CLASSIFICATION AND ATTRIBUTES OF NON-TEXT-SEARCHABLE DOCUMENTARY MATERIAL AND ITS TREATMENT IN THE LICENSING SUPPORT SYSTEM (LSS)

Prepared for

Nuclear Regulatory Commission
Contract NRC-02-88-005

Prepared by

Center for Nuclear Waste Regulatory Analyses
San Antonio, Texas

September 1992

DRAFT

CLASSIFICATION AND ATTRIBUTES OF NON-TEXT-SEARCHABLE DOCUMENTARY MATERIAL AND ITS TREATMENT IN THE LICENSING SUPPORT SYSTEM (LSS)

Prepared for

Nuclear Regulatory Commission
Contract NRC-02-88-005

Prepared by

Charles Acree, Rawley Johnson, Bob Brient, Steve Specto.

Center for Nuclear Waste Regulatory Analyses
San Antonio, Texas

September 1992

DRAFT

CONTENTS

CONTENTS (CONT)

FIGURES

TABLES

ACKNOWLEDGMENTS

This report was prepared to document work performed by the Center for Nuclear Waste Regulatory Analyses (CNWRA) for the U.S. Nuclear Regulatory Commission (NRC) under Contract No. NRC-02-88-005. The activities reported here were performed on behalf of the NRC Office of the Licensing Support System Administrator. The report is an independent product of the CNWRA and does not necessarily reflect the views or regulatory position of the NRC.

The authors gratefully acknowledge the technical support provided by members of the CNWRA and Southwest Research Institute (SwRI) staffs and appreciate the graphic and clerical help provided by the SwRI Publications & Photographic Services group.

1 BACKGROUND

The Licensing Support System Administrator (LSSA) has requested the Center for Nuclear Waste Regulatory Analyses (CNWRA) to classify the various types of documentary material which are unsuitable for entry into the Licensing Support System (LSS) in the form of searchable full text, to describe the attributes of each defined category, to identify the types of material that may be resident on non-imageable media, and to specify the characteristics of non-imageable media that should be recorded to assist information retrieval.

The LSSA has also asked the CNWRA to propose a definition for a "technical investigation package" (TIP) which would contain non-text-searchable materials, to discuss how TIP segments could be submitted to the LSS if an investigation is lengthy, to propose a format for a TIP table of contents, and to suggest what attributes should be recorded in a TIP's bibliographic header.

The CNWRA submitted a draft report to the LSSA in March 1992 which addressed some of these topics. This report represents a revised and expanded draft, including the modifications and additional sections requested by the LSSA.

This report was prepared by the authors in collaboration with several technical experts, representing different technical disciplines within the CNWRA, and is based upon their combined experience.

2 CLASSIFICATION OF NON-TEXT-SEARCHABLE DOCUMENTARY MATERIAL

To fulfill its classification assignment, the CNWRA has considered all of the items that its technical staff has actually observed—in its everyday work as well as investigations in pursuit of this task. It has also considered materials mentioned in the LSS Rule (NRC, 1989), materials listed as Department of Energy (DOE) and Nuclear Regulatory Commission (NRC) document-type codes, materials listed in quality assurance program documents, and materials that it believes could emerge in future years. Essentially, the CNWRA has expanded upon the model it produced in its report entitled "Alternative Ways of Making Packaged Documentary Material Accessible Within the Licensing Support System" of September 1991 (CNWRA, 1991b).

Using physical form or "document type" as the basis of its classification scheme, the CNWRA has defined the following categories of non-text-searchable documentary material:

- Map
- Photograph
- Design Drawing
- Figure
- Data Sheet
- Data Plot
- Computer Program
- Handwritten Notes.

Document type was selected as the basis of the classification scheme because this distinction will determine appropriate processing and storage requirements for these kinds of material and will provide a requestor with an additional way to find them in the LSS aside from subject content.

Ideally, classification should be both exhaustive and mutually exclusive. The CNWRA cannot verify that its classification is totally exhaustive because this voluminous material is extremely varied and has not been completely inventoried. As the LSS in loaded, it is possible that new materials may emerge that will suggest additional categories. Overlaps complicate the pursuit of mutual exclusivity. It is useful, for example, to make a distinction between data sheets, data plots, and handwritten notes, as the CNWRA has done. However, many of the data plots that have been observed are surrounded by handwritten notes and/or numeric data. Data sheets have sometimes appeared as computer printouts and sometimes as handwritten entries on printed forms. Individual documents often contain interleaved pages of data sheets, figures, notes, and other graphic material.

For more thorough classification, an effort was made to define sub-categories that would be as specific as possible, while maintaining reasonable exclusivity and exhaustivity. This effort resulted in the identification of a number of potential sub-categories, but they were inevitably based upon the content (technical subject matter) of the material, rather than upon its form. The subject content of an LSS unit (document) will be adequately described in the Title, Abstract, Descriptors, and Identifiers fields of the LSS bibliographic header. Thorough subject-based classification will be employed by the Descriptors field, which will be governed by an extensive Thesaurus of controlled, subject-related terms. Redundancy between the LSS Document Type field and the LSS Descriptors field would, in the CNWRA's opinion,

be undesirable. The CNWRA believes that technical subject will usually be the primary bibliographic search criterion within the LSS and that Document Type will usually be a secondary criterion which might be used to refine a search to more specific information, by reducing the number of "hits".

Southwest Research Institute library science specialists were enlisted to review and validate the CNWRA classification. They found it "user friendly" and commented that a sub-classification that would provide a second or even a third "tier" would be desirable if it could be supported. The CNWRA was not satisfied that this goal was feasible, given its observation that subject-based distinctions inevitably intrude.

The CNWRA's efforts to identify sub-categories are discussed more specifically in Appendix A. It is recommended that the eight primary categories be represented in the LSS Document Type List, which the CNWRA understands will be developed by the LSSA.

Technical investigation packages (discussed in subsequent sections of this report) will contain large collections of non-text-searchable documentary material which, in the context of these packages, is commonly known as "raw data." Most of the non-text-searchable material that will be entered into the LSS is expected to be contained within these packages. For this reason, the term "package" should also be an option on the LSS Document Type List, as any given package is likely to contain various types of documents.

Appendix B provides descriptive sheets that define the attributes of the eight primary categories of non-text-searchable documentary material, with "examples" based on the CNWRA's efforts to sub-divide the primary categories. The individual sheets also define the respective criteria that will normally be used to retrieve these materials and recommend unitization guidelines, related-document guidelines, and submission requirements. They employ the format that was suggested by the LSSA (LSSA, 1991a), with some rearrangement. It will be noted that entries under the heading "Criteria Normally Used for Retrieval Purposes" are identical for all eight categories, as one would expect, given the prime importance of technical subject matter, date, and origin—validated to some extent by the LSS Prototype Test Report (SAIC, 1990a). The precepts under the heading "Unitization Guidelines" are consistent with the CNWRA's earlier recommendations on this matter (CNWRA, 1991b).

3 "NON-IMAGEABLE" MATERIAL

The LSS Rule (NRC, 1989) distinguishes between two kinds of non-text-searchable documentary material. First, there is "graphic oriented" material, which is not appropriate for entry into the LSS in the form of searchable full text but can be converted by a digital scanning device into electronic "bit-mapped" images that may be subsequently viewed on LSS computer screens (and reproduced by graphic printers). This material may be handwritten, as well as graphic, in nature. Second, there is material which is not only unsuitable for entry into the LSS in the form of searchable full text, but is also unsuitable for entry into the LSS in the form of viewable images. Bibliographic headers will be created for such items, so that a requestor may find them through a bibliographic search; but, since there will be no way to examine them "on-line," they will have to be physically retrieved. A requestor will either have to order copies of the items or make other arrangements to examine them.

3.1 FACTORS AFFECTING SCANNABILITY

The CNWRA has given long consideration to this important distinction and concluded that a given item's suitability for scanning, or what might be called its "scannability" or "imageability," appears to depend upon four factors, which may overlap:

(1) Is it technically feasible to produce an image of a given item? Predictably, there will be miscellaneous items that will be inherently unsuitable for conversion to an electronic image. For example, a map with colored flags pinned on it (representing significant locales) would be hard to convert into a recognizable image. The Department of Energy calls an item such as this "one of a kind." Predictably also, there will be items that are of such poor quality that they cannot be converted by any reasonable means into clear images. For example, a dark, smudged field map with barely discernable handwritten notations upon it would be unlikely to produce a satisfactory image. Presented with an indistinct image of something important, a requestor would no doubt demand to see the original item in order to make a personal judgment regarding its information content.

(2) Would a given item's image be meaningful and useful to a requestor? For example, numeric data residing on a magnetic tape could be printed out as a data sheet and subsequently scanned to produce an electronic image (and even searchable "text"). But it would be pointless to print such data in most instances, particularly if the data are voluminous, because the information content is so thoroughly dependent upon a computer-based program for presentation and interpretation. If the data relate to a dynamic computer model, printing the data would be even less useful, because the data's meaning would depend on how various factors are given weight in manipulating the model; there would, in fact, be multiple meanings that could be derived.

(3) Can a given item be reduced to a still image without loss of information content? For example, a series of slightly different photographs resident on a motion picture reel could be individually printed and scanned to produce discrete images. But this, too, would be pointless, unless a certain frame is of special importance. The information content, in this case, is the intrinsic relationship between succeeding frames, which produces an illusion of continuous motion.

(4) Can a given item be reduced to a two-dimensional image without loss of information content? For example, if a three-dimensional computer model intended to depict fluid flow were reduced to a two-dimensional image, significant loss of information content would occur.

Other factors have been considered. Color-significant items (e.g., color-coded graphs) would not create useful images if the LSS is confined to black-and-white screens. Extra-large items (e.g., fold-out maps) would not be imageable if LSS equipment is incapable of scanning them—although they could be "tiled" (cut apart) for scanning if the separataion is limited to a maximum number of tiles, which is best set at four, allowing a maximum size of 17×22 inches for the item itself. In the CNWRA's opinion, these other factors, which would prohibit the scanning of a sizable amount of material, should be ruled out. The CNWRA has recommended (CNWRA, 1991b) that the LSS use color monitors (which are becoming standard for systems with graphic capabilities) and that items be scanned regardless of size (as they are for significant systems with graphic capabilities).

The CNWRA, furthermore, believes that all LSS workstations should be capable of displaying images. Prevailing LSS design is based not only on the presumption that "oversize" items and color-significant items are not scannable; it is based on the presumption that some LSS users will have workstations that lack the capability to display images, although the system designer assumes that all workstations will have that capability by the time of the repository license application is submitted in 2001 (SAIC, 1990c). The CNWRA believes that this will occur by the time the LSS becomes operational and that LSS planning should proceed accordingly. Graphic capability will be essential to LSS success, while its added cost, if any, will likely be insignificant. The inability of an LSS user to view images on a workstation would require the user either to borrow someone else's workstation for serious inquiry or to order desired images (based on searches of bibliographic headers and/or package tables of contents) from the LSSA. The LSSA would have to print and mail those items, probably in great quantity.

The four factors governing scannability have a common theme: the probable loss of information content (and possibly context) if an inappropriate attempt is made to convert an item into an electronic image. In a system that will be used for a potentially contentious licensing proceeding, even a minor loss of information content would be undesirable. If a perception should arise among LSS users that potential losses of information content are being tolerated in order to promote the conversion of material to electronic images, the LSS would suffer a serious loss of trust.

Even a straightforward transformation of an item that is undertaken to allow it to be captured as an LSS image poses the risk of information loss or adulteration. For example, the flagged map (in the above example) could be intentionally altered so that large colored dots would replace the flags, thereby permitting a clear image to be made; but what if one of the little flags is lost or misrepresented in the process? If the illegible map (above) were re-drawn, some of its content might very well be inadvertently changed. Changes, when detected, may give rise to suspicions that they were in fact intentional.

Thus, there is good reason for the CNWRA to recommend that non-text-searchable documentary material, which will exist mostly as raw data in support of technical investigations, should be kept in its original form, as it was "acquired" by the principal investigator. Records managers and others who process LSS material should be instructed to avoid transforming the original versions for the purpose of improving their scannability.

There is another common theme among the factors governing scannability: machine dependence. Most of the items that will be unsuitable for conversion to viewable images will be resident on media that must be examined with the help of electronic or optical/mechanical equipment. Those that are not resident on such media will occasionally appear as miscellaneous items that are disqualified on the basis of the first factor alone (being unsuitable for reasons of poor quality or extraordinary character). The rest of the items, however, which are expected to constitute the vast majority, can be usefully characterized as items residing on "machine-dependent" media.

Increasingly, documentary material is appearing on electronic media, rather than on the conventional medium of paper. The CNWRA expects that, by the time the LSS becomes operational, much of its information will be permanently and exclusively resident on electronic media, such as magnetic tape. However, other machine dependent media, relating to photography, will also retain information.

3.2 TYPES OF MACHINE-DEPENDENT MEDIA

The CNWRA has identified twelve machine-dependent media types that either store documentary material today or may be used to store it in the future:

- Magnetic Tape
- Magnetic Disk
- Solid-State Device
- Optical Disk
- Videocassette/Videodisc
- Audio Recording
- Motion Picture Film
- Microform
- Photographic Slide
- Photographic Negative
- Develocorder Film
- Radiograph.

Just as all possible types of non-text-searchable documentary material cannot be exhaustively predicted, neither can all possible types of media. The CNWRA has identified all of the types that appear to be reasonable, conceptually, at this time. Solid-state device was included, for example, although there is probably no documentary material that actually exists on such a device today.

It is recommended that these twelve media types become entries in a controlled Storage Media List for the benefit of the proposed Media Type field in the LSS bibliographic header, for which recommendations are made in Appendix E. The media types are individually defined in Appendix C.

Table 3.1 illustrates the probable relationships between the eight categories of non-text-searchable material described in Section 2 of this report and the twelve machine-dependent media types listed above. The categories which are likely to exist on a specific medium have a 2 at the matrix intersection. Those which may exist on a specific medium have a 1. Those which are unlikely to exist on a specific medium have a 0.

Table 3.1. Likely Use of Media Types for Non-Text-Searchable Materials

Media Types	Non-Text-Searchable Documentary Material Categories							
	Map	Photograph	Design Dwg	Figure	Data Sheet	Data Plot	Computer Program	Handwritten Note
Magnetic Tape	2	1	2	2	2	2	2	0
Magnetic Disk	2	1	2	2	2	2	2	0
Solid-State Device	2	1	2	2	2	2	2	0
Optical Disk	2	1	2	2	2	2	2	0
Videocassette/Videodisc	1	2	1	1	1	1	0	0
Audio Recording	0	0	0	0	1	0	0	0
Motion Picture Film	1	2	1	1	1	1	0	0
Microform	2	2	2	2	2	2	2	2
Photographic Slide	1	2	1	2	0	1	0	1
Photographic Negative	1	2	1	2	0	1	0	1
Develocorder Film	0	0	0	0	0	2	0	0
Radiograph	0	2	0	0	0	0	0	0

2 Likely to exist
1 May exist
0 Unlikely to exist

3.3 RETRIEVAL OF NON-IMAGEABLE MATERIALS

An LSS requestor will usually find non-imageable materials of interest through a search of bibliographic headers, since all such materials will have headers according to the CNWRA's recommendation (CNWRA, 1991b). A requestor may also find them by conducting a text search among the tables of contents of technical investigation packages (discussed below), or by browsing through the viewable materials in these packages, which will include the original header input forms inserted as "slip sheets" to provide descriptions of items that have been stored apart from the package.

Once an LSS requestor has found non-imageable materials of interest in one of those ways, the requestor will presumably want to examine them as soon as possible. Since they cannot be viewed on the LSS screen, the requestor will need to know how to retrieve them and what possible difficulties may be presented by their machine dependencies. More specifically, the requestor will need to know what organization has stored them, how to contact that organization for retrieval assistance, and what specific hardware/software requirements are involved.

Those specifics will be included within LSS bibliographic headers (and also on the "slip sheet" header input forms that replace these materials within packages). The CNWRA has proposed two fields that will serve this purpose: the controlled Media Type field mentioned above (supplemented by the existing free-form Comments field) and a Storage Location field. See Section 7.

Appendix C, which defines the machine-dependent media that will hold non-imageable material, specifies the particular hardware/software dependencies and other critical attributes that should be described for the different media within bibliographic headers. It also specifies the particular conversion considerations that should apply to them.

After learning what is necessary to examine non-imageable materials of interest, the requestor will have to contact the organization that produced them to determine how the materials may be examined. It is expected that most of the electronic and photographic-based materials may simply be copied and sent to the requestor. Those which cannot be copied will either have to be sent on loan (if permitted by LSSA protocols) or examined on the site of the originating organization.

In its investigations on this subject, the CNWRA has heard concerns that, based on past experience, an originator's response may be inadequate. Suggestions have been made that non-imageable materials should be permanently stored at a central location (such as the University of Nevada) for more timely response, and that a central "library" for these materials should be established under the control of the LSSA. The CNWRA sees no need for central or LSSA-controlled storage, provided that the LSSA can assure LSS users that these materials will be made available by responsible organizations in a timely fashion—within a few days at most—and will adhere to acceptable standards of compliance.

The LSSA will have to provide guidance to the responsible organizations that will govern how materials must be stored for proper safekeeping. They must be stored so that they may be expeditiously retrieved, yet in a way that will prevent deterioration, loss, theft, and damage. Physical access to the vaults in which they are permanently kept should be strictly controlled through approved access authorization lists.

The LSSA will also have to provide guidance to deal with the problem that is expected to occur as various types of equipment on which these materials are dependent fade from common use and become gradually obsolescent (e.g., equipment using eight-inch diskettes or seven-track magnetic tapes).

It is generally no problem to set standards which will ensure that information is not lost as a result of the normal deterioration of media—various kinds of magnetic media in particular. Media can simply be regenerated periodically, according to an acceptable schedule, with care to avoid generation loss, especially with videotape and photographic film. The problem that might arise when a particular reader/copier on which dormant material is dependent breaks down, and cannot be repaired or replaced, is also preventable. Policies can be set within originating organizations dictating that material must be

transferred to new media in anticipation of equipment obsolescence, when this can be done without loss of information content.

3.4 THE PROBLEM OF EQUIPMENT AVAILABILITY

The real problem in this area is that requestors of non-imageable materials may not have the necessary equipment available to examine them. The CNWRA has delineated the following options.

(1) A requestor could be provided copies of the materials, with adequate specifications for use of the media on which they are resident, and be expected to secure the equipment necessary to examine them. This places all of the burden on the requestor.

(2) A requestor could be provided both the materials and the requisite equipment. This would work only with small, simple, spareable, equipment which could be provided on loan.

(3) A requestor could be invited to examine the material on site if requisite equipment is unavailable to the requestor. This could be both inconvenient and costly to some requestors.

(4) The originating organization which stores the materials could be required to convert requested material onto media which the requestor can examine with the requestor's own equipment. This would place an unacceptable, open-ended burden on originators. It could also result in a loss of information context.

(5) The originating organization could be required to convert the requested material onto media which the requestor can examine with equipment that would be available to everyone at a central location (the University of Nevada, for example). This would be less burdensome for the originator, since commonplace equipment would apply, but (like the previous option) this would involve conversion of data in response to a specific request, which might inadvertently alter the data or be suspect.

(6) The originating organization could be required to routinely, repeatedly, and carefully convert old media to new forms as outdated machines are continually replaced. This will be done by some organizations, to some extent, in the normal course of operations. A policy enforcing it, however, would be both difficult to standardize and costly.

(7) The LSSA could offer a requestor the option of using an LSSA-retained contractor to convert selected material onto media which the requestor can examine with available equipment. This might result in undetected modification of data.

The CNWRA recommends a straightforward, combined approach to the problem. If the materials cannot be copied (or loaned) to a requestor who either has or can obtain access to requisite equipment, the requestor would have to examine the material on the site of the originating organization where the material is stored. Conversions in response to a request (in contrast to copies provided on request) should be avoided.

4 DEFINITION OF A TECHNICAL INVESTIGATION PACKAGE (TIP)

The LSS Rule (NRC, 1989), 10 CFR 2.1003(c)(3), requires that upon the occurrence of certain conditions discussed below, a "bibliographic header shall be submitted for a table of contents describing that package of information, and documentary material contained within that package shall be named and identified." (emphasis added) The provision relates to packages involving "technical work or investigation." Therefore, this report will hereafter use the phrase "Technical Investigation Package" to refer to packages identified under the above citation. Technical Investigation Packages (or "TIPs") may focus upon activities such as: the production of a technical report, the completion of a map, the drafting of an engineering design, the creation of a computer model or program to organize or assess technical data, or simply the compilation of technical data on a given subject area of investigation. A TIP may also result from a technical review or analysis, whether internal or by one party of any of the above products prepared by or for another party. TIPs would not, however, focus upon administrative matters, involving procurement, financial, travel, personnel, or training matters.

All participants and potential parties would be required to organize material in a TIP when two or more individual items of documentary material have been or are being "collected or used in conjunction with other such information to analyze, critique, support or justify any particular technical or scientific conclusion, or relates to other documentary material as part of the same scope of technical work or investigation" (10 CFR 2.1003(c)(3)). These items may consist of sets or groups of various types of documentary material, whether textual, graphic-oriented, or unsuitable for entry as images into the LSS, which have been created, obtained, used, captured and/or stored by a party or potential party to the HLW proceeding. The complete TIP, its Table of Contents, and accompanying header are required to be submitted to the LSS upon the issuance of a final product, or upon the conclusion of the specific investigation without a final product. Documentary material prepared, obtained or generated after the completion and submission of a TIP would not be added to an existing TIP. However, if such material is generated, it would be submitted either individually or included within a subsequent TIP and would reference and/or incorporate a previous TIP in whole or in part.

The August 8, 1988, Final Draft of the Rule and Minutes of the July Meeting from the facilitator of the Negotiating Team contains discussion on the issue of including raw data in packages. It is clear from this discussion that the rule was designed to address the types of "packages" which the DOE was already assembling at various sites including Yucca Mountain. The discussion by the DOE representative on this issue contains the following:

> ". . . he stated that DOE had some concerns with the requirement for 'reasonably contemporaneous' entry of this type of documentary material. He acknowledged that the reasonably contemporaneous requirement was consistent with the requirements for entering other types of documentary material into the LSS, but he explained that this would not make sense for raw data because it is often collected at separate points in time and is not used until a complete set or 'suite' of data has been collected. Furthermore, almost without exception, the data must be subject to quality assurance procedures before it is used by DOE in a study or assessment. Thus, he proposed that the words 'reasonably contemporaneous with their creation or acquisition,' as found

> in Sections 2.1003(c)(1) and (c)(2), be changed to read 'in a time frame to be established by the access protocols under Section 2.1011(d)(10).'
>
> "In response to this suggestion, representatives of the State of Nevada wanted it to be clear that as long as this meant that the types of documentary material to be covered by this section are entered into the LSS after the principal investigator decides that the data is in a form that it can be used, including the completion of quality assurance procedures, this change would be acceptable.
>
> "The Committee agreed to make the language change suggested by the DOE for Sections 2.1003(c)(1) and (c)(2). In addition, the Committee agreed that the Supplementary Information (SI) to the rule should specify that the access protocols should make every attempt to ensure that any collection or 'package' of documentary material, as the term is used in Section 2.1003 (c)(3), which relates to a study, should be submitted reasonably contemporaneous with the completion of such a 'package', including any quality assurance that might be necessary."

The above agreement was reflected in both the Proposed and Final Rulemakings [53 Fed. Reg. 44411 (1988); 54 Fed. Reg. 14925 (1989)]. The proposed rule reflects the language of the August 8th draft and states at page 44415, in the preamble:

> "Graphic-oriented material includes raw data, computer runs, computer programs and codes, field notes, laboratory notes, maps, and photographs which have been printed, scripted, handwritten or otherwise displayed in any hard copy form and which, while capable of being captured in electronic image by a digital scanning device, may be captured and submitted to the LSS Administrator in any form of image, along with a bibliographic header. Proposed § 2.1003(c) also addresses documentary material that is not suitable for entry into the Licensing Support System in either image or searchable full text. . . . The access protocol should ensure that any collection or 'package' of documentary material, as the term is used in proposed § 2.1003(c)(3), which relates to a study, should be submitted reasonably contemporaneous with the completion of such a 'package,' including any quality assurance that may be required."

The final rule repeated the preamble language of the proposed rule [54 Fed. Reg. 14934-35 (1989)].

The intent throughout the negotiations and rulemaking was both simplicity, in order to provide for uniform compliance, and completeness, in order to provide all parties with access to all discoverable material and utilize existing cataloging processes then in use by DOE. This intent is captured by the use of packages to both organize background and supporting data in a more orderly fashion and enable researchers greater assurance that they have identified the total documentary record on a given discreet area. There is no explicit or implied intent to "require" the assembly of "packages" that would not naturally result from the scope or content of the technical material. However, when a package has been assembled as a result of a specific scope of work, a package table of contents and header for that table of contents are required to be created and submitted to the LSS. Long term studies have always been intended to be assembled in identifiable units; however, individual inquiries which may relate to the same general subject, may or may not be part of consolidated packages depending on factors such as: authors, intended uses, and timing of preparation. The same raw data may be used by different individuals in

behalf of the same party and not be part of a single package if factors such as ultimate use are not common. Several examples of DOE packages were previously provided by the CNWRA (CNWRA, 1991b).

In summary, a TIP is a collection of documentary materials which are part of the same scope of technical work or investigation, or which are collected or used with other information in relation to a specific technical or scientific conclusion.

5 TIP SUBMISSION

The purpose of a TIP, as envisioned by the LSS Rule (NRC, 1989), is to gather, for convenient storage and expeditious retrieval, all of the materials pertaining to a clearly identified investigative activity, project or study. The aim is to keep closely-related materials together as an integral, mutually-dependent collection, on a consistent basis, so that the individual items will not be inconveniently scattered or improperly disassociated from one another. Packaging is intended not only to simplify the storage and retrieval of raw data resulting from an investigation, but also to keep that raw data closely associated with any finished products, such as technical reports, that are based upon it.

Section 2 of this report identified eight categories of non-text-searchable documentary material. A TIP may include material from any of the defined categories. It may reside on imageable media, such as paper, or on non-imageable media, which is mostly machine-dependent as described in Section 3. While most of the non-text-searchable material that will be entered into the LSS is expected to be contained in TIPs, this material will not appear exclusively in TIPs. A map, a photograph, or any of the other materials might very well be submitted to the LSS by its originator as a stand-alone item, having no association whatever with a TIP.

The inclusion of a given item within a TIP is determined not by its type or appearance, but by the answer to a simple question: Was it produced as the result of technical work that is documented by a package? To use the more precise words of the LSS Rule, an item would be included in a package if it "has been collected or used in conjunction with other such information to analyze, critique, support, or justify any particular technical or scientific conclusion, or relates to other documentary material as part of the same scope of technical work or investigation" [10 CFR 2.1003(c)(3)].

After stating what belongs in a package (which this report is calling a TIP), the LSS Rule further stipulates that, when material has been thus collected for submission, "an appropriate bibliographic header shall be submitted for a table of contents describing that package and documentary material contained within that package shall be named and identified" [10 CFR 2.1003(c)(3)]. Through those final actions, a package is completed.

A TIP may not be treated as an open-ended collection. It remains "open" only while the investigation that it documents is still incomplete. After all of the work has been done, after all of the pertinent data have been accumulated, after all of that data has been subjected to applicable quality assurance standards, and after a finished report or other product, if any, has been written, internally reviewed, and published, the TIP will then be closed. A table of contents will be created for the package and the package will be indexed. No further materials will be subsequently added to it. Later reviews (whether internal or external) and subsequent data collections or reports concerning the subject investigation and its resulting materials will be related to the closed package through customary means of reference provided by LSS bibliographic headers, using pointers. Corrections to packaged material, of course, may be made at any time, in accordance with record-correction protocols that will apply on a general basis to all LSS units. The bibliographic headers for errata sheets relating to package-units will be linked to those units through the "Package ID" field. See Section 7.

All of the above requirements and observations regarding a TIP appear to be straightforward. However, there is an inherent problem in the practice of packaging documentary material which the parties drafting the LSS Rule most certainly recognized but chose not to address specifically. They no doubt expected the LSSA to address this problem, as well as other troublesome implementation details, when it establishes access protocols for non-text-searchable data.

The inherent problem is this: A given investigation may take several months, possibly a year or more, to complete. If the responsible organization is allowed to withhold from the LSS all of the information that is gradually generated by an investigation until it is finally concluded so that a comprehensive package of the relevant materials can be prepared for LSS entry with a complete table of contents, the resulting delay may be unacceptable to other parties interested in obtaining all or parts of that information. Though the bulk of packaged materials will consist of raw (primarily non-text-searchable) data, its timely entry into the LSS will surely be essential to the system's success.

What, then, is the best way to make accumulating packaged material available to LSS requestors during the course of a lengthy investigation, prior to its completion? Should an otherwise delayed TIP be divided, in some manner, into parts for more timely, periodic submission? If a package is divided, how would it eventually become a consolidated entity? Would it have a comprehensive table of contents? How would the required bibliographic header leading to that table of contents be affected by the division?

The LSS Rule offers limited guidance on this matter. It makes no suggestion that a package be divided. The Rule itself says only that the LSSA will set "time frames" when it establishes access protocols for the submission of all kinds of non-text-searchable material. Supplementary information to the Rule, referring to Section 2.1003, says only that a package "should be submitted reasonably contemporaneous with the completion of such a 'package', including any quality assurance that may be required."

As previously reported (CNWRA, 1991a & b), the Department of Energy's Yucca Mountain Project Office (DOE/YMPO) has struggled with this problem, to satisfy its own internal needs as well as the anticipated mandates of the LSSA access protocols. The DOE/YMPO is the only LSS participant that has adopted and implemented formal packaging procedures. It is also the acknowledged producer of the vast bulk of documentary material, including non-text-searchable material. While it is likely that other organizations will conduct studies and investigations that will cause them to submit packaged materials to the LSS, it is the DOE/YMPO that will be most affected by any procedural changes that must be imposed to satisfy LSS packaging requirements.

The DOE/YMPO has been employing what it calls "data-record packages" to combine materials relating to a particular investigative activity, in compliance with the LSS Rule. It has, in fact, been storing nearly all of its non-text-searchable documentary material within packages that correspond basically to the CNWRA's definition of a TIP.

To speed the introduction of packaged material into its central records system, the DOE/YMPO at first required that material be submitted quarterly, in what it has called package "segments". It soon discovered that a rigid time-driven requirement was difficult to impose, because it resulted in numerous requests for specially-approved (delayed) package submission schedules, which, in practice, were seldom denied.

By this time, the DOE/YMPO may have devised a more workable solution. (The CNWRA has been unable to discuss the matter with the DOE/YMPO since August 1991 and therefore may not be up to date regarding its current operating procedures.)

It must be pointed out that prolonged submission of packaged DOE/YMPO material from its many, geographically dispersed contractors will not be as much of a problem for the DOE/YMPO as it will be for other LSS participants wishing to inspect DOE/YMPO material. When packaged material is retained at the site of an investigation, the DOE/YMPO and its contractors will know precisely what items are available and how they can be retrieved expeditiously, using internal procedures. This knowledge will come from the enlarged Data Catalog that the DOE/YMPO will publish quarterly to describe the status and location of all of the materials that its contractors have produced during the course of their investigations. This knowledge will also come from the Office's Automated Technical Data Tracking System (ATDT), which will not only update the Data Catalog, but will be invaluable to DOE/YMPO contractors in between updates, providing them knowledge of available materials on a "real time" basis.

Whatever solution the DOE/YMPO may adopt to satisfy its own needs with regard to the prolonged-investigation problem, the solution that is finally adopted for the LSS must satisfy LSS needs and be applicable to all TIPs, on a consistent basis, regardless of source.

Much will depend on the "time frames" that the LSSA chooses to set in its access protocols regarding timely submission of non-text-searchable material. The LSSA may decide to retain a schedule that is consistent with the NRC/DOE Site-Specific Procedural Agreement of September 18, 1984, which specified that data be provided on request following QA checks that normally occur within 45-days from the date of acquisition (DOE, 1990). While it seems doubtful that the LSSA would set a more stringent deadline than that, it might establish a more lenient one of a few months, or even longer. Even a six-month allowance, however, would be insufficient for some investigations, making it necessary to consider the practicality of divided package submission.

There are many conceivable ways that packaged materials might be submitted to the LSS under such circumstances, with tables of contents and bibliographic headers to describe them. Only a few options, however, are viable. The pros and cons of three basic alternatives will be discussed.

5.1 ALTERNATIVE 1: SUBMIT A TIP TO THE LSS ONLY WHEN IT IS COMPLETE

Under this alternative, no package segments would be submitted to the LSS (Figure 5-1). However, the nature and location of quality-assured materials that are accumulating as the result of on-going investigations would be made known through the issuance of periodic reports by the responsible organization. These reports would be entered into the LSS so that available material could be requested.

The TIP itself would be submitted to the LSS, with a descriptive table of contents and related bibliographic header, only when the investigation is concluded and the package is complete, in compliance with all appropriate quality assurance procedures. Upon submission, all of a TIP's collected materials would be scanned to produce electronic images viewable on LSS screens and converted to searchable ASCII text, insofar as the varied nature of the materials allows.

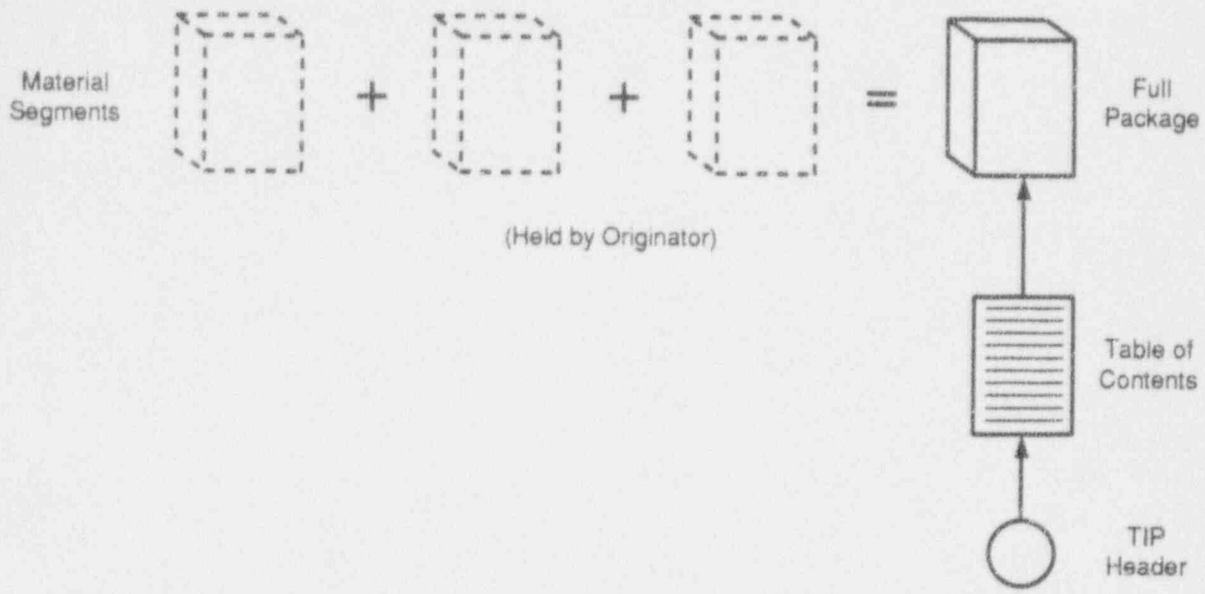

Figure 5-1. Alternative 1: Submit a TIP to the LSS Only When it is Complete

Even if the responsible organization divides TIPs for its own internal record control purposes, those segments would be retained until the entire TIP is finished, at which time the TIP would be submitted to the LSS intact.

The responsible organization would issue periodic reports describing documentary material that is accumulating as a result of on-going investigations, has been acquired in conformance with quality assurance procedures, and may be of use to other organizations. These reports would be entered into the LSS on a scheduled, timely basis as documentary material. Their entire contents would be text searchable. If the responsible organization received a request for items mentioned in its reports, it would either make copies for requestors or allow requestors to examine the materials on site.

The DOE/YMPO appears to be proceeding along these lines. As mentioned above, it will be publishing a quarterly Data Catalog which will describe all of the materials produced by its contractors. Each successive version of the Data Catalog will be entered into the LSS in text-searchable form, permitting requests to be made for items listed within it. In addition, the DOE/YMPO may grant other LSS participants direct access to its Automated Technical Data Tracking System (ATDT), also mentioned above, which can be used in between Data Catalog updates to provide reference to the nature and storage location of accumulating material.

ADVANTAGES:

- The submission of comprehensive TIPs to the LSS is faithful to the concept of packaging that the drafters of the LSS Rule apparently envisioned.

- This alternative avoids the complexities that would result from artificial division of investigative activity and/or the TIPs they produce. It would also avoid complexities that

would result from the construction of multiple tables of contents concerning related activity, and from the creation of bibliographic headers leading to those tables of contents.

- The additional cost which complication and increased work inevitably imposes would be avoided.

DISADVANTAGES:

- Packaged material awaiting the conclusion of an investigation and the resulting completion of a TIP would be unavailable to LSS requestors as viewable images on their LSS screens (if that material is imageable) and would have to be obtained through special request. This would be less convenient and would involve delay.

- The need to order unsubmitted material from the originating organization would create an undesirable dependence upon that organization, alerting it to particular requestor concerns which a requestor might not wish to divulge.

- The descriptions of unsubmitted but available data in periodic reports would probably be less satisfactory than descriptions that are contained in tables of contents or bibliographic headers.

- LSS users would not have assurance that the materials they have examined as a result of requests for early release based on a periodic report would not change prior to the materials' final submission to the LSS within a TIP.

- This alternative would require all organizations to issue periodic data-availability reports such as the DOE/YMPO's Data Catalog, although it is less likely that other organizations would have prolonged investigations requiring such reports.

- To the extent that materials within unfinished packages cannot be duplicated for examination by requestors and would consequently have to be examined on site, there would be travel expense and increased time and effort involved on the part of requestors.

5.2 ALTERNATIVE 2: SUBMIT A TIP IN SEGMENTS AND EXTENSIVELY INDEX ITS CONTENTS

Under this alternative, a package would be submitted to the LSS in periodic segments meeting time frames that will be established by the LSSA (Figure 5-2). The submission of a complete table of contents for the TIP and a bibliographic header leading to that table of contents, both required by the LSS Rule, would accompany the submission of the final segment. Multiple headers describing all of the individual items (or item-groupings) that a segment contains would be created by the responsible organization and submitted with each segment. Those headers would permit LSS users to find and retrieve any of the submitted items through the convenient means of bibliographic search.

A segment's materials would be scanned upon receipt to produce viewable electronic images for the LSS and would be converted to searchable ASCII text, insofar as the nature of the individual items permits.

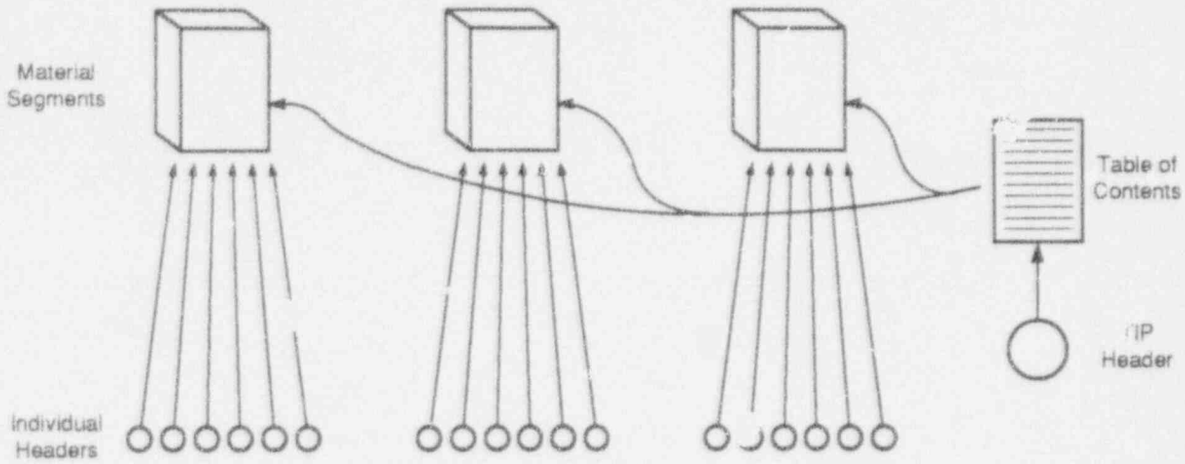

Figure 5-2. Alternative 2: Submit a TIP in Segments and Extensively Index its Contents

While the images and text of the successive segments would never be co-located physically within the LSS, electronic linkage using the available Package ID field within the bibliographic headers of the individually indexed packaged items would relate all of the items to one another. This linkage would be available both before and after package completion and would effectively consolidate a TIP from the viewpoint of the LSS requestor, rendering its continued physical division within the LSS data base "invisible."

ADVANTAGES:

- This alternative would make documentary material directly available through the LSS on a timely, periodic basis and avoid all of the disadvantages that apply to Alternative 1.

- It could also eliminate the need for manual drafting of a comprehensive table of contents when a TIP is completed, in view of the fact that a table of contents could be generated electronically at any stage in the growth of a package by bringing together the titles of the headers that have been created thus far for the individual items within the TIP, satisfying the requirement of the LSS Rule in that manner.

- There would be enhanced searchability of packaged materials as the result of more thorough utilization (sub-division) of a TIP's contents, as discussed in the CNWRA report on this topic (CNWRA, 1991b). A requestor would be able to find all viewable raw data items through bibliographic search, without also having to conduct text searches of TIP tables of contents.

- The formation of adequate descriptions for packaged items would be more strictly controlled through the discipline and automated error-checking routines that are inherent in the creation of individual headers for materials. Listing them on a table of contents will necessarily be less restrictive, even if the instructions are flawless.

DISADVANTAGES:

- The primary disadvantage of this alternative would be the cost of the extensive additional labor that would be required to apply headers to the thousands of individual raw data items (or groups of items) appearing within completed and future TIPs, even if the items were grouped to a greater extent than they have been by the DOE/YMPO. It is difficult to measure the incremental cost because the CNWRA has seen no sound estimates on the number of TIPs that are expected to be entered into the LSS or of the average number of items that will be included within them. However, in view of the fact that the labor-intensive indexing function is recognized to be one of the most expensive aspects of any information retrieval system, increasing its cost by any factor must be considered significant. It was largely for reasons of this anticipated cost that the CNWRA recommended (CNWRA, 1991b) that individual indexing of packaged raw data items should be limited to those that are non-imageable (machine-dependent for the most part).

- There are other disadvantages in extensive headering, aside from cost. Whenever more labor is involved, opportunity for error is increased. For instance, a vital packaged item might, through indexing error, be disassociated from its package. Also, the application of multiple references to the same package could mislead LSS users into thinking that they had uncovered more expansive information on a given subject (investigation) than is actually available and complicate the straightforward retrieval of a TIP.

5.3 ALTERNATIVE 3: DIVIDE INVESTIGATIONS INTO PHASES FOR TIMELY SUBMISSION OF TIPS

Under this alternative, an investigation that is expected to be prolonged, taking several months to complete, would be divided into practical phases (sub-investigations, in effect) that would permit submission of the resulting packaged material to the LSS with time frames and standards of compliance that would be established by LSSA protocols in accordance with the LSS Rule (Figure 5-3).

Packages would not be submitted in "segments," since it would be the investigations, not the packages, that would be divided. Each TIP, documenting a completed investigative phase, would contain all of the materials produced during that phase which are in conformance with appropriate quality assurance procedures. Each TIP would receive its own title, table of contents, and package header. Each one would be immediately scanned to produce viewable images and searchable text for the LSS, insofar as possible.

The "Related Documents" field in the headers for these TIPs would be used to interrelate these common-subject packages as they are submitted, through a special prefix. Each successive TIP concerning the broader, prolonged investigation would "point" through this field to the last one that was entered into the LSS. The "Package ID" field, in contrast, would be reserved to connect packaged items having their own headers (e.g., non-imageable raw-data items) to their parent TIPs, in the same way that this field would be used for the first two alternatives. This will be explained further in Appendix E.

When the prolonged investigation is completed, the final package would logically contain the finished product for that broader investigation, if any, together with appropriate approval and review documents. Those important products and documents would receive individual headers (as would non-imageable raw-data items), according to the CNWRA's recommendation (CNWRA, 1991b).

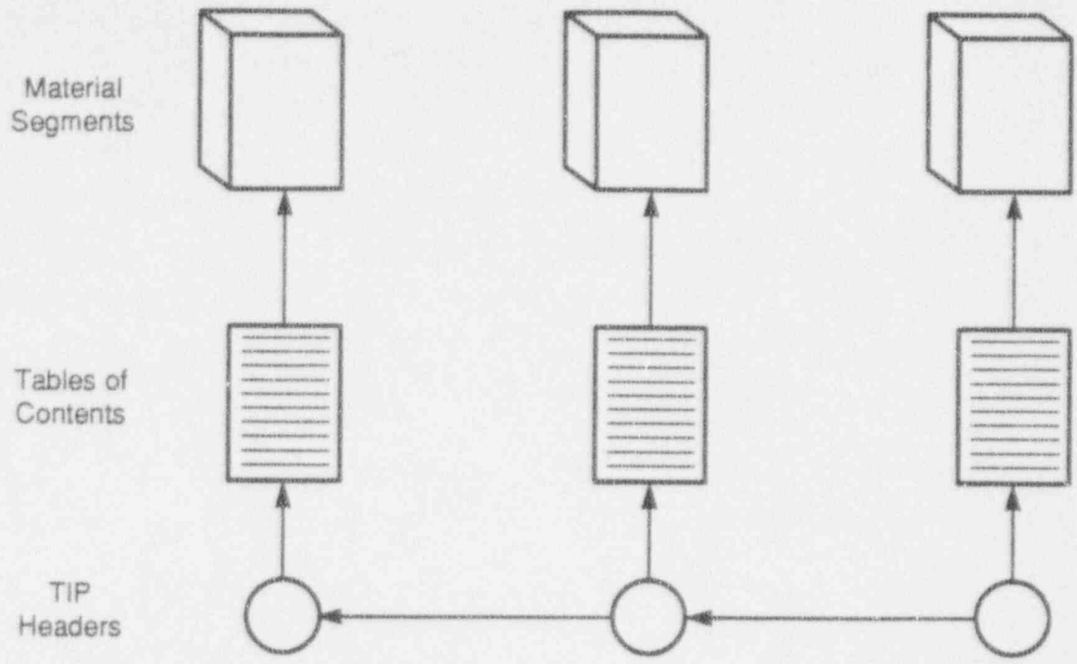

Figure 5-3. Alternative 3: Divide Investigations into Phases for Timely Submission of TIPs

This alternative would appear to deprive the LSS requestor of a single, comprehensive table of contents that would describe all of the materials relating to a prolonged investigation, and of a bibliographic header leading to a comprehensive table of contents. However, the investigation would retain its integrity by means of links (in the form of related-document pointers) which would permit the tables of contents for closely related packages to be electronically combined. This would permit a requestor to gain immediate access to materials in the packages through a composite, computer-generated table of contents. The existence of multiple TIPs concerning the broader investigation would be transparent to the user.

ADVANTAGES:

- This alternative would make documentary material directly available through the LSS on a timely basis and avoid all of the disadvantages that apply to Alternative 1.
- It would avoid the increased costs inherent in Alternative 2.

DISADVANTAGES:

- The division of investigative activity into phases would sometimes be problematic, if not impossible. For example, a test that would require measurements of the same kind to be made for a year or more could be divided into phases only by artificial time slices that would conform to the upper limits for material submission imposed by LSSA protocols. Forced submission of that kind would serve no purpose, since it would not provide useful information to the LSS. Therefore, if this alternative were adopted, exceptions would probably have to be made, allowing some prolonged investigations to produce TIPs only

when they are concluded. This allowance could be made on a general, circumstantial basis by the protocols themselves or on an approved, case-by-case basis by the LSSA.

- This alternative would lack the advantages of greater searchability of packaged materials and the more tightly controlled descriptions of them that Alternative 2, with its exhaustive indexing, would afford.

- Compared with Alternative 1, the opportunity for error is increased by the need to incorporate extra links.

5.4 RECOMMENDATION

The CNWRA considered other alternatives, but found them to be either impractical or overly complex. Any alternative that would hamper investigative activity or result in the premature release of investigative results would be intolerable. Any alternative that would require the replacement of earlier collections of images by subsequent accumulations would be unacceptable, because the LSS must maintain strict document integrity. Any alternative that would involve the replacement of obsolete tables of contents or obsolete bibliographic headers for TIP segments by up-to-date or final versions would be unworkable. Leaving obsolete headers or obsolete tables of contents relating to TIP segments in the system would create undesirable, if not unacceptable, complication and clutter.

Table 5-1 summarizes the comparative advantages and disadvantages of the three alternatives that have been discussed. On balance, the CNWRA recommends Alternative 3 as the option that would solve the problem of delayed material availability from prolonged investigations, while avoiding extensive additional cost. This alternative appears to satisfy the LSS Rule with regard to material submission and provides adequate retrieval capability for LSS users. The impact of this recommendation on the DOE/YMPO and on other organizations that submit TIPs should not be excessive.

The thorough indexing of individual (or grouped) packaged materials that forms the basis for Alternative 2, of course, could be applied to TIP submission whether or not packages are divided for timely submission. It has been presented as an alternative for this discussion because, in the CNWRA's view, it offers the only viable way in which TIPs could enter the LSS in segments. If the LSSA decides to adopt protocols requiring thorough indexing of the viewable raw materials within TIPs, that approach could be incorporated into the procedures of Alternatives 1 or 3, as well as Alternative 2. The CNWRA report that focused on the subject of TIP indexing (CNWRA, 1991b) debated thorough versus selective indexing of packaged materials and recommended the latter.

Table 5-1. Summary of the Comparative Advantages and Disadvantages of the Three Alternatives

	Alternative 1 (No Segments)	Alternative 2 (Segment & Index)	Alternative 3 (Use Phases)
Material Availability	Delays	Timely	Timely
Relative Cost Differential	Small	Large	Small
Simplicity of Retrieval	Least Complex	Redundant Hits	Links Required
Consolidated Table of Contents and Header	Drafted	Automatic	Links Required
Enhanced Searchability	No	Yes	No
Discipline of LSS Entry	Less Controlled	Tightly Controlled	Less Controlled
Opportunity for Indexing Error	Least	Most	Some
Need for Extra Data-Availability Reports	Yes	No	No
Dependence on Originator for Interim Access	Yes	No	No
Need to Sub-Divide Investigations	No	No	Yes

6 DESCRIBING A TIP IN A TABLE OF CONTENTS

The LSS Rule requires that a table of contents be written for every package of documentary material, describing the items contained. This listing will be made by the originator of the package, usually by the principal investigator, assisted if necessary by the responsible organization's records management staff.

Typically, the bulk of a TIP will be composed of the numerous items of non-text-searchable "raw data" that have been produced by the subject investigation. These data may include a wide variety of materials: surveys, test results, logs, notes, measurements, calculations, graphs, sketches, etc. Many will exist in rough, handwritten form; others will be computer generated. If a finished product, such as a technical report, structural design, or a map, is created as a result of the investigation, it will be featured as the primary document within the package. Letters of approval for the investigation, circulated drafts of the product, and internal reviews of the product are other documents that will be included in the package when they are forthcoming from the investigation. All of these materials will be listed on the TIP table of contents as individual items, except for raw data, which may be generically grouped (e.g., drilling reports, pumping test data).

It is imperative that a table of contents be thoughtfully written because it will serve as the most comprehensive description of the package within the LSS. A bibliographic header will be created for every package, but there is only so much information that a header can impart about the content of a large collection of diverse materials. A retrieved package header will lead a requestor to the more detailed table of contents, which will be displayed on the LSS screen as a menu, in effect, for selective viewing of the assembled materials. As in the case of other LSS documents, an alternative route to this on-screen menu will be through full-text search, which will yield tables of contents that contain specified words or phrases. The CNWRA has recommended that TIP tables of contents be kept in a separate partition of the LSS textual data base to assist this type of search (CNWRA, 1991b).

Careful construction of a table of contents is also of critical importance because it will be the only avenue of retrieval for most of the raw data that packages will contain. The finished products, approvals, reviews and circulated drafts within a package may be individually retrieved, independently of the table of contents, by means of the bibliographic header that will be assigned to each of these documents. They may also be individually retrieved by means of full-text search on the basis of selected words or phrases, since they will all be textual in nature. In contrast, according to the CNWRA's recommendation, few raw data items will be assigned headers of their own, because of their integral relationship to a packaged collection, their essential interdependence, and their inability to be properly understood outside the context of their package (CNWRA, 1991b). Being primarily graphic or handwritten in nature, these items (lithologic logs, for example) will generally lack searchable text which would permit alternative full-text-search retrieval.

The CNWRA has recommended (CNWRA, 1991b) that an exception be made in the case of non-imageable raw data—the data which is mostly machine-dependent (resident on magnetic tape, for example) or cannot otherwise be usefully scanned to produce viewable images, and which must therefore be stored on the premises of the originating organization, requiring requestors to make special access arrangements. These non-imageable items should be assigned their own headers to assist common retrieval by document type and/or media type when desired. A principal investigator would provide specifics for these headers by filling out input forms detailing the precise content, nature, location, and means of obtaining the stored

items. These input forms, which would not have to adhere to an LSS standard, would be used by indexers to create descriptive headers leading to the materials. They would serve additionally as "slip sheets," replacing the items themselves within a TIP. After their conversion to electronic images, they would become viewable pages, supplementing the necessarily brief table of contents descriptions. The DOE has been using such slip sheets for its microfilm-based records system. (Input forms for non-imageable materials that do not belong to a package may be destroyed after the header is created.)

Someone interested in finding substantiating evidence for a particular investigative conclusion will presumably wish to see, first of all, an itemization of the materials supporting the investigation's findings. A package table of contents is intended to serve that critical function. It provides a convenient overview. The items it lists will be clearly "named and identified," to use the words of the LSS Rule, so that a requestor may proceed to inspect those of possible interest. Sometimes, the requestor will be able to view selected materials directly on the LSS screen. In other instances, when the data is on non-imageable media, the requestor will have to make arrangements with the originating organization to examine them.

The CNWRA has recommended (CNWRA, 1991b) that, when a table of contents is retrieved through the LSS, the capability of selecting items from it for on-screen viewing should be enhanced through hypertext indexing. Such indexing will enable a requestor to "turn" immediately to the beginning page (image) of any selected item by using a pointing device or a single keystroke. Viewable items and slip sheets for non-imageable items may be selected. Hypertext capability will obviate the need for consecutive numbering of package pages in order to place beginning page numbers beside each listed item on the table of contents. (Page numbering for the purpose of selective reading from a table of contents is ordinarily a necessity for any large document or assemblage of material.) If hypertext indexing were not employed, page numbers would have to be applied to the individual pages of a TIP and to the table of contents listing, to permit a requestor to view selected items conveniently without having to flip tediously through numerous unwanted images before finding them.

A mandatory format, for use by all LSS participants, should be provided for a TIP table of contents because this description/selection vehicle should have a consistent appearance across organizational lines, thereby enabling LSS users to find the materials they seek quickly and without complication.

Accordingly, Appendix D provides instructions for the completion of a TIP table of contents, with a sample blank format and a completed example. The instructions and format are intended for use by all LSS participants that submit documentary material in package form.

All of the items listed under the "Raw Data" section of the table of contents will belong to one of the eight categories defined in Appendix B of this report. As mentioned earlier, the criteria normally used for retrieval purposes for all of these categories is expected to center upon technical subject. Since all of the items within a given package have a common subject (which makes them a package) and the subject is featured prominently in the title of the table of contents, there is obviously no need to repeat the subject for each individual item. For instance, in the example table of contents provided in Appendix D, there is no need to say anything more about the "Borehole Geophysical Log." The package title indicates that it pertains to a certain test well which was investigated by the U.S. Geological Survey (USGS). The fact that no sub-contractor is noted next to the item indicates that none was used.

In contrast to a bibliographic header, there is no "field" to identify an item's document type in a table of contents. That identification will be made on the listing if it is essential to the adequate description of an item. In the case of the field notebook in the example, it is inherently identified by type (as a likely

sub-category of the handwritten notes category). In the case of the lithologic log (which also happens to be also a set of handwritten notes), it is not, because this type of log is quite specific and adequately described for what it is.

The proposed table of contents format, it will be noted, is fundamentally the same as the "model" developed at the USGS to encourage consistency in the diverse tables of contents that have been employed by the DOE's Yucca Mountain Project Office (DOE/YMPO). The DOE/YMPO's general procedure that includes instructions for the preparation of a package table of contents (DOE, 1992), in contrast, contains minimal requirements. Aside from cosmetics, the CNWRA's proposed format differs from the USGS model and DOE practice in only three essential respects:

(1) All of the persons mentioned in the table of contents (contributing authors, etc.) would be identified using their full first names and middle initials, not merely initials preceding the surname.

(2) Any participating subcontractor organizations would always be named next to the items that they produced.

(3) All internal YMPO/USGS control numbers and symbols are excluded in view of the fact that the proposed form will be used by all organizations that submit packages, including the YMPO and its contractors. Each organization may add internal control numbers to the form as necessary.

7 DESCRIBING A TIP IN THE LSS BIBLIOGRAPHIC HEADER

A TIP may contain a wide variety of documentary material. As envisioned by the LSS Rule, the bulk of a package will ordinarily be composed of non-text-searchable "raw data," including closely associated handwritten items (such as notebooks), graphic items (such as maps), and non-imageable items (such as numeric data on magnetic tape)—all of which have been classified and described in Sections 2 and 3 of this report. Typewritten, text-searchable documents are logically and conveniently included in TIPs when the investigation has resulted in a technical report and in associated approval and review documents, including circulated drafts.

As the LSS Rule requires, every package will be digitally scanned to produce a set of electronic images which may subsequently be viewed on LSS computer screens. If any items in a package are not already printed on paper or other "hardcopy" media that will permit scanning, they will be converted to paper, from the original film, mylar, or other duplicable media, so that they may be transformed into electronic images—provided that this can be done without loss of information content.

The items in a package that are inherently unsuitable for scanning will be stored by the originating organization for examination upon request, as discussed in Section 3. The header input forms that will be written to describe those items will be inserted into the package in their place, as "slip sheets," for scanning with the rest of the materials.

Following the scanning process, an LSS capture station will convert to ASCII form any full-text items that may exist in a package (reports and approval/review documents in particular), so that they may be retrieved through text search. The LSS Rule states that "text embedded within (graphic-oriented) documents) need not be separately entered in searchable full text." However, to broaden text search capability, the CNWRA has recommended that an attempt be made to optically recognize the text of all packaged materials scanned by a capture station, including graphic-oriented raw-data. It has done so with the caveat that manual (human) intervention, which normally resolves interpretive ambiguities that confuse automatic character recognition, be avoided altogether for packaged raw data. Otherwise, interruptions calling for operator resolution could easily overwhelm the capture process, requiring unjustifiable expenditures for the limited amount of typewritten text likely to be captured within these materials. The suggested process would be expected to capture printed explanatory comment accompanying a map, for example, while it would presumably ignore miscellaneous notations on the map itself, depicting the names of mountains, rivers, etc.

A bibliographic header will be submitted for every package, based upon its table of contents, in accordance with the LSS Rule. As an integral, mutually-dependent collection of material, a package will thus be indexed, in effect, as a single unit, receiving a descriptive header which will lead an interested party to its more thoroughly descriptive table of contents. The package header will thus identify a TIP, providing reference to it, while its table of contents will describe the items it holds.

The CNWRA has recommended (CNWRA, 1991b) that finished products and non-imageable (mostly machine-dependent) items within a package receive bibliographic headers of their own, independent of the header that will be applied to the package as a whole, in order to facilitate their individual retrieval apart from associated packaged materials. In consultation with the LSSA, the CNWRA has since agreed that approval/review documents, including circulated drafts, should also be indexed for this same reason.

Headers for TIPs and for the individually indexed items they contain will be created by the organization that originated the package and will be finished, through the completion of certain fields (e.g., Accession Number and Pointers), by personnel working at the LSS capture station.

Entries for headers describing materials resident on non-imageable media will be taken from the input forms (slip sheets) completed by TIP investigators. If a given tape, disk, etc., contains information on more than one subject, a separate header should be created for each subject.

There is only one circumstance in which the CNWRA would recommend the provision of a header to a raw data item within a TIP that has been successfully converted to an electronic image: when an item that was included in an earlier TIP is included in a subsequent TIP. The Related Documents and Pointers fields should provide mutual reference in such an instance.

The CNWRA foresees that a TIP is most likely to be retrieved on the basis of three essential attributes: its technical subject (the primary criterion), its origin, and its date. A requestor will want to know first and foremost what investigation was undertaken on a particular subject, what organization performed the work, and when that work was done. Searches based on these criteria are possible using corresponding bibliographic header fields that have already been approved by the LSS Advisory Review Panel (LSSARP): Title, Abstract, Descriptors, and Identifiers (for technical subject), Author Organization (for origin), and Document Date.

After retrieving a TIP through a search of these fields (and possibly others, such as Author Name or Document Type), individually or in some combination, a requestor will be presented with a table of contents, which will provide a convenient overview of the package, as well as more descriptive detail than was available in the bibliographic header leading to it. The table of contents will permit thorough or selective examination of the package's finished product, its approval/review history, and/or its associated raw ("backup") data.

The imageable materials contained in a package, whether they are typewritten, handwritten, or graphic in nature, will be immediately viewable on the LSS screen through the use of hypertext indexing. When a requestor wishes to inspect non-imageable items (resident for the most part on machine-dependent media), however, the requestor will have to make other arrangements to obtain them, having learned their content, specifications, and storage location from descriptions contained on the viewable "slip sheets" that will replace those items within the package. The slip sheet descriptions will have served as the basis for the bibliographic headers that will be created for each of the items. Those headers will always be available as alternative references to non-imageable items.

Several LSSARP-approved bibliographic header fields are expected to be of use to a requestor in seeking packaged documentary material, in addition to the primary ones mentioned above (Title, Abstract, Descriptors, Identifiers, Author Organization and Document Date). In particular, the Package ID and Related Documents/Pointers fields will be helpful. The Package ID field was adopted by the LSSARP to relate all parts of a given package to one another—those having their own individual headers, such as technical reports and non-imageable materials. The CNWRA proposes that the Related Documents field (converted to standardized form in the Pointers field) be used to interrelate packages relating to a prolonged investigation, as mentioned in Section 5 of this report, under Alternative 3.

The CNWRA sees no need for the LSSARP to consider additional fields to describe a TIP itself in an LSS bibliographic header. However, it does see a definite need for added fields to describe material

resident on non-imageable media, whether or not that material is included within a TIP. Those fields are a Media Type field and a Storage Location field. In addition, the CNWRA recommends that a Quality Assurance Status field be added to the LSS header for all materials.

The most orderly and comprehensive way to describe the LSS header's applicability to TIP-units themselves and to individually indexed materials that are packaged or unpackaged (not collected within a TIP) will be to comment upon each of the fields that the LSSARP has approved (HWG, 1990 and HWG, 1991), supplementing the LSSARP-approved descriptions and the instructions contained in the draft LSS Cataloging Manual (SAIC, 1990b). This is done in Appendix E.

The Appendix E recommendations are intended to apply as noted to headers for TIPs, for individually-indexed packaged items, and for individually-indexed, unpackaged non-text-searchable materials. Little will be said in Appendix E about the way that packaged technical reports and approval/review documents should be treated in the header, as these materials should all be indexed in basically the same way that other reports, memoranda, and correspondence are indexed for the LSS.

8 CONCLUSION

Section 2 of this report defined eight categories of documentary material that are unsuitable for entry into the LSS in the form of searchable full text: maps, photographs, design drawings, figures, data sheets, data plots, computer programs and handwritten notes. Appendix A discussed the CNWRA's attempt to classify these "document types" further, into sub-categories. Appendix B described the individual attributes of the eight categories and suggested submission requirements.

Section 3 discussed the kinds of non-text-searchable material that are likely to be found on non-imageable media. A matrix illustrated the probable relationships between the eight document types and twelve identified non-imageable media types (all machine-dependent): magnetic tape, magnetic disk, solid-state device, optical disk, videocassette/videodisc, audio recording, motion picture film, microform, photographic slide, photographic negative, develocorder film, and radiograph. Appendix C described the respective attributes of these media types which must be conveyed to requestors. Section 3 also discussed how non-imageable items could be retrieved from dispersed storage locations. It recommended that when requestors lack access to necessary equipment to read these items they be advised to examine them on site, in view of the fact that any conversion might result in an unacceptable loss of information content.

Section 4 proposed a working definition for a "Technical Investigation Package" ("TIP"), to help clarify the intent of the LSS Rule's stipulations regarding the submission of a "package of information." TIPs are expected to contain most of the non-text-searchable material that will become available through the LSS.

Section 5 discussed how TIP material could be submitted to the LSS, within LSSA-mandated time frames, when a lengthy investigation occurs. It suggested three alternatives, recommending that a lengthy investigation be divided into phases that will permit timely submission of material.

Section 6 recommended a method of creating a table of contents for a TIP. Appendix D provided proposed instructions, with a sample format and a completed example.

Section 7 recommended a coherent way in which attributes of TIPs and individually indexed non-text-searchable materials could best be recorded within the LSS bibliographic header. Appendix E provided detailed suggestions for the use of approved header fields and recommended three additional fields: Media Type, Storage Location, and Quality Assurance Status.

Within this report, certain recommendations made by the CNWRA in its last report (CNWRA, 1991b) were repeated:

(1) All imageable material, packaged or not, should be scanned to produce LSS images, including "oversize" and color-significant items. The CNWRA assumes that, by the time it becomes operational, the LSS will employ technology to scan and display such items and that all of the LSS workstations will incorporate graphic-display capability.

(2) In addition to the creation of a bibliographic header for a TIP itself, headers should be created for finished products and for non-imageable items within a package. The present

report added that approval/review documents within packages should also be individually indexed.

(3) TIP tables of contents should be placed in a separate partition of the LSS database to assist full text search. The earlier report also recommended that the LSS be designed to permit a simultaneous search of text and header fields, if desired, a suggestion that was supported by the LSS Prototype Test Report (SAIC, 1990a).

(4) Hypertext technique should be adopted by the LSS to facilitate the display of viewable packaged materials.

(5) Imageable raw data within TIPs, which contain only occasional "embedded," typewritten passages suitable for text search, should be subjected to automatic text indexing that avoids manual intervention.

The CNWRA believes that the impact of its recommendations on LSS participants, including the DOE, would not be excessive, even if they are applied retroactively, as the CNWRA recommends.

9 REFERENCES

Center for Nuclear Waste Regulatory Analyses (CNWRA). 1991a. *Definition of Technical Data and Analysis of Infrastructures Within Participant Organizations for Providing LSS Access*. San Antonio, TX: CNWRA.

CNWRA. 1991b. *Alternative Ways of Making Packaged Documentary Materials Accessible Within the Licensing Support System*. San Antonio, TX: CNWRA.

Department of Energy (DOE). 1990. *Technical Data Management Plan*. YMP/88-18, Revision 1, Las Vegas, NV: DOE.

DOE. 1992. *Records Management: Record Source Implementation*. YMPO SP 1.36, Revision 8, Las Vegas, NV: DOE.

Header Working Group (HWG). 1990. *Recommended Fields for LSS Header Records*. Marshall, VA: HWG.

HWG. 1991. *Additional Fields for Headers*. Marshall, VA: HWG.

Licensing Support System Administrator (LSSA). 1991a. *LSSA Comments on Definition of Technical Data and Analysis of Infrastructures Within Participant Organizations for Providing LSS Access*. Washington, DC: U.S. Nuclear Regulatory Commission (NRC).

LSSA. 1991b. *Technical Direction on CNWRA Task 1—Access Protocols*. Washington, DC: NRC.

LSSA. 1992a. Comments on CNWRA March 1992 Report, *Classification and Attributes of Non-Text-Searchable Documentary Material*. Washington, DC: NRC.

LSSA. 1992b. *LSSA Program Element Plan for CNWRA for FY 92 Through FY 94*. Washington, DC: NRC.

NRC. 1989. *10 CFR Part 2, Subpart J, The LSS Rule*. Washington, DC: NRC.

Science Application International Corporation (SAIC). 1990a. *Licensing Support System Prototype Test Report*. McLean, VA: SAIC.

SAIC. 1990b. *Licensing Support System Cataloging Manual (draft)*. McLean, VA: SAIC.

SAIC. 1990c. *Revised Cost Analysis for the LSS*. McLean, VA: SAIC.

APPENDIX A

EFFORTS TO IDENTIFY SUB-CATEGORIES OF NON-TEXT-SEARCHABLE MATERIAL

EFFORTS TO IDENTIFY SUB-CATEGORIES OF NON-TEXT-SEARCHABLE MATERIAL

A group of eight CNWRA technical staff members representing the geosciences, materials, mining engineering and quality assurance disciplines met to identify potential sub-categories of the eight primary categories previously identified by the CNWRA. The group initially attempted to think of as many subcategories as possible, then considered whether the total of the sub-categories under a major category could be exhaustive, and whether the individual sub-categories would be exclusive of each other.

Later, the sub-categories identified by the technical staff were supplemented by adding all of the previously identified examples and by appropriate examples (likely to be non-text-searchable) of Lifetime Quality Assurance records from standard ASME NQA-3-1989, "Quality Assurance Program Requirements for the Collection of Scientific and Technical Information for Site Characterization of High-Level Nuclear Waste Repositories." DOE Quality Assurance program documents were also reviewed to identify other possible sub-categories.

The list then included a fairly broad section of the types of non-text-searchable materials that are anticipated. The listed sub-categories ranged widely in their level of detail, or specificity, so tiers of categories were developed, the members of succeeding tiers being increasingly specific. For each tier, the criteria of exhaustivity and exclusivity were applied. Finally, the tiered structure was examined to determine if the classification scheme could be used as a document type scheme, holding true to the form and format characteristics of document type. Classification schemes based on other characteristics, subject matter in particular, are covered by the Descriptors field in the LSS bibliographic header, which has an extensive controlled Thesaurus. The Title, Abstract, and Identifiers fields will also provide valuable subject-related search capability.

A summary of the non-text-searchable document category analysis follows.

Map: Five possible first tier sub-categories were identified that were considered to be exhaustive. Most maps could fit into more than one of these sub-categories because mixing of information on maps is common. Applying the final criterion, the first tier (and hence all lower tiers) sub-classification of Map was based on content (subject matter) of the map, which violates the essential characteristic of document type. Second tier sub-categories were identified for four of the five first-tier sub-categories. Being more specifically subject oriented, these were more exclusive, but exhaustive lists would be difficult to develop and exceedingly lengthy.

Photograph: The four first tier sub-categories included three fairly specific types of photographs, and one much more general sub-category to catch the balance, so the categorization scheme is not uniform. Some more specific second tier sub-categories were identified, but as with the first tier, these classifications are based on content rather than form or format.

Design Drawing: Six sub-categories that appeared exhaustive were identified, however exclusivity was poor, that is, any drawing could be likely to have the characteristics (subject) of more than one sub-category. No lower tiers were identified.

Figure: Two tiers of sub-categories were developed, but none of the lists seemed to be exhaustive because of the wide variety of possible type of figures. Classification was based on content or subject matter.

Data Sheet: A relatively large number of first tier sub-categories were identified because of the wide range of possible subjects of data sheets. Even so, the lists were probably not exhaustive or mutually exclusive because data content may cross into several categories. Sub-tiers were more exclusive, but less exhaustive. Sub-classification again tended to be based on subject matter.

Data Plot: As an alternative method of presenting data, the same comments apply as to data sheets.

Computer Program: While this category included five first-tier sub-categories (two of them with a second tier) that seemed exhaustive and exclusive, all were subject matter related.

Handwritten Notes: Three first tier (subject related) subclasses were identified that were possibly exhaustive and exclusive. A second tier of notebooks had three subject-based designations that had relatively poor exclusivity.

Conclusions: The technical staff had varying success in developing exhaustive and mutually exclusive sub-categories of non-text-searchable documentary materials. However, the subclassification schemes invariably resorted to subject-matter-based distinctions, which conflicted with the form and format characteristics of a document type classification scheme.

APPENDIX B

CATEGORIES OF NON-TEXT-SEARCHABLE DOCUMENTARY MATERIAL

CATEGORIES OF NON-TEXT-SEARCHABLE DOCUMENTARY MATERIAL

CATEGORY: Map

DEFINITION: Representation of a geographic area, often highlighting certain natural or constructed features (e.g., topography, geologic structure, physical properties, roads, buildings). Map overlays and sketches are included.

PHYSICAL ATTRIBUTES: A map may be recorded on scannable media, such as paper, or on non-imageable media, such as magnetic tape.

EXAMPLES: Geologic, geographic, topographic, hydrologic, and geophysical maps, including cross-sections, site surveys, probe locations, etc.

CRITERIA NORMALLY USED FOR RETRIEVAL PURPOSES: Subject, origin, and date.

OTHER CHARACTERISTICS: Specific extent of geographic area, scale, size, color, legibility, and Quality Assurance (QA) status.

UNITIZATION GUIDELINES: If a map exists as a stand-alone item (apart from a document or package) it should be indexed as an independent unit. If it exists within a collection of maps on the same subject (apart from a document or package) and is best interpreted within the context of that collection, the collection should be indexed as a single unit. If a map occurs within a document and would be both useful and fully understood outside the context of that document, it should be independently indexed. If a map is part of a package, it should be independently indexed only if it is recorded on machine-dependent media or cannot be converted to an electronic image without loss of information content.

RELATED DOCUMENTS: If a collection of maps is separated into multiple units for indexing purposes, the units should be related to each other through this LSS bibliographic header field.

SUBMISSION REQUIREMENTS:

ELECTRONIC IMAGE: Every printed map should be scanned, if possible, to produce a viewable image. If a map is not imageable, an effort should be made to convert it to scannable paper for LSS input, provided that the conversion will not result in a loss of information content. If conversion is undertaken, overlays for maps should be individually converted to paper covering the map they augment, so that the combination may be scanned to produce an electronic image.

ASCII TEXT: Embedded text that may exist on a map will be automatically converted to ASCII form for text search, insofar as possible, without manual intervention.

BIBLIOGRAPHIC HEADER: An independent header will be submitted if the above unitization guidelines require it.

CATEGORY: Photograph

DEFINITION: Still picture of an object or scene, obtained through the process of photography.

PHYSICAL ATTRIBUTES: A photograph may be recorded on scannable media, such as paper, or on non-imageable media, such as magnetic tape.

EXAMPLES: Aerial photographs, satellite images, photomicrographs, and simple camera shots depicting terrain features, physical samples, etc.

CRITERIA NORMALLY USED FOR RETRIEVAL PURPOSES: Subject, origin, and date.

OTHER CHARACTERISTICS: Color, size, legibility, and Quality Assurance (QA) status.

UNITIZATION GUIDELINES: If a photograph exists as a stand-alone item (apart from a document or package) it should be indexed as an independent unit. If it exists within a collection of photographs on the same subject (apart from a document or package) and is best interpreted within the context of that collection, the collection should be indexed as a single unit. If a photograph occurs within a document and would be both useful and fully understood outside the context of that document, it should be independently indexed. If a photograph is part of a package, it should be independently indexed only if it is recorded on machine-dependent media or cannot be converted to an electronic image without loss of information content.

RELATED DOCUMENTS: If a collection of photographs is separated into multiple units for indexing purposes, the units should be related to each other through this LSS bibliographic header field.

SUBMISSION REQUIREMENTS:

ELECTRONIC IMAGE: Every printed photograph should be scanned, if possible, to produce viewable image. If a photograph is not imageable, an effort should be made to convert it to scannable paper for LSS input, provided that conversion will not result in a loss of information content.

ASCII TEXT: Embedded text that may exist on a photograph will be automatically converted to ASCII form for text search, insofar as possible, without manual intervention.

BIBLIOGRAPHIC HEADER: An independent header will be submitted if the above unitization guidelines require it.

CATEGORY: Design Drawing

DEFINITION: Depicted schematic arrangement of the features of a plan for an object or facility.

PHYSICAL ATTRIBUTES: A design drawing may be recorded on scannable media, such as paper, or a blueprint, or on non-imageable media, such as magnetic tape.

EXAMPLES: Architectural, structural, mechanical, facility, instrumentation and electrical designs.

CRITERIA NORMALLY USED FOR RETRIEVAL PURPOSES: Subject, origin, and date.

OTHER CHARACTERISTICS: Size, scale, color, legibility, and Quality Assurance (QA) status.

UNITIZATION GUIDELINES: If a design drawing exists as a stand-alone item (apart from a document or package) it should be indexed as an independent unit. If it exists within a collection of design drawings on the same subject (apart from a document or package) and is best interpreted within the context of that collection, the collection should be indexed as a single unit. If a design drawing occurs within a document and would be both useful and fully understood outside the context of that document, it should be independently indexed. If a design drawing is part of a package, it should be independently indexed only if it is recorded on machine-dependent media or cannot be converted to an electronic image without loss of information content.

RELATED DOCUMENTS: If a collection of design drawings is separated into multiple units for indexing purposes, the units should be related to each other through this LSS bibliographic header field.

SUBMISSION REQUIREMENTS:

ELECTRONIC IMAGE: Every printed design drawing should be scanned, if possible, to produce a viewable image. If a design drawing is not imageable, an effort should be made to convert it to scannable paper form for input, provided that the conversion would not result in a loss of information content.

ASCII TEXT: Embedded text that may exist on a design drawing will be automatically converted to ASCII form for text search, insofar as possible, without manual intervention.

BIBLIOGRAPHIC HEADER: An independent header will be submitted if the above unitization guidelines require it.

CATEGORY: Figure

DEFINITION: Graphic representation of a form. The category excludes maps, photographs and design drawings (which are separate categories), but encompasses all other diagrams, drawings, sketches, etc.

PHYSICAL ATTRIBUTES: A figure may be recorded on scannable media, such as paper, or on non-imageable media, such as magnetic tape.

EXAMPLES: Planning chart, organizational chart, geological diagram, technical drawing, flora/fauna drawing, flow chart, and presentation material.

CRITERIA NORMALLY USED FOR RETRIEVAL PURPOSES: Subject, origin, and date.

OTHER CHARACTERISTICS: Size, scale, color, legibility, and Quality Assurance (QA) status.

UNITIZATION GUIDELINES: If a figure exists as a stand-alone item (apart from a document or package) it should be indexed as an independent unit. If it exists within a collection of figures on the same subject (apart from a document or package) and is best interpreted within the context of that collection, the collection should be indexed as a single unit. If a figure occurs within a document and would be both useful and fully understood outside the context of that document, it should be independently indexed. If a figure is part of a package, it should be independently indexed only if it is recorded on machine-dependent media or cannot be converted to an electronic image without loss of information content.

RELATED DOCUMENTS: If a collection of figures is separated into multiple units for indexing purposes, the units should be related to each other through this LSS bibliographic header field.

SUBMISSION REQUIREMENTS:

ELECTRONIC IMAGE: Every printed figure should be scanned, if possible, to produce a viewable image. If a figure is not imageable, an effort should be made to convert it to scannable paper for LSS input, provided that the conversion will not result in a loss of information content.

ASCII TEXT: Embedded text that may exist on a figure will be automatically converted to ASCII form for text search, insofar as possible, without manual intervention.

BIBLIOGRAPHIC HEADER: An independent header will be submitted if the above unitization guidelines require it.

CATEGORY: Data Sheet

DEFINITION: Numeric array in listed, tabular, or electronic form.

PHYSICAL ATTRIBUTES: A data sheet may be recorded on scannable media, such as paper, or on non-imageable media, such as magnetic tape (in which case it would appear on a "sheet" if printed). It is sometimes handwritten.

EXAMPLES: Test, drill-hole, geophysical, radioactivity, trench, atmospheric, environmental, and quality assurance data, including instrument readouts, calibrations, etc.

CRITERIA NORMALLY USED FOR RETRIEVAL PURPOSES: Subject, origin, and date.

OTHER CHARACTERISTICS: Legibility, page-count, and Quality Assurance (QA) status.

UNITIZATION GUIDELINES: If a data sheet exists as a stand-alone item (apart from a document or package) it should be indexed as an independent unit. If it exists within a collection of data sheets on the same subject (apart from a document or package) and is best interpreted within the context of that collection, the collection should be indexed as a single unit. If a data sheet occurs within a document and would be both useful and fully understood outside the context of that document, it should be independently indexed. If a data sheet is part of a package, it should be independently indexed only if it is recorded on machine-dependent media or cannot be converted to an electronic image without loss of information content.

RELATED DOCUMENTS: If a collection of data sheets is separated into multiple units for indexing purposes, the units should be related to each other through this LSS bibliographic header field.

SUBMISSION REQUIREMENTS:

ELECTRONIC IMAGE: Every printed data sheet should be scanned, if possible, to produce a viewable image. If a data sheet is not imageable, an effort should be made to convert it to scannable paper for LSS input, provided that the conversion will not result in a loss of information content.

ASCII TEXT: Embedded text that may exist on a data sheet will be automatically converted to ASCII form for text search, insofar as possible, without manual intervention.

BIBLIOGRAPHIC HEADER: An independent header will be submitted if the above unitization guidelines require it.

CATEGORY: Data Plot

DEFINITION: Plot or "graph" of digital numeric or analog data, depicting a comparison, series, progression, equation, measurement, etc., using coordinates.

PHYSICAL ATTRIBUTES: A data plot may be recorded on scannable media, such as paper, or on non-imageable media, such as magnetic tape. When printed, it is usually drawn by a digital or analog computer, but it may be hand-drawn as well. If the plot exists as a "strip chart," it may be folded or rolled.

EXAMPLES: Plots of data provided on data sheets or plots of analog data. All strip charts, helicorder rolls, seismic reflection record sections, and borehole geophysical logs are included.

CRITERIA NORMALLY USED FOR RETRIEVAL PURPOSES: Subject, origin, and date.

OTHER CHARACTERISTICS: Form, scale, color, size, dimensions, legibility, page-count, and Quality Assurance (QA) status.

UNITIZATION GUIDELINES: If a data plot exists as a stand-alone item (apart from a document or package) it should be indexed as an independent unit. If it exists within a collection of data plots on the same subject (apart from a document or package) and is best interpreted within the context of that collection, the collection should be indexed as a single unit. If a data plot occurs within a document and would be both useful and fully understood outside the context of that document, it should be independently indexed. If a data plot is part of a package, it should be independently indexed only if it is recorded on machine-dependent media or cannot be converted to an electronic image without the loss of information content.

RELATED DOCUMENTS: If a collection of data plots is separated into multiple units for indexing purposes, the units should be related to each other through this LSS bibliographic header field.

SUBMISSION REQUIREMENTS:

ELECTRONIC IMAGE: Every printed data plot should be scanned, if possible, to produce a viewable image. If a data plot is not imageable, an effort should be made to convert it to scannable paper for LSS input, provided that the conversion will not result in a loss of information content.

ASCII TEXT: Embedded text that may exist on a data plot will be automatically converted to ASCII form for text search, insofar as possible, without manual intervention.

BIBLIOGRAPHIC HEADER: An independent header will be submitted if the above unitization guidelines require it.

CATEGORY: Computer Program

DEFINITION: A sequence of coded instructions/algorithms/procedures ("source code") to be loaded and executed on a computer.

PHYSICAL ATTRIBUTES: A computer program may be recorded on scannable media, such as paper (ordinarily printed by a computer), or on non-imageable media, such as magnetic tape.

EXAMPLES: Computer code, simulation models, database programs, and spreadsheet programs.

CRITERIA NORMALLY USED FOR RETRIEVAL PURPOSES: Subject, origin, and date.

OTHER CHARACTERISTICS: Hardware/software dependencies if recorded on machine-dependent media; page-count and legibility otherwise. Release number and Quality Assurance (QA) status in either instance.

UNITIZATION GUIDELINE: If a computer program exists as a stand-alone item (apart from a document or package) it should be indexed as an independent unit. If it exists within a collection of computer programs on the same subject (apart from a document or package) and is best interpreted within the context of that collection, the collection should be indexed as a single unit. If a computer program occurs within a document and would be both useful and fully understood outside the context of that document, it should be independently indexed. If a commuter program is part of a package, it should be independently indexed only if it is recorded on machine-dependent media or cannot be converted to an electronic image without loss of information content.

RELATED DOCUMENTS: If a collection of computer programs is separated into multiple units for indexing purposes, the units should be related to each other through this LSS bibliographic header field.

SUBMISSION REQUIREMENTS:

ELECTRONIC IMAGE: Every printed computer program should be scanned, if possible, to produce a viewable image. If a computer program is not imageable, an effort should be made to convert it to scannable paper for LSS input, provided that the conversion will not result in a loss of information content.

ASCII TEXT: While a progr.m is textual in nature, there is no point in converting its text to ASCII for keyword search. However, automatic conversion, without manual intervention, should proceed with these materials as well as others when they are contained in a package.

BIBLIOGRAPHIC HEADER: An independent header will be submitted if the above unitization guidelines require it.

CATEGORY: Handwritten Notes

DEFINITION: Any set of handwritten notes made for any purpose. Mixed handwritten/printed materials are included, as are printed forms that are filled out in handwriting. Marginalia on predominantly printed material is excluded, because the material is considered text-searchable. Often, handwritten notes will be compiled into notebooks.

PHYSICAL ATTRIBUTES: A set of handwritten notes is usually recorded on scannable media, such as paper, but it is conceivable that handwritten notes may be recorded on non-imageable media, such as magnetic tape.

EXAMPLES: Field, laboratory, and scientific notebooks, logs, memoranda and review comments.

CRITERIA NORMALLY USED FOR RETRIEVAL PURPOSES: Subject, origin, and date.

OTHER CHARACTERISTICS: Legibility, page-count, and Quality Assurance (QA) status.

UNITIZATION GUIDELINES: If a set of handwritten notes exists as a stand-alone item (apart from a document or package) it should be indexed as an independent unit. If the set exists within a collection of handwritten notes on the same subject (apart from a document or package) and is best interpreted within the context of that collection, the collection should be indexed as a single unit. If a set of handwritten notes occurs within a document and would be both useful and fully understood outside the context of that document, it should be independently indexed. If a set of handwritten notes is part of a package, it should be independently indexed only if it is recorded on machine-dependent media or cannot be converted to an electronic image without loss of information content.

RELATED DOCUMENTS: If a set or larger collection of handwritten notes is separated into multiple units for indexing purposes, the units should be related to each other through this LSS bibliographic header field.

SUBMISSION REQUIREMENTS:

ELECTRONIC IMAGE: Every printed set of handwritten notes should be scanned, if possible, to produce viewable images. If the notes are not imageable, an effort should be made to convert them to scannable paper for LSS input, provided that the conversion will not result in a loss of information content.

ASCII TEXT: Printed text that may accompany handwritten notes should be automatically converted to ASCII for text search, without manual intervention. No effort should be made at the capture station to transcribe (interpret, clarify, or type) handwritten notes for the purpose of replacing those notes as submitted material. Any transcription that is considered essential should be treated as a new (text-searchable) document, related to the notes, and clearly identified for what it is.

BIBLIOGRAPHIC HEADER: An independent header will be submitted if the above unitization guidelines require it.

APPENDIX C

MEDIA ATTRIBUTES

MEDIA ATTRIBUTES

MEDIA: Magnetic Tape

DEFINITION: A flexible strip coated with a magnetic substance, used as a storage medium for information recorded by a computer. It could be wound on a reel or enclosed in a cartridge or cassette. It may contain programs, data, models, graphics (maps, designs, etc.), or any other information that is retained in electronic form.

ATTRIBUTES REQUIRING DESCRIPTION: Specific hardware/software dependencies including computer type/model, operating system, application program,, width, number of tracks, recording density, character code, record length, block size, file description, physical file characteristics, recording mode, record layout (field names, sizes, starting positions, numeric-data forms), and series number.

CONVERSION CONSIDERATIONS: There should be no conversions to paper for the purpose of creating viewable LSS images.

MEDIA: Magnetic Disk

DEFINITION: A round flat plate coated with a magnetic substance used as a storage medium for information recorded by a computer. It may be in the form of a hard disk or a flexible ("floppy") disk. Hard disks are usually housed in a container of some kind and are sometimes combined in a spindled pack. A flexible disk may be enclosed in a rigid or soft jacket, which is usually square in shape. A magnetic disk may contain programs, data, models, graphics (maps, design drawings, etc.), or any other information that is retained in electronic form.

ATTRIBUTES REQUIRING DESCRIPTION: Specific hardware/software dependencies including computer type/model, operating system, application program, character code, record length, block size, file description, physical file characteristics, record layout (field names, sizes, starting positions, numeric data forms), and series number.

CONVERSION CONSIDERATIONS: There should be no conversions to paper for the purpose of creating viewable LSS images.

MEDIA: Solid-State Device

DEFINITION: An assembly of electronic circuitry capable of functioning as a permanent storage medium for information recorded by a computer, which is removable from the computer. It may exist in many forms (a chip, a board, a module, etc.) and may be known by various names, which continue to evolve as computer technology advances (e.g., "ROM"—meaning read-only-memory). Although this medium is not currently employed as a removable storage device for documentary material, it is possible that it will be used in future years. A solid-state device may contain programs, data, models, graphics (maps, designs, etc.), or any other information that is retained in electronic form.

ATTRIBUTES REQUIRING DESCRIPTION: Specific hardware/software dependencies including computer type/model, operating system, application program, character code, record length, block size, file description, physical file characteristics, record layout (field names, sizes, starting positions, numeric-data forms), and series number.

CONVERSION CONSIDERATIONS: There should be no conversions to paper for the purpose of creating viewable LSS images.

MEDIA: Optical Disk

DEFINITION: A round flat plate having a plastic coating, which is used as a storage medium for information recorded by a computer. It is read by using a laser device. It may be housed in a spindled pack or within an inseparable container. (Its use as a videodisc, rather than a computer disk, is separately described. An optical disk may contain programs, data, models, graphics (maps, designs, etc.), or any other information that is retained in electronic form.

ATTRIBUTES REQUIRING DESCRIPTION: Specific hardware/software dependencies including computer type/model, operating system, application program, character code, record length, block size file description, physical file characteristics, record layout (field names, sizes, starting positions, numeric-data forms), and series number.

CONVERSION CONSIDERATIONS: There should be no conversions to paper for the purpose of creating viewable LSS images.

MEDIA: Videocassette/Videodisc

DEFINITION: An encased tape or round flat disk, having a plastic coating, upon which images have been recorded for the purpose of playback on a television set. It may contain motion or still pictures. A videocassette is a case containing a video tape, which relates to a television set by means of a videocassette recorder/player. At least three formats have been applied to videocassettes: VHS, Beta, and U-Matic—the former being the most common. A videodisc, in contrast, is similar in appearance and use to a phonograph record and relates to a television set by means of a videodisc player. Optical disk technology is employed in its recording and playback. (The use of optical disks as computer storage media is separately described.) A videocassette or disk may depict terrain, structures, weather conditions, physical samples, etc.

ATTRIBUTES REQUIRING DESCRIPTION: Specific hardware dependency (videocassette or videodisc player) and, in the case of a videocassette, its format (VHS, Beta, U-Matic).

CONVERSION CONSIDERATIONS: There should be no conversions to paper for the purpose of creating viewable LSS images.

MEDIA: Audio Recording

DEFINITION: A physical medium upon which sounds have been recorded for the purpose of playback on a suitable device. There are five common types of audio recording media: (1) magnetic tape, which may be stored on a reel or cassette, (2) phonograph record, (3) compact disk, (4) digital audio tape (a relatively new product), (5) voice-mail record (another emerging technology), and (6) dictation media, which may use various types of dictation devices. (Tapes and disks employed by computers are separately described.) An audio recording may document untranscribed speeches, readings, conversations, notes, sounds or nature, etc.

ATTRIBUTES REQUIRING DESCRIPTION: Specific hardware dependency (audiocassette player, phonograph, compact disk, digital audio tape, dictation device).

CONVERSION CONSIDERATIONS: If the recording is verbal, an effort should be made, prior to submission, to transcribe the recording—typing it on paper suitable for conversion to viewable LSS images, so that the information it contains will be more readily available to requestors in both image and textual form. This should be done, however, only if no loss of information content would occur. If it is done, the audio recording itself would not be submitted to the LSS.

MEDIA: Motion Picture Film

DEFINITION: A strip of photographic film on which images have been recorded for the purpose of playback on a movie projector. It is composed of multiple frames of film connected in a sequence on a strip, which is usually would on a reel and punched for a sprocket drive. A motion picture film may depict terrain features, physical samples, structures, weather conditions, etc.

ATTRIBUTES REQUIRING DESCRIPTION: Specific hardware dependency (8mm, 16mm, etc.).

CONVERSION CONSIDERATIONS: There should be no conversions to paper for the purpose of creating viewable LSS images.

MEDIA: Microform

DEFINITION: Microimage of a document that has been optically reduced onto film to a degree that it cannot be easily read without magnification. There are three common types of microforms: (1) roll microfilm (15 mm and 35 mm) which may be stored on a reel, cartridge or cassette, (2) microfiche, which is a sheet of film containing multiple images in a grid pattern, and (3) aperture card, which is a paper card (usually an electric accounting machine "punchcard") into which a rectangular aperture has been cut to mount film (usually one 35 mm frame). Magnification devices have been specifically designed to read each of these forms. A microform may depict maps, photographs, designs, figures, plots, notes, etc.

ATTRIBUTES REQUIRING DESCRIPTION: Specific hardware dependency (reel, cartridge, cassette, microfiche, aperture-card reader, etc.) and size.

CONVERSION CONSIDERATIONS: An effort should be made, prior to submission, to submit a document to the LSS in its original paper form, suitable for conversion to a viewable LSS image, so that the information it contains will be more readily available to requestors. This should be done only if no loss of information content would occur. If a microform is converted to paper for LSS submission, the microform itself would not be submitted to the LSS.

MEDIA: Photographic Slide

DEFINITION: A photographic positive-image transparency (usually 35mm) intended for projection on a screen, its image having been optically reduced onto film to a degree that it cannot be easily read without magnification. It is ordinarily mounted on a conventional cardboard or plastic holder that is designed to be inserted into a projection device but will also permit viewing by means of a slide or microform reader. A photographic slide may depict a scene, design, physical sample, structure, etc.

ATTRIBUTES REQUIRING DESCRIPTION: Size.

CONVERSION CONSIDERATIONS: If feasible and there would be no loss of information content, images on slides should be converted, prior to submission, to paper form, suitable for conversion to viewable LSS images, so that the information they contain will be more readily available to requestors. If a slide is converted to paper for LSS submission, the slide itself would not be submitted to the LSS.

MEDIA: Photographic Negative

DEFINITION: A negative photographic image on transparent film used for printing positive pictures, which is reduced to a degree that its reversed light/dark image cannot be easily read without magnification. It ordinarily appears on a film strip containing several images, but that may not always be the case. It may usually be viewed in distorted light/dark form through projection or with the help of a microform reader. A photographic negative may depict the same things that a printed photograph depicts—terrain features, physical samples, structures, etc.

ATTRIBUTES REQUIRING DESCRIPTION: Size.

CONVERSION CONSIDERATIONS: An effort should be made, prior to submission, to print from the negative a positive rendition on paper that will be suitable for conversion to a viewable LSS image, so that the information it contains will be more readily available to requestors. This should be done only if no loss of information content would occur. If a photographic negative is converted to paper for LSS submission the negative itself would not be submitted to the LSS.

MEDIA: Develocorder Film

DEFINITION: A photographic film strip similar to a motion picture film reel except that it is not punched for a sprocket drive. Made of plastic, it is 16 mm in width and requires a develocorder film reader for viewing. A develocorder film is most commonly used to record earthquake seismographic data.

ATTRIBUTES REQUIRING DESCRIPTION: None.

CONVERSION CONSIDERATIONS: There should be no conversions to paper for the purpose of creating viewable LSS images.

MEDIA: Radiograph

DEFINITION: A photographic image on a radiation-sensitive film, commonly made by x-ray, gamma-ray, or alpha-ray. It is similar in appearance to a photographic negative, but is usually larger (e.g., 8" by 10"). It requires a light table and sometimes a microscope for viewing. A radiograph is most commonly used to record the distribution of radioactive species in both natural and engineered materials.

ATTRIBUTES REQUIRING DESCRIPTION: Radiation type, size.

CONVERSION CONSIDERATIONS: There should be no conversions to paper for the purpose of creating viewable LSS images.

APPENDIX D

INSTRUCTIONS FOR COMPLETING A TIP TABLE OF CONTENTS

INSTRUCTIONS FOR COMPLETING A TIP TABLE OF CONTENTS

D.1 GENERAL INSTRUCTIONS

General instructions for completing a TIP table of contents follow.

1. The TIP table of contents must be typewritten, so that it may become text searchable within the LSS. It will always appear as the first page of a package.

2. Authors should be identified by the name printed on the title page of the finished product to which they contributed or on the signature line of approval/review documents. A full identification of the person (providing a first name, middle initial, and surname) will assist the retrieval of a particular author's products and approvals/reviews when a full-text search is conducted. If an author belongs to an organization other than the one responsible for producing the package, the author's affiliation should be appended to the author's name.

3. When included in an entry, dates will always be completely spelled out (e.g., April 15, 1991) to avoid possible ambiguity. When the specific date is unavailable, the month and year will be sufficient.

4. The description of each item (product, review, raw-data item, etc.) will be followed by the number of pages it contains—including all covering pages, prefaces, etc. When raw-data items are grouped (e.g., drilling reports), a combined page-count will be provided. At the bottom of the table of contents, the page-counts of all of the items in the package will be totaled, including slip sheets (which are usually written on one page).

5. Every item will be located in the package in the order in which it is listed on the table of contents, which will normally be in the order of date. There will be no need to number the pages in a package and to place beginning page numbers beside each listed item, as would be done in a normal table of contents, because the use of hypertext indexing by the LSS will permit a requestor to move automatically from the package table of contents to the initial page (image) of a selected item using a pointing device or a single keystroke.

6. There is provision in the following instructions for packages that contain no finished products (which are expected to occur frequently) and for packages that contain more than one finished product (which are expected to occur infrequently).

7. The heading, "PACKAGE TABLE Of CONTENTS", should appear at the top of every listing.

D.2 SPECIFIC INSTRUCTIONS

The required sections of the TIP table of contents should be completed as follows.

PACKAGE NUMBER: Provide a unique identifying number for the package which will distinguish it from other LSS packages. The numbering system will be centrally controlled and

will use the same number that is designated for the bibliographic header. Obtain this "Package ID" number from LSS Process Control. It will be used primarily to relate individually indexed materials within a package to their parent package and to each other.

RESPONSIBLE ORGANIZATION: Provide the full name of the originating organization in the format listed in the LSS Organization Name Authority List, followed by the acronym or abbreviation listed there (which will assist text search).

TITLE: Provide a succinct description of the subject investigation which will distinguish it from others. (The term "package" should not appear in the title.)

FINISHED PRODUCT: If the investigation has resulted in a finished product (technical report, design, map, etc.), provide its full title in single-paragraph format, including, in order: subject, report number (in parentheses), date, and the names of all contributing authors who are listed on the product's own title page, appending organizational affiliation (using the controlled acronym/abbreviation) if it differs from the organization responsible for the package. If a package contains more than one finished product, provide the full title of each, in the same way, using a separate paragraph for each. If a package does not contain a finished product, type "None".

APPROVAL AND REVIEW DOCUMENTS:

LETTERS OF APPROVAL: List all approvals that the investigation has received, providing in each instance the full name (using middle initial) of the person in authority, organizational affiliation using the acronyms/abbreviations found in the LSS Organization Name Authority List, and the date of approval. If there were no formal approvals for the investigation, type "None".

CIRCULATED DRAFTS: If the investigation produced one or more "nonfinal" reports or other finished products which were circulated for supervisory concurrence and failed to receive it, these drafts should be included in the package. This must be done whether or not a finished product is subsequently published. Provide the full title of each such draft, in the format specified above for finished products. If a package contains more than one finished product, do the same for any circulated drafts that preceded them, using a separate paragraph for each. If the investigation did not produce a circulated draft, type "None".

DOCUMENT REVIEWS: If the investigation resulted in a product that was reviewed internally within the responsible organization, provide the full names of the reviewers (with middle initial) and the dates of review. If a package contains more than one finished product, identify the product reviewed by adding its subject within parentheses. If there were no document reviews, type "None".

RAW DATA:

List each item of raw data included in the package (notes, logs, test results, measurements, calculations, graphs, surveys, sketches, etc.), either individually or by group if, in the judgment of the principal investigator, the item belongs to a set of items having the same subject, author organization, and document type (e.g., "Pumping test data" in the example).

Use consistent, unabbreviated terminology in describing these materials to facilitate subsequent retrieval.

Refer to the LSS Document Type List in the LSS Cataloging Manual (taken from Appendix B of this report), which describes each type. Use those terms within the descriptions of the items when it would help to identify them. Describe the items in accordance with the essential characteristics of their respective document type, but avoid attributes (such as legibility, color and scale) that will be immediately evident to the requestor upon viewing the items on the LSS screen.

If a listed item was produced by someone unaffiliated with the responsible organization (e.g., "Drilling reports" in the example), the author (if identified) and the producing organization must be named.

Add dates of origination to the listed items in parentheses.

When a slip sheet (header input form) is inserted within a package in place of non-imageable material stored separately (e.g., material on a magnetic tape), add the parenthetical term "(slip sheet)" to the item's description. A proper description for machine-dependent items should include both the item's content (e.g., "injection test data") and the medium on which it is recorded e.g., "magnetic tape"). Detailed specifications of an item's machine dependency (or more detailed description in the case of other non-imageable items), as well as its origin, creation date, storage location, and retrieval procedure will be viewable on the slip sheet, not on the table of contents itself.

TOTAL PAGES: Add the page counts listed individually for each finished product, letter of approval, circulated draft, document review, and raw data item (or group, as listed) and provide the total package page count.

SIGNATURE AREA: After the statement: "I attest that this package is accurate and complete", provide the typed name and signature of the principal investigator or project manager, followed by the typed date of package completion.

QUALITY ASSURANCE STATUS: Enter either "YES" or "NO" in the space provided, indicating whether or not the entire package has been produced by the responsible organization under an approved quality assurance program, in accordance with 10CFR Part 60 Subpart G.

FORMAT FOR A TIP TABLE OF CONTENTS

PACKAGE TABLE OF CONTENTS

PACKAGE NUMBER:

RESPONSIBLE ORGANIZATION:

TITLE:

PAGES

FINISHED PRODUCT:

APPROVAL AND REVIEW DOCUMENTS:

LETTERS OF APPROVAL:

CIRCULATED DRAFTS:

DOCUMENT REVIEWS:

RAW DATA:

TOTAL PAGES ____

I attest that this package is accurate and complete.

_____
(typed name)

QUALITY ASSURANCE STATUS ___
(Either "Yes" or "No")

_____
Package Completion Date

COMPLETED EXAMPLE OF A TIP TABLE OF CONTENTS

PACKAGE TABLE OF CONTENTS

PACKAGE NUMBER: DA12345678

RESPONSIBLE ORGANIZATION: United States Department of the Interior Geological Survey (USGS)

TITLE: Geohydrologic and Drill-Hole Data for Test Well USW H-5, Yucca Mountain, Nye County, Nevada

I attest that this package is accurate and complete.

William Z. Tell (signed)
William Z. Tell

QUALITY ASSURANCE STATUS: Yes
(Either "Yes" or "No")

November 1, 1985
Package Completion Date

guideline should apply (all authors of the package should be mentioned), unless it is exceptionally identified as the creation of another person on the table of contents.

AUTHOR ORGANIZATION: In the case of a package-unit, this field will take the name of the responsible organization from the table of contents. It will also include any other names noted, on an exceptional basis, next to raw-data items that have been produced by other organizations (including subcontractors) and next to authors of finished products. In the case of individually-indexed packaged items, the name of the responsible organization will be entered, unless it is a packaged item identified as originating elsewhere. Names on the table of contents will in all instances conform with the controlled Organization Name Authority List. As stipulated by the draft LSS Cataloging Manual, the correlation between each author and that author's organizational affiliation will be maintained.

ADDRESSEE: This field is not applicable to packages but will be applicable to individually-indexed correspondence that a TIP may contain.

ADDRESSEE ORGANIZATION: This field is not applicable to packages but will be applicable to individually-indexed correspondence that a TIP may contain, using the Organization Name Authority List.

DOCUMENT DATE: In the case of package-units, the date of its completion, located at the bottom of the table of contents, should be used. In the case of individually-indexed packaged items, the particular date provided on the table of contents should be entered.

DOCUMENT/REPORT NUMBER: This number, or numbers, assigned to a unit by the submitting organization for identification or control purposes, should be applied to package-units and to other items in accordance with the organization's internal procedures. Technical reports and approval/review documents will be numbered in accordance with the format rules of the Document Number Authority List, if they apply.

DOCUMENT CONDITION: Packages may have missing pages, illegible pages, or pages bearing marginalia, in which case these characteristics should be entered into this field in accordance with the controlled vocabulary in the Document Condition List, as stipulated by the draft LSS Cataloging Manual. In the case of non-imageable items, their non-imageable status should be indicated.

EDITION/VERSION: This field would have applicability to circulated drafts, some finished products, and some non-imageable materials computer programs in particular. The draft LSS Cataloging Manual excludes this field and would apparently use the Pointers field for this purpose.

EVENT DATE, CODE: If TIPs or individually-indexed materials have a particular dated event (audit, hearing, inspection, or meeting) as their primary topic, this date would be entered using the same controlled vocabulary that will be used for all materials. Otherwise, the Document Date, showing a TIP's completion date will be sufficient, requiring nothing to be entered here.

PROTECTED STATUS: Privileges or exceptions claimed for a TIP or for individually-indexed materials should be explained, using the controlled vocabulary. (The draft LSS Cataloging Manual excludes this field.)

RELATED DOCUMENTS: The original approved description for this field (HWG, 1990) calls for a package and the cataloging units it contains to be identified here, by the submitter, as a "whole/part" type

of relationship, to be translated at LSS input time into standardized form for the Pointers field. The draft LSS Cataloging Manual, which excludes the Related Documents field, suggests that the prefix "PA" from the Pointer Code List be used in the two-way Pointers field to communicate that a unit that "has parts. . . or is part of . . .". This could logically be used for packages and for package parts. However, a Package ID field was adopted by the LSSARP in 1991 (HWG, 1991) making such use redundant for TIPs. The CNWRA recommends that, for TIPs, these fields (Related Documents and Pointers) be used instead to link packages that concern the same prolonged investigation, using a special prefix, as suggested in Section 5, Alternative 3 of this report.

SPECIAL CLASS: This field, intended to further classify units, will employ a controlled vocabulary from a Special Class List according to the draft LSS Cataloging Manual. While that Manual requires this field to be used for package headers, its Special Class List does not include the term "package" as an option. Given the fact that a package-unit will be identified as such in the Document Type field, there seems to be no reason to include a redundancy here. Non-imageable items should be identified in this field as having a "Header Only," which is an option on the List. The listed option, "Header & Image only," will not ordinarily be used for packaged material of that type, since imageable raw data will not be individually indexed according to the CNWRA's recommendation; but a finished product, such as a map or a design drawing, would be so identified.

DOCUMENT TYPE: This field will be governed by a controlled Document Type List, which was mentioned but not included in the draft LSS Cataloging Manual. The intent here is to describe the kind of unit that the header represents, not in terms of information content, but in terms of form (e.g., correspondence, reports, etc.). A "package" should certainly be an option on this List, when it is created. In the case of individually-indexed non-text-searchable materials, the classification that has been provided in Section 2 of this report should be developed into a set of additions to this List. More than one entry should be made if more than one document type applies to a particular unit.

SPONSORING ORGANIZATION: This is the name of the agency or agencies responsible for funding or otherwise sponsoring the investigative work undertaken by the Submitter Center. It, too, will use the controlled Organization Name Authority List.

PUBLICATION DATA: This field, which is intended to capture "bibliographic information that is not covered in other fields but is important in identifying or citing the unit" has no perceived applicability to package-units. However, as suggested by examples in the draft LSS Cataloging Manual, it is a convenient location to record publication data on indexed maps and computer software programs, which would include characteristics that have been defined for those categories of material in Appendix B. Design drawings and commercially available photographs are other items whose publication data could usefully be placed in this field, when those items are indexed.

DESCRIPTORS: These subject terms, selected from the controlled vocabulary in the LSS Thesaurus to identify the information content of a unit, will be particularly important for the retrieval of TIPs. They should be gleaned not only from the title of a package, but also from the package table of contents, the package abstract, the title of any finished product included in the package, and the abstract, table of contents, index, introduction and any other summary included in the finished product. In the case of non-imageable packaged items, the subset of the above descriptors that are pertinent to their specific content should be entered.

IDENTIFIERS: These subject terms, which are not yet contained in the LSS Thesaurus but would serve to further identify the information content of a unit, will also be important for the retrieval of TIPs and individually-indexed materials. They should be gleaned from the same sources from which descriptors have been gleaned.

COMMENTS: This field is intended to capture "information not covered in (other fields) which . . . will be necessary to identify or retrieve the unit." Examples provided in the draft LSS Cataloging Manual are specifics about foreign language content, about missing pages/attachments (for which only an indicator was used in the Document Condition field), or about an omitted image when the page in question was larger than 17 x 22 inches, rendering it "oversize" and thus presumably unimageable. Conditions such as these would apply to packaged materials as well as to other LSS documents. However, the CNWRA has recommended (CNWRA, 1991b) that all items, regardless of size (or color), should be scanned, if possible, to produce viewable LSS images, which would negate the Manual's last example. The CNWRA suggests that, in the case of items resident on non-imageable media, this field be used to supplement the proposed controlled-entry Media Type field, as described below.

ABSTRACT/SUMMARY: This narrative description of a unit's content should be included within headers for package-units themselves, for finished products within a package (but not approval/review documents), and for non-imageable items belonging to a package. Headers for non-text-searchable items having no package association should also include abstracts. An abstract is most successfully written by the producer of the material—by the principal investigator or another knowledgeable technician, rather than by records management specialists who will be less acquainted with its technical content. An abstract for a package-unit should be brief and focus entirely on the subject and results of the investigation. It should make no attempt to summarize the table of contents. An abstract for individually-indexed materials should mention characteristics included in Appendix B of this report (which would presumably be incorporated within the framework of the LSS Cataloging Manual), when those attributes are not covered by other fields (e.g., Document Condition, Submitter, Page Count, and Quality Assurance Status). For instance, the geographic area, scale, size, and color of a map should be described in this field.

LSS SYSTEM ACCESSION NUMBER: This unique identification code, assigned to each cataloging unit by the LSS capture station, has no exceptional applicability to packages or to non-text-searchable material in general.

NUMBER OF IMAGES: This number will be based upon the total in the Submitter Page Count field. If materials are scanned as single LSS images regardless of size, as the CNWRA recommends, this field becomes redundant with Submitter Page Count, because an "oversize" image will not be divided into parts. If a printed rendition is ordered by an LSS requestor, it, too, would be delivered whole.

POINTERS: This field is intended to provide references to Related Documents after the entries in that field have been standardized. The CNWRA's recommendations concerning that field apply.

PACKAGE ID: This field was adopted specifically for the benefit of packages, with the aim of relating all parts of a given package to each other. A controlled two-character prefix to the assigned number (obtained from LSS Process Control and typed at the top of a table of contents) is suggested by the draft LSS Cataloging Manual, for the purpose of distinguishing "data packages" (which this report is calling TIPs) from other types of packaged collections (e.g., training packages, audit packages). The parts of a package that are independently indexed as units (finished products, approval/review documents, and non-

imageable items, according to the CNWRA recommendation) should all bear the number of their assigned package. Errata sheets relating to package-units will be linked to those units through this field.

COPYRIGHT: This field has no exceptional applicability to packaged materials.

In addition to the above-described LSSARP-approved bibliographic header fields, the CNWRA recommends the inclusion of three additional fields to assist the retrieval of non-text-searchable material:

MEDIA TYPE: This field, already included in the draft LSS Cataloging Manual, would designate the media on which non-imageable materials are stored. It would not apply to headers for TIPs, their finished products and their approval/review documents. According to the CNWRA's recommendation, non-imageable items will be individually indexed in all instances (in or out of packages) and require this characterization, since they will be unavailable for viewing on LSS screens. These items are currently numerous and are expected to become increasingly so. The draft LSS Cataloging Manual did not provide a controlled Storage Media List, which it proposed. The machine-dependent media types described in Appendix C of this report should constitute options for that List, which would permit a search by general media type. However, to enable a requestor to understand what specific equipment will be necessary to examine a given item, a full media description will be necessary, amplifying the listed options to detail the hardware/software dependencies contained in Appendix C. In view of the fact that a controlled field like this cannot contain such detail, it is suggested that the Comments field be used for this purpose. The detail would be obtained from information provided on the corresponding slip sheets within a package, which were created as input forms. Miscellaneous non-imageable items that have no machine dependencies could be described as "paper," "mylar," "vellum," "acetate," or "other" in the Media Type Field. In the case of imageable items, the field could be left blank or the term "image" could be entered by default.

STORAGE LOCATION: This field will tell where non-imageable materials can be found and how they can be examined. It should contain an acronym/abbreviation from the Organization Name Authority List, which would link to an updatable table providing a current mailing address, phone number, FAX number, cognizant administrative authority, and procedure for retrieval. Like the Media Type field, this field will have no applicability to TIP headers themselves but will be essential for non-imageable items, which must be kept apart from the LSS computer system at dispersed storage locations. The draft LSS Cataloging Manual suggests that the Media Type field incorporate a code which could be translated through a table into a procedure for access to the particular item. The CNWRA believes that a separate, searchable field (Storage Location) would be preferable. The field could either be left blank in the case of imageable items, or the term "image" could be entered by default.

QUALITY ASSURANCE STATUS: This would be a simple indicator, saying either "yes" or "no", thereby indicating whether or not the unit was produced by the submitting organization under an approved quality assurance program, in accordance with 10CFR Part 60 Subpart G. The field would have applicability to all LSS documents, including packaged material. In the case of package-units, the entry may be taken from the typed notation on the bottom left corner of the TIP table of contents, which identifies its "Quality Assurance Status."