	011 011	;	Semicolon
28	34	011 100	<	Less than
29	35	011 101	=	Equals
30	36	011 110	>	Greater than
31	37	011 111	?	Question mark
32	40	100 000		
33	41	100 001	a	
34	42	100 010	b	
35	43	100 011	c	
36	44	100 100	d	
37	45	100 101	e	
38	46	100 110	f	

APPENDIX III.B.2. CHARACTER SET - PART 2: 6-BIT CHARACTER SET

39	47	100 111	g	
40	50	101 000	h	
41	51	101 001	i	
42	52	101 010	j	
43	53	101 011	k	
44	54	101 100	l	
45	55	101 101	m	
46	56	101 110	n	
47	57	101 111	o	
48	60	110 000	p	
49	61	110 001	q	
50	62	110 010	r	
51	63	110 011	s	
52	64	110 100	t	
53	65	110 101	u	
54	66	110 110	v	
55	67	110 111	w	
56	70	111 000	x	
57	71	111 001	y	
58	72	111 010	z	
59	73	111 011		Shift code, nonstandard set I
60	74	111 100		
61	75	111 101		Shift code, nonstandard set II
62	76	111 110		Shift code nonstandard set III
63	77	111 111		Delete

6-BIT NONSTANDARD SET I

Each character in this set is preceded by the shift code for Nonstandard Set I (73_8).

Dec.	Octal	Binary	Graphic	Name/function
0	00	000 000		Null
1	01	000 001		Start of heading
2	02	00 010		Start of text
3	03	000 011		End of text
4	04	000 100		End of transmission
5	05	000 101		Enquiry
6	06	000 110		Acknowledge
7	07	000 111		Bell
10	10	001 000		Backspace
9	11	001 001		Horizontal tabulation
10	12	001 010		Line feed
11	13	001 011		Vertical tabulation

APPENDIX III.B.2. CHARACTER SET - PART 2: 6-BIT CHARACTER SET

12	14	001 100		Form feed	33	41	100 001	A
13	15	001 101		Carriage return	34	42	100 010	B
14	16	001 110		Shift out	35	43	100 011	C
15	17	001 111		Shift in	36	44	100 100	D
16	20	101 000		Data link escape	37	45	100 101	E
17	21	010 001		Device control 1	38	46	100 110	F
18	22	010 010		Device control 2	39	47	100 111	G
19	23	010 011		Device control 3	40	50	101 000	H
20	24	010 100		Device control 4	41	51	101 001	I
21	25	010 101		Negative acknowledge	42	52	101 010	J
22	26	010 110		Synchronous idle	43	53	101 011	K
23	27	010 111		End of trans. block	44	54	101 100	L
24	30	011 000		Cancel	45	55	101 101	M
25	31	011 001		End of medium	46	56	101 110	N
26	32	011 010		Substitute	47	57	101 111	O
27	33	011 011		Escape	48	60	110 000	P
28	34	011 100		End of file	49	61	110 001	Q
29	35	011 101		End of record	50	62	110 010	R
30	36	011 110		Field terminator	51	63	110 011	S
31	37	011 111	‡	Double dagger (delimiter)	52	64	110 100	T
32	40	100 000	@	Commercial at sign	53	65	110 101	U

III.B.2. CHARACTER SET - PART 2: 6-BIT CHARACTER SET

110 110	V		
110 111	W		
111 000	X		
111 001	Y		
111 010	Z		
111 011	[	Open bracket	
111 100	\	Reverse slash	
111 101	]	Closing bracket	
111 110			
111 111			

STANDARD SET II

acter in this set is preceded by the shift Nonstandard set II (75_8).

al	Binary	Graphic	Namd/function
000 000			
000 001	Ł		Polish L - uppercase
000 010	Ø		Scandinavian O - uppercase
000 011	Đ		Đ with crossbar - uppercase
000 100	Þ		Icelandic thorn - uppercase
000 101	Æ		Digraph AE - uppercase
000 110	Œ		Digraph OE - uppercase
7	07	000 111	ʼ	Miagkii znak
8	10	001 000	·	Dot in middle of line
9	11	001 001	♭	Musical flat
10	12	001 010	®	Subscript patent mark
11	13	001 011	±	Plus or minus
12	14	001 100	Ơ	O-hook, uppercase
13	15	001 101	Ư	U-hook, uppercase
14	16	001 110	ʻ	Alif
15	17	001 111		
16	20	010 000	ʻ	Ayn
17	21	010 001	ł	Polish l - lowercase
18	22	010 010	ø	Scandinavian o - lowercase
19	23	010 011	đ	Đ with crossbar - lowercase
20	24	010 100	þ	Icelandic thorn - lowercase
21	25	010 101	æ	Digraph ae - lowercase
22	26	010 110	œ	Digraph oe - lowercase
23	27	010 111	ʼ	Tverdyi znak
24	30	011 000	ı	Turkish i - lowercase
25	31	011 001	£	British pound
26	32	011 010	ð	Eth
27	33	011 011		

APPENDIX III.B.2. CHARACTER SET - PART 2: 6-BIT CHARACTER SET

28	34	011 100	ơ	o-hook, lowercase
29	35	011 101	ư	u-hook, lowercase
30	36	011 110		
31	27	011 111		
32	40	100 000	?	Pseudo question
33	41	100 001	◌̀	Grave
34	42	100 010	◌́	Acute
35	43	100 011	◌̂	Circunflex
36	44	100 100	◌̃	Tilde
37	45	100 101	◌̄	Macron
38	46	100 110	◌̆	Breve
39	47	100 111	◌̇	Superior dot
40	50	101 000	◌̈	Umlaut or Dieresis
41	51	101 001	◌̌	Hacek
42	52	101 010	◌̊	Angstrom or circle above
43	53	101 011	◌͡	Ligature, first half
44	54	101 100	◌͜	Ligature, second half
45	55	101 101	'	High comma, off center
46	56	101 110	◌̋	Double acute
47	57	101 111	◌̐	Candrabindu
48	60	110 000	◌̧	Cedilla
49	61	110 001	◌̡	Right hook
50	62	110 010	◌̣	Dot below character
51	63	110 011	◌̤	Double dot below character
52	64	110 100	◌̥	Circle below character
53	65	110 101	◌̳	Double underscore
54	66	110 110	◌̲	Underscore
55	67	110 111	◌̦	Left hook/tail
56	70	111 000	◌̧	Right cedilla
57	71	111 011	◌̬	Upadhmaniya
58	72	111 010	◌͠	Double tilde, first half
59	73	111 011	◌͡	Double tilde, second half
60	74	111 100		
61	75	111 101		
62	76	111 110		High comma, centered
63	77	111 111		

GREEK, SUBSCRIPT, AND SUPERSCRIPT CHARACTERS (6-BIT)

A. GREEK

The following Greek characters are preceded by the escape sequence shift ESCg (73_8 33_8 47_8). This is a locking sequence and must be ended by the escape sequence shift ESCs (73_8 33_8 63_8) to return to the standard set.

Dec.	Octal	Binary	Graphic	Name

III.B.2. CHARACTER SET - PART 2: 6-BIT CHARACTER SET

100 001	α	Alpha
100 010	β	Beta
100 011	γ	Gamma

IPTS

ing subscript characters are preceded by the quence shift ESCb (73_8 33_8 42_8) and must be he escape Shift ESCb (73_8 33_8 63_8) to return ndard set.

l	Binary	Graphic	Name
	010 000	0	Subscript 0, etc.
	010 001	1	
	010 010	2	
	010 011	3	
	010 100	4	
	010 101	5	
	010 110	6	
	010 111	7	
	010 000	8	
	010 001	9	
	001 000	(	Left parentheses
	001 001	)	Right parenthesis
	11 13	001 011	+ Subscript plus
	13 15	001 101	- Subscript minus

C. SUPERSCRIPTS

Superscript characters have the same values as the subscript set. However, the escape sequence to enter the superscript set is shift ESCp (73_8 33_8 60_8). The escape sequence to return to the standard set is shift ESCs (73_8 33_8 63_8).

APPENDIX III.C. FILL CHARACTER

The key to retaining the MARC structure, while simultaneously reducing required coding specificity is the "fill character". For U.S. MARC records the use of this fill character is limited to fixed field data elements, such as the 008 Fixed Length Data Elements Field. It may not be used in the Leader or in tags, indicators, or subfield codes. Presence of a fill character in a fixed field indicates that the full MARC format specifies a value, but the creator of the record has not attempted to supply a value. By way of comparison, presence of a code value for "unknown" in a fixed field indicates that the creator of the record has attempted to supply a value, but was unable to determine what the appropriate code should be. Use of the fill character in fixed fields is usually regulated by policy of the inputting agency.

For communication purposes, hexadecimal "7C" in the expanded 8-bit ASCII character set and hexadecimal "3C" (octal "74") in the standard 6-bit character set derived from the 8-bit ASCII character set, is to be used as the "fill character". It is typically displayed as a vertical bar.

In the fixed fields, the code "u" is used for unknown when the field contains alphabetic codes; a hyphen (-) is used for unknown when the field contains numerics.

APPENDIX II

PROPOSALS FOR QUALITY CONTROL FOR THE

NAME AUTHORITY FILE SERVICE

Contents

INTRODUCTION

These proposals for quality control in the Name Authority File (NAF) are based upon the following premises:

- authority data in NAF will be subject to as much quality control as is possible
- insofar as possible, quality control will be supported by the host computer system (there will, however, be some administrative quality control as well)
- any data used for key (access) will be edited comprehensively by the host computer system

In casting these recommendations, the Task Force on a Name Authority File Service has also assumed that authority data will be input to the NAF as logical records containing:

- established form (LC/MARC Authorities tag 1nn)
- see from (x) references (4nn)
- see also from (xx) references (5nn)
- [illegible] notes, etc.

[illegible] data will be [illegible] information text and [illegible] input such that a see [illegible] reference (A see B) [illegible] reference (B see also A) [illegible] input without reference [illegible] of the complete authority [illegible] for [illegible] heading(s) involved

SYSTEM SUPPLIED EDITING AND QUALITY CONTROL (INTERACTIVE, ONLINE)

1. New Data (i.e., any new record or field)

1.1 General Checking

1.1.1 For any field, check to determine:

a. that the tag used for the field is defined to the system

b. when indicators are defined for the field, the correct number of indicator positions is present

c. when the field is subfielded

- that only subfield codes defined for use in that field are used
- that each subfield contains data (i.e., each subfield code is followed by one character that is not blank, a delimiter for the next subfield, or end of field)
- that each subfield [illegible] begins with an invalid blank

d. [illegible] the field [illegible] length of the field [illegible]

[illegible] are [illegible] each [illegible] input [illegible] appropriate error message(s).[1]

NB: Checking for the [illegible] of valid content designation performed during interactive online input and update to the NAF should be supported such that editing rules may be changed easily to accommodate changes in content designation -- e.g., definition of new fields, definition of new subfield codes, cancellation of subfield codes. The introduction of such changes to interactive, online edit routines may also require batch (scheduled) scanning of all existing records to change content designation to agree with new definitions.

1.1.2 In addition, for <u>fields containing headings or references, or that are used for access</u>, check to determine:

a. when indicators are defined for the field, that indicator values are valid

b. when the field is subfielded, that past practices no longer condoned for subfield content are identified and stopped, e.g.,

[illegible] heading with "U.S." as sole [illegible] of subfield "a" is invalid

[illegible]

Brit [illegible] co[illegible]

subfi[illegible] is [illegible]

[illegible] it is [illegible]

th[illegible] operat[illegible] al[illegible] on [illegible] for which attempted.

e. when the field is fixed length, that the code values used for defined character positions are valid; when possible, cross-check between and among character positions

If any errors are found in the input stream, reject the entire input transaction and write the appropriate error message(s).[2]

1.2 Editing Specific to the Addition of Data to the File

Having passed the checks listed in 1.1, the field is a candidate for addition to the file.

Please note: all checks listed in section 1.2 should be performed before the input stream is added to the file; if any test fails, reject the entire input transaction and write the appropriate error message(s).[3]

1.2.1 For any field, perform whatever additional checks the system may require or that are desired.

1.2.2 In addition, for fields containing headings or references, or that are used for access, perform the additional checks specifjed below.

A. Established heading field (1nn)

[2]See footnote 1, page 6.

[3]See footnote 1, page 6.

1. Check to see if the heading is already on the file.[4]

   a. If not on the file, mark that data may be added

   b. If already on the file, reject the entire input transaction and write appropriate error message(s),

   OR

   mark that data may be added only if override is specified (use of override will depend upon matching routines, etc.).[5]

   The check to see if the heading is already on the file may be accomplished in one or more ways:

---

[4] The first test -- to see if a heading is already on the file -- could be performed as a "key" check (the content designated version of the field is translated into search/sort key; the search/sort key value is checked against the file). If a match is found on the on the file, there is a check for a content designated match between the input and the heading on the file. If they match exactly, the heading is defined as "already on file." If the only difference between the input and the heading on the file is in content designation, this would be reported as an error. Otherwise, the heading may be treated as a "new heading" for addition ot the file (this will result in some duplicates because of the differences in diacritics and punctuation).

[5] The host computer system should log the use of the override feature for review by the Library of Congress and by the host system staff. Review of override use by the Library of Congress will help to ensure that override is not being misused as regards data content, e.g., the override feature being used to perform "blind updates." Review of override use by the host computer system will help to ensure that override is not being used in order to circumvent system bugs, etc.

- o access string match (e.g., "normalized" version of heading with content designation removed)
- o character for character match of content designated data

The types of checks performed and the results defined as passing the test will determine if an override feature is necessary.

B. See from (x) references (4nn)

1. Check to see if the heading is already on the file

   a. If not on the file, mark that data may be added

   b. If already on the file, check the use of the heading on the file:

      - o if 1nn, reject the entire input transaction and write the appropriate error message(s)

        This is the case where the "see reference" proposed is already on the file as an established heading.

      - o if 4nn, mark that data may be added

C. See also from (xx) references (5nn)

1. Check to see if the heading is already on the file

a. If not on the file, reject the entire input transaction and write the appropriate error message(s)

This is a case where a "see also from" is proposed for a heading form not on the file.

Alternatively, this type of "see also from" could be added to the file, but it will result in the addition of an incomplete authority record for the "see also from" heading.[6]

b. If already on the file, check the use of the heading on file:

- o if 1nn, mark that data may be added.
- o if 4nn, reject the entire input transaction and write

[6] Library of Congress Authority Data, to be used for the Base File for the Name Authority File Service, will include many such "See Also From" headings (some of these will be topical subjects used as "See Also From" references for name headings). Thus, the initial batch loading of the system (with the Base File) may need some different machine editing routines than those used for ongoing, interactive additions and changes.

Some technique of record flagging or diagnostics should be provided by the host system so that incomplete authority records for such "See Also From" headings may be easily and quickly identified.

the appropriate error message(s)

This is a case where a "see also from" is proposed for a heading form already on the file as a "see from."

D. Other fields, not headings or references

Depending on what the fields or character positions are that are used for access, perform appropriate tests.

1.2.3 Other

A. Required input content

If the data added are a new record, check to ensure that the fields required in each record are present -- e.g., 1nn, 008, control number, etc. -- and check to ensure that mandatory field content is present.

B. Internal consistency of input content

If the data added are a new record, check that the data are internally consistent -- e.g., 1nn and 4nn content are not the same, 1nn and 5nn content are not the same, 4nn and 5nn content are not the same, etc.

If the data added are a new field to an existing record, check to see that the addition of these data to the existing record is internally consistent -- e.g., if data

added are 4nn, check to see that it is not the same as 5nn in the record.

Depending upon the structure of the file and the input requirements, the following may also be required/desirable:

C. Warning or error message(s) if duplicate search/sort keys are place on the file

D. Checks, etc., to ensure that if the corporate subdivision is placed on the file that the main body is already established, e.g.,

110 ǂaMaryland. ǂbDept. of Mines, Geology and Water Resources.

cannot go on the file until

110 ǂaMaryland

is present in the file. (This test seems more crucial than some others.)

2 Updates that Change Existing Fields

Input that updates a record by adding a new field (i.e., a field previously not in the record) is treated under Section 1 (New Data), above.

Input that updates an existing field is treated in this section (Section 2). An update to an existing field may be viewed as being:

a. Deletion only

The sole action is to remove the field (i.e., nothing is put in its place).

b. Replacement

The action is to replace existing data with something else and two steps are required:

(1) Delete existing data

(2) Add new data

(It is assumed that the system will differentiate between update actions to fields that are deletions only, those that are replacements.

2.1 Deletions Only

2.1.1 General Checking

A. Check to see that the tag used to request the deletion is defined to the system.

1. If not defined, reject the entire update transaction and write appropriate error message(s).

2. If not defined, flag the record for inclusion in the "electronic mail" for review by the host system computing staff.

2.1.3 Editing Specific to Deletion of Data From File

A. Is the field defined to the system as deletable?

Some fields will be required in a record, and may not be deleted (e.g., inn, control number, etc.).

1. If the field is not deletable, reject the entire update transaction and write the appropriate error message(s).

2. If the field is deletable, check the data to be deleted to determine if it matches exactly what is already on the file.

a. If it matches what is already on the file, mark for deletion.

b. If it does not match what is already on the file, reject the entire update transaction and write the appropriate error message(s).

2.2 Replacements

2.2.1 General Checking

A. Check to see that the tag used to request replacement is defined to the system.

1. If not defined, reject the entire update transaction and write the appropriate error message(s).

2. If not defined, flag the record for inclusion in the "electronic mail" for review by host system computing staff.

2.2.2 Editing Specific to Replacement of Data on File

A. Is the field defined to the system as replaceable?

Some considerations: Some fields will not be replaceable (e.g., control number). Is it possible that 1nn should not be replaced, but that the existing record should be deleted and a new record added?

1. If the field is not replaceable, reject the entire update transaction and write the appropriate error message(s).

2. If the field is replaceable, perform the checks listed in section 1 upon the proposed replacement data.

   a. If all tests are passed, mark the field for replacement.

   b. If any test fails, reject the entire update transaction and write the appropriate error message(s).

3 Deletion of Records

The circumstances under which it will be permitted to delete records from the Name Authority File need to be specified.

Depending upon what those circumstances are, editing rules will be developed. These rules should probably be quite strict and tied to file security in such a way to prevent erroneous deletions.

(Deletions will proceed through the "proposed change cycle.")

4- System Characteristics

The quality control features supported online for the Name Authority File by the host computer system will include conversational subsystems for addition, change, and deletion of authority records and data in those records.

For example, when a contributing source wants to change an existing record, the following steps seem appropriate:

- The operator searches and identifies the record to be changed
- The system displays data in the appropriate update format
- The operator enters the proposed changes.
- The system responds to the proposed changes (i.e., the system performs the tests outlined above and responds).
- If the system responds that the data pass all tests, the operator is asked to respond as appropriate:

- to specify that the data go to the electronic mail for the "proposed change cycle"

- to specify that the data be added to the Name Authority File

- to specify that the system send appropriate data to the "proposed change cycle" and add to the Name Authority File those data that do not need to go through the "proposed change cycle"

This will permit the contributing source to review changes before passing them to the electronic mail or adding them to the Name Authority File. At this point the operator may respond "NO" which means that proposed changes will not be added to the electronic mail or the NAF (this would be the response, for example, if the operator saw an error in the proposed change, such as a misspelled word, that should be corrected before proceeding further).

SYSTEM SUPPLIED EDITING AND QUALITY CONTROL (SCHEDULED REPORTS)

Even though it is assumed that all input and update to the Name Authority File will be accomplished through interactive, online use of the system, there will be definite advantages to producing batch reports on system activity, system use, etc. These reports may be used to monitor information added to the NAF, to identify the nature of problems that users are having, and probably to isolate areas where conflict resolution may be required.

The first four reports shown below are considered the most important and should be supplied on a regular and timely basis (perhaps daily). Other reports shown below may be useful only under certain circumstances (e.g., by new contributing sources, by the Library of Congress on a "spot check" basis, or under certain operating conditions of the host computer system (e.g., diagnostics from scheduled batch editing would be used regularly only if all editing is not provided during online, interactive system use).

1. List of headings added/changed using override

   If override features are permitted, it will be useful to show headings "forced" into the file.

2. List of duplicate search/sort keys

   If duplicates are permitted, it will be extremely useful to know the circumstances under which they are occurring (cf. section 1.2.2.A).

3. List of headings deleted

4. List of records deleted

5. Statistics of system use, system performance, Name Authority File growth, etc., e.g.;

   - number of records added, changed

   - number of headings added, changed

   - number of transactions ge[illegible]y type

   - average response time, response time per different types of transactions

   - use of "proposed change cycles," responses to proposed changes, etc.

6. List of headings added

Experience with online, interactive input/update editing of data shows that even the most well-considered series of tests will have some loopholes, or it will be decided not to install some types of tests online until a sense of the magnitude of the problems encountered is identified. There are some types of errors that it is probably senseless to search for through online, interactive editing, and these should be searched for in scheduled batch proce[illegible] (daily, weekly, etc.). (cf. report 7, below)

A method of quality control that seems potentially quite useful is the listing of headings added during a day, week, etc. Such listings could be used by the conflict resolution agency to monitor headings added to the file. (A knowledgeable cataloger reviewing such a list can often identify problems early; the earlier

problems are corrected, the better. Some problems will be coding problems.)

Lists broken out by heading type (100, 110, 111, 130, etc.), and within that by form of name indicator, would be useful non-system quality control for human review. (Initially, such lists might be produced daily, then on some sample basis; it may be desirable to list all headings added by new contributing sources until it is clear that the contributing source is using the system properly.)

Diagnostics on scheduled batch processing

Certain types of editing may be appropriate for scheduled batch processing. Reports of results of such processing must be prepared. (Such reports would be especially important if not all additions to the Name Authority File are accomplished through online, interactive use of the system.)

For example, batch editing might identify headings that incorporate subfield patterns rarely used (or that seem peculiar; e.g., a corporate name with 3 or more subfield "b" in data).

8. Other reports

For example, by institution, lists of error messages received could be generated. Such lists would indicate the understanding of the system and its use; this information could be used as a basis for retraining or, if applicable, removal from system use.

SYSTEM SUPPLIED EDITING AND QUALITY CONTROL (SAMPLING)

In order to review and monitor system use and Name Authority File growth, it seems desirable to have the host computer system support mechanisms for deriving samples of data in the NAF and making such samples available for analysis (either by the system or by personnel responsible for conflict resolution and/or quality control).

As a means of system-supported quality control it would be desirable for the system, on a predetermined schedule (weekly, monthly, quarterly, etc.), to draw a representative sample of Name Authority File data. Such data would include (a) records created by contributing sources during the specified report period, and (b) records changed by contributing sources during the specified report period.

These samples (random or stratified) could be drawn using audit trail data or other data in the Name Authority File. The sample would be available as a special machine-readable file, for review by appropriate personnel and for subsequent analysis by statistical program packages (which would include the production of statistical reports).

The sampling technique should also be available upon request -- e.g., to satisfy a request to sample the work of only one contributing source for some period of time. In this instance, the sample should be drawn and produced through batch processing, and made available within 24 hours.