

Enclosure 8

From: Dan Graser
To: Chris Berlien, David Hunt, Dennis Bechtel, Elaine Ezra, Englebrecht
vonTiesenhausen, Glen Foster, Harry Leake, Harvey Spiro(...)
Date: Fri, Feb 11, 2000 6:35 PM
Subject: Revised Functional Requirements

TWG Members:

Attached is a file containing all the revisions suggested to date by the various commenters (with exception of one of Lew Robertson's comments "does the version of SQL matter?". Labat is still working on the document volume tables back in the TAB sections. That should be ready next Monday or Tuesday. If you have comments back to me, and I get those tables updated, we will be in a position to provide the draft to the ARP and discuss with them the next steps needed.

Please review this draft as the issue will be discussed at the next ARP meeting.

CC: Chris Hoxie, G Paul Bollwerk, internet:rusmith@fdic.gov, Jack Whetstine, John Sullivan, Matthew Schmit

Mail Envelope Properties (38A49CDA.4A1 : 1 : 18974)

Subject: Revised Functional Requirements
Creation Date: Fri, Feb 11, 2000 6:35 PM
From: Dan Graser

Created By: DJG2.TWF2_PO.TWFN_DO

Recipients	Action	Date & Time
co.clark.nv.us dax (Dennis Bechtel) evt (Englebrecht vonTiesenhausen)	Transferred	02/11 6:36 PM
cs.unlv.edu taghva (Kazem Taghva)	Transferred	02/11 6:36 PM
fdic.gov rusmith CC (internet:rusmith@fdic.gov)	Transferred	02/11 6:36 PM
gfoster.com gfoster (Glen Foster)	Transferred	02/11 6:36 PM
isri.unlv.edu tom (Tom Nartker)	Transferred	02/11 6:36 PM
labat.com joseph_speicher (Joseph Speicher) tony_neville (Tony Neville)	Transferred	02/11 6:36 PM
Bob Wells BC (internet:Bob.Wells@rw.doe.gov)	Transferred	02/11 6:36 PM
chris berlien (Chris Berlien)	Transferred	02/11 6:36 PM
elaine ezra (Elaine Ezra)	Transferred	02/11 6:36 PM
ymp.gov david_hunt (David Hunt) harry_leake (Harry Leake) john_gandi (John Gandi) lew_robertson (Lew Robertson) sam_hobbs (Sam Hobbs)	Transferred	02/11 6:36 PM

twf1_po.TWFN_DO	Delivered	02/11 6:36 PM
GPB CC (G Paul Bollwerk)	Opened	02/12 11:22 AM
jgw CC (Jack Whetstine)	Opened	02/14 7:30 AM

twf2_po.TWFN_DO	Delivered	02/11 6:35 PM
CLH1 CC (Chris Hoxie)	Opened	02/14 8:31 AM
hjs (Harvey Spiro)	Opened	02/14 12:18 PM
JAS2 CC (John Sullivan)	Opened	02/14 7:32 AM

twf3_po.TWFN_DO	Delivered	02/11 6:36 PM
MRS3 CC (Matthew Schmit)	Opened	02/14 7:01 AM

Post Office	Delivered	Route
co.clark.nv.us		internet
cs.unlv.edu		internet
fdic.gov		internet
gfoster.com		internet
isri.unlv.edu		internet
labat.com		internet
Bob		internet:rw.doe.gov
chris		

internet:terraspectra.com

elaine

internet:terraspectra.com

ymp.gov

internet

twf1_po.TWFN_DO

02/11 6:36 PM

twf2_po.TWFN_DO

02/11 6:35 PM

twf3_po.TWFN_DO

02/11 6:36 PM

Files	Size	Date & Time
rule_requires_revised_fixed	180794	Friday, February 11, 2000 6:28 PM
MESSAGE	1161	Friday, February 11, 2000 6:35 PM

Options

Auto Delete:	No
Expiration Date:	None
Notify Recipients:	No
Priority:	Standard
Reply Requested:	No
Return Notification:	None

Concealed Subject:	No
Security:	Standard

To Be Delivered:	Immediate
-------------------------	-----------

Status Tracking:

Delivered & Opened

**Licensing Support Network Functional Requirements
Revised Draft for TWG Consideration**

February 11, 2000

The Licensing Support Network (LSN) functional requirements are derived from a series of statements that are found in 10 CFR 2, Subpart J (the LSN Rule). **Section references found within this document that are not otherwise identified are references to applicable sections of the LSN Rule.** Capabilities addressed in the LSN Rule include:

- (1) Internet-based search and retrieval capability;
- (2) functionality to be provided by the participants' systems independent of capabilities of the LSN;
- (3) adjunct capabilities of the Nuclear Regulatory Commission's (NRC) separate electronic docket;
- (4) adjunct capabilities of the NRC's separate electronic information exchange system; and,
- (5) procedural direction external to the LSN that facilitates its uniform operation.

The functional requirements described in the LSN Rule may be subject to subsequent guidance or clarification insofar as they are only broad directives on what the system is intended to do. Additionally, these statements are strongly influenced by the technology solutions that are subsequently applied to the problems.

1.0 FUNCTIONALITIES STATED IN THE LSN RULE

This section lists the functionalities that are directly identified in the LSN Rule and allocates them in one of the five categories noted above. **Subsequently, this document examines functionalities required for core search and retrieval, the NRC Docket system, and the LSN Administrator (LSNA) Audit System.**

1.1 CORE SEARCH AND RETRIEVAL CAPABILITY

The term "Licensing Support Network" means the combined system that makes documentary material available electronically to parties, potential parties, and interested governmental participants to the proceeding for a license to receive and possess high-level radioactive waste at a geologic repository operations area pursuant to 10 CFR, Part 60, as part of the electronic docket or electronic access to documentary material, beginning in the pre-license application phase. The specific method of providing electronic access to documentary material is not specified in the LSN Rule in order to allow flexibility in accommodating current and future technology advances. [63 FR 71735, SUPPLEMENTARY INFORMATION].

The system must be available in time to allow the Department of Energy (DOE) and NRC to meet **their** obligations to make documents available 30 days after the submission **of the site recommendation** to the President (**all** other participants **must** make documents available 30 days after the selection decision becomes final after review by Congress) [§2.1003].

The NRC must provide electronic access to the LSN at the NRC Web site, [HTTP://www.nrc.gov](http://www.nrc.gov) and/or at the NRC Public Document Room (PDR) beginning in the pre-license application phase.

1.2 FUNCTIONALITIES TO BE PROVIDED BY THE PARTICIPANT SYSTEMS

Each participant must:

- obtain the computer system necessary to comply with the requirements for production and service [§2.1011(b)];
- provide a **bibliographic** header record with each document or other material submitted [§2.1003(a)(1)];
- make available an electronic file of the searchable text of its documents. In this context, a file is an entity of data available to system users that is capable of being manipulated as an entity [§2.1003(a)(1)];
- make **text of documents available, including** a bibliographic header with a statement of where an authenticated image **copy** of the document can be obtained [§2.1003(a)(1)];
- make images for non-textual, graphic-oriented documentary material available in lieu of text [§2.1003(a)(2)];
- make headers (only) available for items not suitable for image or text [§2.1003(a)(3)];
- make headers (only) available for privileged, confidential, safeguards and other types of limited access **documentary material** [§2.1003(a)(4)].

DOE must additionally:

- make **documentary material** available 30 days after the submission **of the site recommendation** to the President (**all** other participants **must** make documents available 30 days after the selection decision becomes final after review by Congress) [§2.1003(a)];
- provide electronic access to the LSN at DOE Headquarters and at all DOE **Local Public Document Rooms (LPDRs)** in the vicinity of the site, including Las Vegas, Reno, Carson City, Nye County and Lincoln County) [§2.1007(a)(1)].

1.3 FUNCTIONALITIES ASSOCIATED WITH THE NRC'S DOCKET IN THE AGENCYWIDE DOCUMENTS ACCESS AND MANAGEMENT SYSTEM (ADAMS)

The NRC, in its management of the official docket, must:

- provide a docket that receives, stores, distributes, and maintains pre-license application docket material beginning 30 days after DOE submission of the site recommendation to the President [§2.1013(a)(2)];
- provide a Protective Order File [§2.1006(b)];
- deliver all answers, orders and decisions per **§2.1013(c)** (e.g., electronically, using secured process) [§2.1010(d)];
- provide the computer system necessary to comply with service requirements [§2.1011(b)];
- maintain the docket (NRC SECY) [§2.1013(a)(1)];
- provide an electronic docket that contains a list of all exhibits, **showing where in the transcript each was marked for identification and where it was received into evidence or rejected** [§2.1013(b)];
- enter **hearing** transcripts into the docket on a daily basis in order to provide next-day availability at the hearing [§2.1013(b)];
- establish a mechanism whereby all filings are able to be submitted/received electronically and to require that a password security code for transmission of these documents; [§2.1013(c)(1)];
- provide a mechanism by which **all Presiding Officer and Commission** issuances and orders are transmitted electronically [§2.1013(c)(5)];
- provide a mechanism by which the Presiding Officer and all counsel of all parties have access to the electronic docket (including Protective Order File) during the hearing [§2.1013(d)];
- identify a means by which the unavailability of the electronic docket for more than 4 hours in any day is communicated to the Presiding Officer so that the day is not counted in the computation of time (NRC SECY) [§2.1017];
- provide a mechanism to receive electronically transmitted depositions (**including questions, cross-questions, and answers**) and enter them into the docket file (NRC SECY) [§2.1019(d)&(e)]; and,
- identify a means by which only a part or parts of a deposition may be offered into evidence (NRC SECY) [§2.1019(f)].

1.4 FUNCTIONALITIES ASSOCIATED WITH THE NRC'S ELECTRONIC INFORMATION EXCHANGE SYSTEM

Each participant must:

- use electronic submissions (see §2.1013(c)) for all their motions practice [§2.1010(d)];
- submit all filings electronically using a password security code for transmission of these documents [§2.1013(c)(1)]; and,
- transmit depositions to SECY in electronic form for entry into the docket [§2.1019(d) and (e)].

1.5 PROCEDURAL OR EXTERNAL SYSTEM CAPABILITIES

Each participant must:

- identify all documents not made available in electronic form per §2.1003 in an electronic notice; and make them electronically available **within five days** after someone requests a document off that list [§2.1004];
- designate an official who is responsible for the administration of **its responsibility to make documentary material** available [§2.1009(a)(1)];
- establish its own procedures to make its own **documentary material** available [§2.1009(a)(2)];
- train its own staff on how to make its **documentary material** available [§2.1009(a)(3)];
- ensure each document has a unique ID [§2.1009(a)(4)];
- have the designated official certify that procedures have been implemented, that documentary material has been made electronically available, and, update these certifications at 12 month intervals [§2.1009(b)];
- follow format standards for providing electronic access [§2.1011(e)(2)(i)];
- follow procedures and standards for motions practice [§2.1011(e)(2)(ii)];
- demonstrate substantial and timely compliance in order to be granted party status [§2.1012(b)(1)];
- ensure that any document intended to be used as an exhibit is available before the commencement of that portion of the hearing where it will be offered [§2.1013(b)];

- capture the event of a document being placed in a recipient's mailbox (**delivery receipt**) in order to provide proof of the completion of service [§2.1013(c)(3)];
- provide and show a proof of service for each document filed into the docket [§2.1013(c)(4)];
- when being deposed, submit an electronic index of all documents in his/her possession relevant to the subject matter of the deposition to all parties, identifying which were already made available electronically [§2.1019(l)(1)]; and
- when being deposed, bring along paper copies of all such documents and deliver paper copies with the certified deposition [§2.1019(l)(5)].

DOE must additionally submit the license application to the docket in electronic form 2.1013(a)(2).

2.0 FUNCTIONALITY FOR CORE SEARCH AND RETRIEVAL

Anticipating that an Internet-based software solution will connect the individual participants' systems, there are forty eight (48) core functions that provide a framework for providing search and retrieval of evidentiary **documentary material** housed in the systems where the participants have made their evidentiary **documentary material** "web accessible". This section details each core function, with a narrative that explains the objectives it meets. Core Functional Requirements are as follows:

2.1 The system must provide the ability to access and retrieve documentary material from a variety of repository systems.

This core capability of the LSN site provides general data availability. The LSN Rule requires that each participant make diverse types of documentary material available to other participants. The participants must make available an electronic file of the searchable text of relevant documents, **if available**. Each participant system must provide the capability to store components (header, text and/or image) for subsequent retrieval, but the LSN Rule does not prescribe a uniform hardware or software solution. Rather, it anticipates that participants will establish externally accessible collections on hardware and software of their own choosing. Within their own system, each participant system additionally provides a capability to logically link the header, text and image of a document, where available, for subsequent retrieval, whether these document components are stored at the same time or at different times.

The LSN Rule allows for solutions that could provide users with individual access to each of these participant systems, but the problems of learning and working with multiple software packages, multiple databases, multiple sets of search results, etc., make a unified interface essential. Therefore, the LSN must provide the capability to read text and header data from participant collections and build indices for full text searching in a single, common user interface at a single location. To optimize this **approach**, the central search capability should be **designed to** generate an index or similar mechanism compiled from information extracted from each of the participant sites. Once this central index identifies participant files **that match** a user's request, a request is passed back to the participant system to deliver the original version of the files responding to the query. This implies that each participant **system** must be capable of responding directly to the user for a request generated through the portal for the delivery of requested components (header, text and/or image); or, that each participant system must provide a capability to pass back, through the portal, object files (text, image, or header records) for an end-user to be able to load onto their workstation for further manipulation.

2.2 The system must provide the ability to index, access and retrieve the full text of documents.

This core capability of the LSN site provides text retrieval capability. The LSN site must provide users with the capability to **electronically search and retrieve** document text, **if available**. The LSN Rule allows for solutions that would parse a query from the LSN

user interface and pass it to the individual participant systems. **This introduces higher levels of complexity in normalizing the query logic and in aggregating result sets where different relevancy algorithms may be encountered in participant systems. To simplify these problems, the text made available at the participant collection must be capable of being read and indexed by LSN portal software, which acts as the single, uniform text engine for processing queries and returning relevant result sets.**

2.3 The system must provide the ability to import/link-to documents in their native format.

This core capability of the LSN site provides participants and users with the ability to retrieve and further manipulate documents in web-viewable formats while at the same time minimizing reprocessing costs. Each participant **system** is required to make image versions of textual documents accessible. In addition to image versions of text documents, each participant **system** must make available images, in lieu of text, for non-textual, graphic-oriented documentary material and provide a header record for those image files. This non-textual material may be generated by specialized systems for specialized needs; therefore, each participant collection may contain relevant documentary material in other file formats such as motion, video, sound. This may include vector representations as a document, graphics, mapping formats, and photographs.

The LSN site must provide users with the capability to electronically retrieve the digital image associated with each page in a document. Web browser software accommodates most standard image formats, with the added benefit of displaying textual files in their native word processing or authoring software formats. Therefore, the system should provide the ability to utilize text and image files in their native formats in order to maximize the benefits of standard browser capabilities and reduce reprocessing burdens on the participants. **The web browser software may not handle files in non-standard formats, such as Geographic Information Systems (GIS), and the user will need either the application software resident on their machine, or will need to be able to contact the originating party to make arrangements for viewing the document.**

For image files associated with text, it is desirable to identify standards that will be easiest for browsers to deal with. The electronic image of documentary material in the LSN should use compressed TIFF Group 4 or PDF for bitonal images and compressed JPEG format for color and grey scale images. Compression ratios for image files stored at participant systems should be selected such that an image can be printed at its original size without any degradation detectable by the unaided eye (**e.g., JPEG lossy at or above a quality level of 75 - on a 1 to 100 scale - for color images such as color keyed maps**). The electronic image representation of documentary material should be stored at the following minimum resolutions:

Bilevel (bitonal) images	300 dpi (1 bit representation)
Grey-level images	150 dpi (8 bit representation)
Color	150 dpi (24 bit representation)

2.4 The system must provide the ability to index and recognize metadata from a variety of data formats.

This core capability of the LSN site provides participants and users the ability to retrieve and further manipulate structured data found in bibliographic header records. Each participant collection is required to provide a header record with each document or other **documentary** material submitted. Although the LSN Rule allows for individual participants' systems to process queries, this would introduce similar query normalization and result set presentation problems to those noted for text searches. To simplify these problems, the bibliographic headers made available at the participant collection must be capable of being read and indexed by the LSN software that acts as the single, uniform structured data engine for processing queries and returning relevant result sets. To do this, the LSN must provide the ability to store indices **that it has built by reading** the contents of participant collections. This implies that the structured data in bibliographic headers made available at the participant collection must be capable of being read and indexed by LSN software.

2.5 The system must provide the ability to map metadata across multiple repositories.

This core capability of the LSN site provides a means to normalize structured data and allow for a single user interface. The LSN header will, at a minimum, include the fields and formats defined in TAB A. Because of variations that may be encountered in naming conventions and **database management systems (DBMSs)** used by the various participants, the LSN software should incorporate a mechanism such as an embedded data dictionary to make data structures and names transparent to the user at the LSN user interface.

2.6 The system must provide the ability to map both standard and custom metadata properties from multiple repositories.

This core capability of the LSN site provides a means to normalize structured data and allow for a single user interface to be used for unique data management implementations. The LSN must present bibliographic headers to users in a standard format regardless of the collection from which they originate. At the same time, the LSN must organize bibliographic headers by use of structured data that is compatible with existing, legacy relational database management systems (RDBMS) that are already being used (or plan to be used) by participants.

This core capability of the LSN site also provides the flexibility to incorporate custom fields used for packages and other referencing. Federal records are organized into packages of related material. This is especially the case with DOE's QA records packages. Therefore, participant systems must provide the ability for package-identifying information **to be** associated with each individual document. The LSN site and the docket files at the NRC must be capable of accepting packages that will be transferred via NRC's EIE capability into the Docket, even if only one component document of a participant's package is selected for inclusion into the docket. For search and retrieval purposes, these relationships must be presented to users by the LSN in a

simple, straight-forward way that can transfer from one database to another (when entered into the docket). This necessitates a bibliographic treatment.

Some of the participant legacy system implementations may have developed unique relationship structures to accommodate non-standardized processes such as unitization and aggregating into packages. The LSN system must provide for unique treatments that may be part of the participants' legacy collection architectures. Participants who have implemented package treatments must identify each document that is a component of a package. To ensure consistency within the LSN environment, packages made available in the participant collections should be defined based on uniform rules and should include a listing of all documents in the package. Those systems, additionally, must include the capability to link the bibliographic headers for the records package and the bibliographic headers for the components of the package within the legacy environment. Documents made available in the participant collections should have a uniform bibliographic treatment for identifying where they belong in a package of other documents (parent/child). For example: participant legacy systems should include the capability to identify where a component document begins and ends in relation to other component documents in a package, and, the capability to identify where within a package a document would be placed in a paper-reconstruction of the package's context.

2.7 The system must provide a methodology or framework for supporting future and/or custom document types and repositories.

This core capability ensures that the system can be adapted for collections offered by "late arriving" participants. The LSN must be designed to allow the addition and augmentation of documentary and official record fields.

2.8 The system must provide users with the ability to define specified data sources as subclass collections (e.g., identify actions against a specified collection such as the DOE collection but not against all the other collections).

This core capability allows users to focus on certain source collections in order to narrow the focus of their searching. A participant site may have categorized documents (e.g., DOE's LA itself) and this will facilitate locating information when you know the information "is in the License Application" etc.

2.9 The system must support custom data sources that may have independent (and varied) administrative settings and security attributes.

This core capability ensures that centralized search and retrieval can be performed without being disrupted by security settings allocated to the participants' underlying databases.

2.10 The system must maintain a hierarchical directory to the location of content files in repository systems.

This core capability ensures that the centralized search and retrieval capability knows where a file is supposed to be located in one of the participants' underlying databases.

This capability will be key for audit and compliance as well as for troubleshooting and for help desk/hotline support.

2.11 The system must be able to aggregate the content of multiple document types from multiple repository systems.

This core capability ensures that bibliographic, text, and image files can be browsed or searched by type (e.g., search all text -- including title -- for the term "perched"). This capability also implies that, to the LSN search software, the ASCII text file of one participant is comparable to the HTML document of another participant and to the native WordPerfect or Word file of yet another participant.

2.12 The system must provide users with the ability to browse its directory via a web browser.

This core capability allows users to browse collection content as an alternative to performing structured and unstructured searches.

2.13 The system must provide selected users with the ability (using standard windows-based workstations and standard Internet browser technology) to create and maintain personalized, web-based homepages.

The LSN has two classes of users: (1) those associated with the licensing proceedings, and (2), the general public. Each group has different needs and capabilities. This core capability allows selected users (e.g., licensing, legal, technical staff of the parties) to customize their interface to the system. This allows key users who are participating in the licensing process to optimize the system for their usage preferences. Conversely, members of the general public will have access to a standard, non-customizable user interface that can not be customized by a user. The general public LSN interface must provide consistent presentations of user-selectable functions, system messages, menus, on-line help and data presentation for all user interface screens within the major user functions (search, retrieval, navigation, and system administration).

Implementing a web-based solution for the LSN, and therefore a web-based user interface, encompasses a range of capabilities that are introduced by the use of browsers on windows-based user workstations. To effect the web-based presence, the LSN client software must be operational with Netscape 4.0 or higher, and Internet Explorer 2.0 or higher. Additionally, to accommodate visually and otherwise handicapped individuals, the LSN site should be engineered to allow non-graphic, text-based browsers and allow non-GUI based web "spiders" or "robots" to programatically access and process the materials on behalf of the user.¹ If any portion of the client-side application must reside on a user workstation, that portion of the code must be distributable (downloadable) via the Internet. The following are consequences of using this technology:

¹ Meeting this requirement may constrain other web capabilities such as client-side programming (e.g., JAVA) and scripting (Javascript). Additionally, other aspects must be carefully implemented (e.g., making sure all inline images have ALT tags to provide a textual description of the image object).

- Using the browser interface, a participant-affiliated user will have the capability to change the user's authorized password provided that the user has logged onto the system with an authorized password.
- The client software based on Internet browser technologies should be relatively "thin". Because of the nature of thin clients, rapid changes can be made in the general layout of the user interface to accommodate user preferences discovered during system pre-operational testing and during system operation.
- The LSN user interface will employ graphical user interface techniques that are consistent with the acceptable user interfaces available at the time of the LSN design. This provides out-of-the-box user controls to perform display and object manipulation functions, including capability to:
 - move between and among full text, images, and bibliographic entries within the browser environment;
 - display the bibliographic data of a document within the browser environment;
 - show all fields for each record retrieved within the browser environment;
 - allow retrieval and display of pages of text with the associated images within the browser environment;
 - display one or more pages of the text of a document within the browser environment;
 - browse (scroll) up or down through result set of bibliographic records within the browser environment;
 - locate the terms in the document that satisfy a full text query and to move the term indicator from one term to the next or previous term without displaying intermediate text;
 - view selected text documents and to move to and display (highlight) references to the terms searched in the documents in the participant collections;
 - navigate (next/previous) between highlighted search terms in the document selected for viewing from a query results list;
 - use (or associate) a file with a viewer that can present object text files stored in their native formats;
 - perform image enlargement, reduction, scrolling and rotation;
 - perform point-and-click image display capability for call-up and image manipulation;

- display the images of a document, page by page, including full page views of the images of 8 ½ x 11 inch pages up to E size pages;
- display bilevel (bitonal), gray-level, and color images of documentary material. The gray-level representation should allow up to 25 shades of grey;
- view the following combinations: (1)header; (2) image; (3) text; (4) header and text; (5) header and image; and, (6) text and image;
- synchronously display a page image of a document and its associated text as supported by the LSN software;
- scroll up and down full-text displays of documents by the line, screen, page, or by entire document;
- move forward or backward “n” pages for images as supported by the LSN software;
- scroll up and down image displays by page;
- zoom-in, zoom-out when viewing an image; and,
- rotate an image when viewing an image.

The system should present these capabilities to a browser, whether or not the user's browser can perform the process.

2.14 The system must provide key users with the ability to select content to appear on his or her personalized homepage without administrator intervention.

This core capability will allow key users (e.g., licensing, legal, technical staff of the participants) to customize their interface to the system. This allows key users to optimize the desktop interface to reflect their areas of interest, query form preferences, additional desktop tools, etc. Members of the general public will have access to a standard interface that can not be customized. Participant users could, if they so choose, also use the general public user interface.

2.15 The system must provide key users with the ability to organize the presentation of content on their personalized homepage without administrator intervention.

This core capability allows key users (e.g., licensing, legal, technical staff of the participants) to customize their interface to the system. The LSN user interface should allow a variety of interfaces, their complexity/power reflecting the experience and

familiarity of the searcher with the data.² This allows key users to optimize the system for their usage preferences. Members of the general public will **have access to** a standard, non-customizable user interface; the non-customizable interface should present a form-based (fill in the blanks) query interface as its default.

2.16 The system must provide authorized users with the ability to automatically publish updates via e-mail or other push technologies.

This core capability allows system designers to develop innovative, non-labor-intensive processes to make the system updates more timely. It would allow selected users (system administrators, compliance certifying officials, and technical POCs of the participants) to more quickly and uniformly coordinate with the LSNA. The LSN, per 10 CFR 2, Subpart J, must provide a mechanism to post the occurrence of changes to items in participant collections. Procedurally, the participants will be required to notify the LSNA of changes to **documentary material** so that the change may be posted on the LSN site and be communicated to other users. Each participant system will be required to provide the capability to identify and mark a document as superseded by another document, and to identify the successor document using the unique LSN ID number. Additionally, the LSN must provide a mechanism to post listings of all the documents not made available in electronic form per 2.1003 in an electronic notice.

2.17 The system must provide authorized users with the ability to publish or submit content to the portal.

This core capability allows the LSNA staff to use the LSN site to post notifications, lists of recently superseded documents, updates to guidance documents and procedures.

² Ad hoc query tools often fall short of satisfying user reporting requirements, which is why a form-based user interface (and, perhaps a half dozen "canned" queries) is the recommended default interface for non-participant users. Users need to have some understanding of the following factors in order to successfully utilize an ad hoc query tool:

Boolean logic - A computer's perception of "and/or" is fundamentally different from human perception. For the occasional database user, perhaps only one in nine can consistently use Boolean logic, and it often requires a time-consuming thought process.

Knowledge of basic Structured Query Language (SQL) and relational theory - With very few exceptions, query tools are front-ends with SQL. If users do not understand the concept of a table and a foreign key, they cannot be expected to intelligently use tools such as "group by" options on their main screen.

Understanding the ad hoc query tool itself - We have done a fairly good job of training users to push the right buttons but query tools are deserving of extended professional training.

Understanding the structure and nuances of the database being queried - For the occasional user in a large homogeneous collection, failure to filter using the structured data will be deadly.

Understanding iterative querying - Users don't do this naturally. They must be trained to think in those terms.

The world of ad hoc query tools is an order of magnitude more complex than is generally perceived. Not only are there many tools, but many types of tools identified in previous LSS requirements represent functionality approaching an OLAP (online analytical processing) implementation, which is multidimensional as opposed to RDBMS structures which are inherently two-dimensional.

2.18 The system must provide some form of content control or content staging to prevent unauthorized users from submitting content to the portal.

This core capability provides a mechanism to ensure that only licensing-related material is included in the LSN. To prevent inclusion of non-licensing-related material, the LSN must prevent the retrieval of document text or images that are not associated with an LSN header. The LSN must enforce strict header, document identification and numbering disciplines to prevent confusion in the identification and use of documents during the proceeding. This implies that the LSN site must minimally support a system-assigned unique LSN accession number and have DBMS software that supports the key field association with the participant's unique local accession number.

Each relevant document in a participant's collection must have a unique LSN ID, adhering to a uniform numbering system that includes a document source identifier, reflected in its header record. These numbers must be able to be reflected within the LSN information presentation schema. Similarly, no document should be available in a participant's collection of relevant documents that does not have a unique ID which is reflected in the associated header.

2.19 The system must provide mechanisms to easily develop and add custom components to extend the system.

This core capability allows modular, COTS products to be added to its core functionality to address deficiencies identified by the users during testing.

2.20 The system must accommodate third-party component providers by adhering to open system standards.

This core capability allows new software components to be integrated into the system without seriously impacting other hardware or software components. This implies that the LSN system should be comprised of products that are standards compliant.

Image Format Standards are TIFF Version 6.0 or PDF. Current industry standards for compression must be used.³ For documents, Consultative Committee International Telephone and Telegraph (CCITT)⁴ Group 3 and 4 compression must be supported. For pictures (color images), the Joint Photographics Expert Group (JPEG) must be supported at a JPEG lossy quality level at or above 75 on the scale of 1 to 100. For video, Moving Pictures Experts Group (MPEG) MPEG -1 and MPEG-2 must be supported.

Participants systems and the LSN should use **Transmission Control Protocol/ Internet Protocol (TCP/IP)** as communications protocol. TCP/IP is the basic packet protocol used to connect machines globally on the Internet and on most local area networks (LANs). The Transmission Control Protocol (TCP) interacts with the Internet

³ Some compression methods are "Lossy". This means that part of the detail is lost when the compression occurs. Lossy compression methods are unsuitable for line drawings or images which must be reproduced exactly, such as maps.

⁴ CCITT is now named International Telecommunications Union - Technical (ITU-T).

Protocol (IP) to provide an application protocol interface. Software developers write their products to interface with or use TCP/IP. In doing so they ensure that their product will run on any network using TCP/IP and on different types of computers that are connected in TCP/IP networks.

Data Base Software should be ODBC (Open Data Base Connectivity) compliant. ODBC compliant applications can use **products such as, but not limited to, ORACLE, SYBASE, INFORMIX, and SQL Server** databases interchangeably. Database software must be SQL compliant.

Microsoft Windows has become the defacto standard desk top operating system for much of general business and government. Microsoft Windows is so widely used that almost all software developers are writing their applications to work in a Windows environment.

There are two presentation or user interfaces currently in use by Microsoft. Windows 3.1 and Windows NT use a Graphical User Interface (GUI) and Windows 95 uses an Object Orient User Interface (OOUI). OOUI is also used by OS/2, NextStep and the Mac. The important fact here is that there is now an industry standard for the way applications are "launched" or started on a computer and a standard look and feel to "Windows" applications. All the applications that run on the Windows desk top can exchange data and utilize services provided by the operating system, such as printing or LAN access. The operating systems and the applications working with it have developed a consistent "look and feel". The consistency provided by these de facto standards enhances the user's ability to learn new systems quickly.

2.21 The system must include authoring tools for developing custom components.

This core capability allows server-resident application software customization after initial installation, by use of extensions (or other techniques) that do not invalidate core software licenses.

2.22 The system must include a repository for content included in the portal which contains both physical (all content replicated in the system) and/or virtual (only links/metadata stored in the system database) storage capabilities.

This core capability allows the LSN site to quickly identify documents responsive to queries, and, the flexibility/option to minimize storage requirements, yet consistently and quickly deliver associated image files. Response time is optimized when indices to documents are maintained on the search server; conversely, not all files need be stored at the LSN site and indeed may be optimally stored in their original participant repository. Therefore, the system must allow for using whatever referential schema is optimal for the particular file. In LSN operations, for each electronic file of searchable text made available in a participant collection, the bibliographic header must contain a statement of (or link to) where an authenticated image of the document may be obtained. The LSN must be able to operate with (or initiate) the schema and its operating software. Although participant collections need to make headers (only) available for items not suitable for image or text and for material that is privileged, confidential, safeguards or has otherwise limited access, this capability is crucial to

meeting the requirement that for each item of documentary material not suitable for imaging and conversion to a standard text file, the participant collection must make available a header that includes a reference to the storage location of the material. This reference must be descriptive enough for users (or software) to identify the location of the material and how to access the material.

2.23 The system must host its database in an Open Database Connectivity (ODBC) compliant RDBMS.

This core capability ensures a standard method for accessing relational (structured) data, which is how the bibliographic data (headers) is characterized. The LSN site must organize structured data using a relational database management system (RDBMS), and that RDBMS must be ODBC-compliant to provide a standard method of accessing relational data. See requirements 2.1, 2.4, 2.5, and 2.18 above.

A second core capability is implicit with the requirement for an ODBC-compliant RDBMS. ODBC is a standard or open application programming interface (API) for accessing a database. ODBC is based on and closely aligned with the standard SQL Call-Level Interface. It allows programs to use SQL requests that will access databases without having to know the proprietary interfaces to the databases. ODBC handles the SQL request and converts it into a request the individual database system understands. The LSN's database software must provide access to documentary material through SQL-based structured index searching (on bibliographic header record fields). This standardization provides users with uniformity of the language used in the database query processes even though there are multiple, diverse underlying databases.

Moreover, ODBC/SQL provides this standardization to text as well as structured databases. This will allow the LSN to provide full text, image and bibliographic search and retrieval capabilities in a single search interface screen (although there may be a "single screen" each for simple vs complex searches); to do so without requiring complex navigation or differing protocols; and to move between and among full text, images, and bibliographic entries without having to iteratively open and close different applications packages. These underlying standards are what will allow users to simultaneously search for key words and terms against both bibliographic databases and full text files.

2.24 The system must be capable of adding additional nodes to address high user demand.

This core capability is required to address scalability requirements. The LSN should be designed using modular design techniques for both its hardware and software, and have well documented software interfaces. To meet this requirement, the Operating System (OS) software should be a mature, robust operating system and be interoperable, capable of working on multiple platforms. (~~specifically, UNIX?~~)

Server platforms must also be scalable. Because of the number of documents related to the licensing, each component of the web-based system -- participant site, portal site, and docket site -- must be scalable. Server hardware must be high-speed, high-

performance and support or be upgradable to multi-processors (dual, quad, or eight-way, depending on application resident on the server.

The LSN central site shall be sized to support a maximum of 150 simultaneous users. The profile of users consists of approximately 500 routine users associated with the licensing itself, an additional 500 occasional users who may routinely monitor the site because of subject content interest, and an additional unknown number of users who may access the site because of web-generated "hits".

2.25 The system must be capable of adding additional nodes to address increased numbers of documents.

This core capability is required to address scalability requirements as noted in the previous section. The LSN should provide an expandable storage capacity capable of storing index files for retrieval of the volumes of documents identified in Appendix B.

2.26 The system must isolate content collection and processing from user service (i.e., search service not interrupted if a scanning agent fails.).

This core capability is required to ensure uninterrupted search and retrieval performance while database building is under way. This means that the server operating system software for the LSN servers must have the ability to perform multiple functions simultaneously while search queries are being processed, e.g., the Operating System Software for the LSN servers must be multi-tasking.

2.27 The system must provide a mechanism for manually securing categories of content within the portal software directories for authorized individuals, groups, and repository collections.

This core capability is essential to meet requirements for providing a Protective Order File. The LSN must provide a Protective Order File, which in all likelihood will reside within the NRC docket server. To provide this capability, the LSN site must provide login and access control tools to get access to the stored text and images in the underlying collection repository. Implicit in system operations, there may be other administrative files that are located on the LSN server to which general users (e.g., non-administrators) should not be allowed access.

Finally, there may be instances where parties to the proceeding must be given priority access to the system at the expense of wide-open public access.

These three factors mean that the LSN site must provide the capability to initiate a user session and allow access to the LSN system by providing an authorized user account name and associated authorized password. Non-participant users would be "guests" in database terminology. Similarly, even for administrators, the LSN site should not allow access to system functions without the initial entry of an authorized account name and associated authorized password. ~~Once the system is activated by a user, user sessions must be able to be closed out to conserve on user software licenses; etc., needed to support the LSN. Therefore, the LSN should provide the capability~~

~~to terminate a working session on the LSN, leaving the user in a state where a new system login is required to gain access to LSN functions.~~

- 2.28 The system must provide a mechanism for manually securing individual items within the portal software directories for authorized individuals, groups, and repository collections.**

This core capability is essential to meet requirements for providing a Protective Order File as detailed above and for restricting access to administrative files on the LSN server.

- 2.29 The system must provide multi-level access controls, or similar mechanisms, which allows users to see metadata information on documentary files for which they do not have appropriate security rights to view the underlying object.**

This core capability is essential to meet requirements for providing a Protective Order File. The LSN must provide a Protective Order File. If the LSN site is going to successfully prohibit user access to data which the user has not been authorized to access, then the LSN must provide multi-level access control.

- 2.30 The system must support multiple named administrators with different security profiles and different administrative domains.**

This core capability is essential to allow contractor and federal staff to perform administrative functions as required for operational, legal or security reasons. The LSN will have a search and retrieval component and also an audit component. The LSN design must provide a mechanism to control access to audit system functions even for those system staff who may be authorized to support the search and retrieval resources.

- 2.31 The system must support groups of administrators with the same security profiles for different administrative domains.**

This core capability is essential to allow contractor and federal staff to perform administrative functions as required for operational, legal, or security reasons. This allows for staff redundancy and backups for performing administrative functions.

- 2.32 The system must provide a mechanism to secure security profile information stored in the LSN server.**

This core capability provides the LSN Administrator with the necessary tools to ensure the security of the LSN. The LSN must be able to prohibit user access to functions which the user has not been authorized to use. **The LSN will rely on this to provide methods of connection throttling and prioritization based on user security profile information. It should utilize, at a minimum, techniques to secure its databases such as by providing operating system user authentication via ID and password; providing for the creation of Access Control Lists (ACL) at the file level, database management system (DBMS) user authentication that includes the ability to establish specific database grants and permissions, and other similar techniques.**

2.33 The system must provide a mechanism to generate results/reports.

This core capability provides the ability to provide predefined and ad hoc reports that can be selected and run by the users.

2.34 The system must provide a mechanism to schedule (automate) content collection, processing, and other system processes.

This core capability provides a means to routinely and systematically canvass participant collections and to update the LSN indices in an automated fashion. Additionally, these capabilities are the basis of control mechanisms used to identify changes to underlying object files and databases. The LSN must provide a function to allow all users to detect that subsequent revisions to a document exist. For **revised documents that were made available in their previous iteration** for more than five days on a website, the participant is required to add the revised version as a separate document.

This capability, in conjunction with "create date" and "update date" attributes within the LSN RDBMS, would allow the LSNA to implement a capability wherein all LSN users can identify the initial date that a record is posted to a participant's system. Additionally, each document intended to be used as an exhibit must **be made available** before the commencement of that portion of the hearing where it will be offered, and this capability is used to identify that the condition has been met.

2.35 The system must allow for individual administrative processes to be aggregated into jobs which may be scheduled as a unit and must preserve dependence within a job stream.

This core capability provides a means to routinely and systematically canvass participant collections and to update the LSN indices in an automated fashion. It also provides the underlying capability to perform system backup and recovery processes that will be run on a routine basis. The LSN server must be capable of storing and maintaining backups consistent with requirements for data availability, especially during the hearing process. Backup capabilities must execute within a window that will not impact the general user population performing search and retrieval. ~~are 2 Gigabytes of data per hour.~~

Similarly, the LSN server must be capable of being recovered from backup media. Recovery capabilities must execute within a window that will not impact the general user population performing search and retrieval. ~~are 1.5 Gigabytes of data per hour.~~

2.36 The system must allow for distributed administration.

This core capability allows LSN staff and contractors to remotely perform system administration functions **in a secure way over the Internet**. This provides implementation and operations flexibility in locating the LSN system resource and also in quickly responding to service requests from either NRC headquarters or from the site of the hearings.

2.37 The system must allow for performing routine administrative and maintenance activities from a remote console, workstation, or terminal.

This core capability details the specific capabilities needed by LSN database administrators: the ability to perform database administration (start up, shut down, file maintenance, tuning, etc.), monitor session activity and system usage, administer user accounts, backup/restore, and otherwise monitor system performance. The LSN must provide access to the servers, and all of their services, via remote access for NRC staff who are authorized to perform various activities (e.g., search and retrieve, upload transcripts, and generate reports).

The LSN site must provide the LSN Administrator with the necessary tools to ensure LSN availability and the integrity of the LSN database. These capabilities include such basic functions as the capability to initialize the software and hardware necessary to operate the LSN, and the capability for the orderly shut down of the software and hardware components of the LSN. To accomplish file maintenance, the LSN site must provide authorized users with the capability to perform changes to the database structure (adding, deleting, modifying fields). This would include database administration features (which cannot all be predicted here) which include having an editable table of valid field values for the LSN Header and any other header information in the system. The LSN must provide authorized users the capability to adjust database performance parameters or to restrict or disable database features in order to optimize system performance.

For the LSN audit site itself, the system administration staff will require network monitoring tools needed to aid in the discovery or diagnosis of network faults and suboptimal configurations.

System administration staff will require tools to monitor usage and access.

System administration staff will require tools to monitor and aid recovery from attempted or successful security incidents.

The LSN design anticipates that participants will house their relevant documents on a computer hardware and software resource that they provide. However, the LSN design may include its own databases, files, indices, etc., in order to optimize performance or deliver functionality identified later in the design process. Server software (operating and applications) must provide network administrators with the services to monitor database information effectively. Therefore, the LSN will require comprehensive database administration tools regardless of the participant systems' requirements. The core database administration tools for the LSN site should include the capability to:

- edit previously stored documents;
- designate specified fields in the LSN header as protected from or available for update;

- prohibit the editing of header fields designated as protected, except for headers which are being created for the first time within the LSN;
- perform changes to data properties (field values) on an individual record basis;
- perform changes to data properties (field values) on a group of documents via global search and replace commands;
- identify validation checks that may optionally be applied against data properties (field values);
- logically delete documents (header, text and/or image) such that they will not be available to LSN users;
- remove, in real-time, document files and their indices from access to the public;
- prevent the deletion of a header before the associated text **and/or** image have been deleted;
- perform validation against an access control list to ensure that only authorized users may delete files;
- provide a confirmation dialog **box** for all destructive actions;
- remove document files and their indices from backup media;
- make a list of documents that have been logically deleted; restore a document that has been logically deleted; and make the document again available to users;
- store non-document-related items, files, or tables that are related to system administration and database administration activities;
- search and retrieve non-document-related items, files, or tables that are related to system administration and database administration activities; and,
- delete non-document-related items, files, and tables that may have been collected during the course of database and system administration activities.

Even though the database is properly tuned, it may be necessary to take other steps to ensure performance during critical periods. Therefore, the LSN must provide a capability, **should it become necessary**, for an authorized user to monitor user session activity levels and to identify and cancel queries or other system activities in order to **ensure** access and performance during the hearing.

The LSN must have the capability to account for user **activity**. It must provide the basic capability to create new user accounts, modify user accounts, and delete user accounts, including the establishment of the account names and initialization of the account password. In order to facilitate implementation of a Protective Order File capability, the

LSN must provide the capability to establish and deny read, write and delete privileges for each of the following types of information on a per-user account basis: documentary material; header data; text data; password; and image data. Finally, the LSN must provide the capability to establish and modify user access authorizations to system functions unique to system administrators.

To ensure system and data availability without impacting the three year hearing schedule, the LSN must provide the capability to create incremental and full backup copies of all data on the system. This implies, depending on architecture, that the LSN will incorporate an archival store, either near-line or off-line, and provide the capability to copy files from its primary operating platform to an archival store. The LSN must support the complementary **process** - the capability to restore data created by the backup function to the operational system, including partial and full data recovery.

The LSN must provide authorized users (administrators) the capability to monitor the status of system hardware, software, and communications components and to interrupt, restrict or disable system capabilities in order to optimize use of system resources. The LSN must provide performance monitoring software needed to verify compliance with response times shown in TAB B.

Finally, the system must be inherently capable of performing in the time frame envisioned **and furthermore to anticipate that some LSN data may be needed for subsequent licensing (emplacement, closure) actions.** Therefore, the server hardware and software must be year 2000 compliant and must be capable of interacting with browser clients which themselves may, or, may not be year-2000 compliant. Specific attributes of this compliance include that:

- no value for a current date prior to the year value **2037 2049** shall cause an error in calculations in any component of server software;
- it shall accurately represent and interpret dates in the range of January 1, 1900 to December 31, **2037 2049**, and that this date shall remain constant as time progresses;
- it shall accurately search and display dates without ambiguity;
- it shall accurately sort 19xx and 20xx dates in ascending/descending order;
- it shall not use a "00" or a "99" in 2-digit date fields to mean something other than date (i.e., estimated date, unknown date, does not expire, etc.);
- it shall provide unambiguous century designators, or, if a date is represented without a century designator, the century shall be unambiguous for all manipulations involving that element (different rules may be formulated for different elements);
- it shall not use a 2-digit data as part of an indexed key that depends on these records to be returned in chronological order;

- it shall accurately sort, search, and display **date-derived** identifiers without ambiguity and without jumbled sequences on reports or screens;
- it shall return accurate results for date calculations; and
- it shall return accurate results for leap year and Julian date calculations.

2.38 The system must support full-text searching across all content in the directory, including titles, summaries and full text object files stored in their native format.

This core capability of the LSN site specifies underlying formats (and data stores - text and header) against which text retrieval capabilities must operate. The LSN must provide the capability to query the system by specifying one or more character strings in the full text of the document to obtain a list of all documents which satisfy the query. Specific text retrieval operations to be provided include the ability to: search for a phrase in the full text of documents; search for a phrase in selected bibliographic header fields (those containing text); and to exclude specific words from full-text indexing operations (stop words). Given that there will be a large volume of text data with content and terminology not intuitive to users, the LSN must provide vocabulary **tools, such as a dictionary dialog box, that allow** users easy access to the words in the text database without having to conduct a full search to determine a **word's** presence or frequency of occurrence.

2.39 The system must support full text search capabilities using Boolean operators.

This core capability of the LSN site specifies the search methodologies needed to operate against large and complex text files. The search engine(s) used to search the text and header indices in the LSN must provide the essential core Boolean operators "AND", "OR" and "NOT" in a query construct, **including nested queries.**

2.40 The system must support full text search capabilities using Natural Language queries.

This core capability of the text search component allows users, especially less advanced users, to enter a query in the form of a sentence or question.⁵ The software will automatically identify the important words and phrases within the query and remove the "noise words". The text retrieval software that accommodates natural language query capabilities should not remove noise words from a phrase that is within its vocabulary. (Example: What is the state of the art in text retrieval? The software should search for: state of the art AND text AND retrieval).

2.41 The system must support Proximity Searching.

⁵ There are at least three credible suppliers of natural language interface software for databases at this time: English Wizard from Linguistic Technology; a software component called English Query from Microsoft, which is now distributed as part of SQL Server 7.0; and Access ELF, which translates queries on-the-fly from English to SQL and doesn't store information in any form.

This core capability provides that the full text search engine(s) used to search the text indices (and text fields in structured data indices) shall provide proximity operators, "adjacent" and "near", to allow the searcher to find words that occur near each other, from a simply adjacency to a distance of nn intervening words.

2.42 The system must support root searching in both structured and unstructured data.

This core capability provides that users may conduct a search by using wildcards (e.g., "*" or "?") in all query construction modes. (Where "*" is used for multiple characters up to and including a complete word, and "?" is used for a single character). This basic capability **must** ~~should~~ have extensions, such as the capability to specify interspersed single and multiple character wild cards within a query argument, to utilize wildcards in conjunction with proximity searching, to use root searching as part of a full text query, and to combine multiple query statements using Boolean expressions (e.g., AND, OR, NOT).

2.43 The system must support searching for specific metadata values in structured data via basic structured query tools.

This core capability of the LSN site provides users the capability to electronically search and retrieve the bibliographic headers in the system. This means that the LSN must provide the capability to construct a query by using both document profile and full-text search capabilities for execution in a single query. This basic capability should have extensions, such as the capability to query the system by specifying the content of more than one header field to obtain a list of all documents which satisfy the query; the ability to search a single field, multiple fields, all fields, text, or text and any combination of record data, as specified by a user; and the capability to query the system by specifying a combination of header field values and the text query parameters in the full text of the document to obtain a list of all documents which satisfy the query.

Further, **this core capability** specifies the search methodologies and tools needed to operate against large and complex structured data (bibliographic) files. The LSN search engine(s) used to search the header indices must provide the essential core relational operators to aid in searching formatted fields **such as date, numeric, and text fields**. Because user sophistication varies, the software must provide the ability to search records by using data values that are either entered in a search window manually or are selected from a controlled vocabulary (via picklist). This implies that the software will provide the ability for a user to see a list of values that have been entered as a property into a field. (E.g., view the index values for the field). **Additionally**, because the content of the underlying documents is highly technical and scientific in nature, the LSN should provide a controlled vocabulary lookup feature for keywords in a structured bibliographic header field (select from list only, no ad hoc entry by end user). The LSN should provide tools to assist the user in retrieving documents when the unique data elements (identifiers) for the documents are not known to the users. Examples might include synonym processing, thesaurus, natural language queries, or other search aids. The LSN must contain a thesaural capability that provides recommended alternative terms to the user-supplied search words.

Original LSS functional requirements anticipated a closed system where additional query tool could be made available. These were often **custom-developed** functions based on features identified in DBMS technologies in use between 1988 and 1994, and are indicative of the types of functions that would be useful to trained information management intermediaries and paralegals who were conversant with discovery database search techniques. These tools included:

- the capability to save, retrieve, edit and/or re-execute a previously constructed query;
- the ability to tag a specific document in a search hit list for action, both as selected individual items, or, as part of a contiguous set of specified documents;
- the capability to save and retrieve a result set of a previously executed search;
- the capability to perform a query against the results set of the previous query during a single session;
- the capability to query the system for a list of all documents that meet the query criteria and perform ascending/descending sorts of the displayed list on the basis of selected display fields, or the basis of relevancy to the query;
- an indication of the query status during a full text query;
- the capability for a user to terminate queries in process without terminating the session or losing previous result sets; and,
- the capability to cancel a constructed query before execution (abandon).

It is reasonable to expect that some of these capabilities may be inherent in the LSN host software, but their implementation may be constrained by security concerns which **restrict user ability to write data to the web server.**

2.44 The system must support searching using Boolean combinations of full-text and metadata values.

This core capability allows users to conduct a search by combining structured and unstructured search techniques in all query construction modes. Because of the size of the database and the potential impact of issuing queries that do not attain a high degree of precision, the LSN must provide the ability to search for and retrieve both bibliographic header and full text documents.

2.45 The system must rank results by relevance with respect to the search term, without regard to the type of the document returned or the repository the original document resided in.

This core capability is essential given the size of the LSN collection and its homogeneity.

2.46 The system must be capable of application upgrades.

This core capability is required to address system currency, maintainability, and scalability requirements. This capability implies that the system to be developed will not rely on extensive customization which could prevent the system from being routinely upgraded with new vendor releases. Customization engenders extensive regression testing against new releases of software and if extensive, could be cost prohibitive. Therefore the client and server software should be a commercial off-the-shelf product that can be installed with only minimal customization (field definition, screen design, etc.).

2.47 The system must provide on-line documentation.

This core capability provides on-line documentation as part of an on-line help capability. This is a critical capability because of the diversity in the level of sophistication of the users. To support users, the LSN must provide help screens to assist user interaction with the (OS, RDBMS) system processes and to respond to system messages, and help screens to assist user interaction with the application software and to respond to application software messages. The online help features of the LSN Portal and Docket should be field-sensitive and should include narrative, not just a cryptic, system-jargon, numbered E-message. For example, the LSN should provide interactive capabilities to assist the user in retrieving documents when the field values that uniquely define the documents are not known to the user.

2.48 The system must provide other (non-online) help documentation.

This core capability provides paper or CD authored user documentation.

3.0 FUNCTIONALITY FOR DOCKET SYSTEM

The Licensing Support Network is intended to support the hearing process and the NRC resources established for maintaining licensing dockets. The LSN provides no resources to accomplish the docketing process within NRC, but still must integrate with the NRC docket. The following is a roster of functional capabilities associated with NRC's docket system which can be used as a frame of reference for the integration objectives to be considered in the LSN's detailed design phase.

3.1 Docket

The docket system must, first and foremost, incorporate **all underlying system requirements found in NRC's docket and records management directives and the established procedures and standards for motions practice.** For the Yucca Mountain Repository licensing, the docket functionality begins 30 days after DOE submission of **the license application, in appropriate electronic format, to SECY.** The LSN utilizes NRC's existing docket structure to provide a docket which receives, stores, distributes, and maintains pre-license application docket materials. It is a fully featured capability, and allows for creating bibliographic headers for transcripts (and associated exhibit materials) submitted which were not previously made available in a participant collection. It also allows for creating a digital image of each page of electronically submitted text material (depositions, transcripts, and attached exhibits) that were not previously made available as an image in a participant collection. The software environment in with the docket resides preserves transcript formats including page and line numbers in the image format for uniformity in reference. It also provides the ability to link document records in the docket file with their **point-of-reference in the full text transcript.** Additionally, it already includes the ability to display limited access warnings.

For authorized individuals (SECY), NRC's Docket:

- provides a mechanism by which transcripts may be entered **into** the docket on a daily basis in order to provide next-day availability at the hearing;
- provides mechanisms for participants to transmit depositions to NRC (SECY) in electronic form for entry into the docket; it also provides a mechanism to authorized users (SECY) to identify only a part of a deposition that is offered into evidence;
- provides an electronic list of all exhibits with a way of indicating where they were introduced; and
- is the repository for the Protective Order File materials.

NRC's SECY organization is authorized to maintain the docket. The LSN site will contain a link directly into the existing NRC docket files and password access to the Protective Order File. **This is one of a number of mechanisms (such as direct access to ADAMS external server) by which the Presiding Officer and counsel for all parties have access to the electronic docket (including Protective Order File) during the hearing.**

NRC resources are also used to support motions practice, i.e., the process of formally communicating, submitting and responding to legal matters, and which is conducted between the parties and the Board. This is the mechanism whereby all filings are able to be submitted/received electronically. The NRC provides this mechanism - Electronic Information Exchange (EIE) - for **participant use in** electronic submissions for all **motions practice**. The NRC will receive electronically transmitted depositions via EIE and enter them into the docket file. Similarly, the **Presiding Officer and Commission** issuances and orders are also transmitted electronically via EIE.

This core capability provides a means to authenticate transmitted files in support of motions practice. The NRC has established a separate, agency-wide **EIE** capability that ensures the integrity of files being transported across the internet. The LSN site must be able to utilize this existing NRC **capability, which** is based on UWI™ forms and Verisign™ digital signaturing software. The NRC procedure provides a mechanism whereby all filings **can be submitted/received electronically**. **This procedure** incorporates password security code techniques as part of the digital signature certificate issuance procedure, and **applies** digital signaturing technologies for transmission of documents. The NRC procedure also provides a mechanism to deliver all answers, orders and decisions per §2.1013(c) [e.g., electronically, **using a secure process**].

Note: Procedurally, a **participant's representative, when being deposed, is required to serve an electronic index of all documents in his/her possession relevant to the subject matter of the deposition on all parties. This index must identify which documents** were already made available electronically and, by inference, all those not made available in the LSN. This functionality must be met by the participants' own resources.

3.2 Docket Submission Controls.

The docket must be **governed by** procedures and tracking mechanisms to allow document submitters to verify that document information entered into the **NRC** docket database is identical to the document information submitted to the LSN. Additionally, the docket environment must implement certain data management capabilities, including:

- a mechanism to preserve transcript formats including page and line numbers in the image format for uniformity in reference; and,
- the ability to link document records in the docket file with their reference in the full text transcript. *The implementation of this functionality should not use hyperlinks (records retirement issues) and should be accomplished in a single environment that is easily understandable and quick to learn.*

3.3 Docketing Process.

Each participant is required to provide the computer system necessary to comply with service requirements (§2.1011(b)). This includes providing an automated/procedural mechanism to show a proof of service for each document filed into the docket (§2.1013(c)(4)). **This captures the event of** a document being placed in a recipient's mailbox in order provide proof of the completion of service (§2.1013(c)(3)). This may be accomplished by procedurally requiring an echo response to incoming e-mail if received from a defined party to the licensing. The NRC docket must also implement procedures and tracking mechanisms to track documents submitted to NRC from receipt (date, time, accepted/rejected status) through disposition.

3.4 Docket Availability.

The NRC docket environment must provide a mechanism that allows an authorized user (SECY) to identify the unavailability of the electronic docket for more than 4 hours in any day, and a means to communicate unavailability to the **Presiding Officer**, so that the day is not counted in the computation of time. (§2.1017)

4.0 AUDIT SYSTEM

The Licensing Support Network Administrator (**LSNA**) is the individual within the NRC responsible for coordinating access to **participant** data via the LSN. The LSNA provides technical support to the **Pre-License Application** Presiding Officer in the matter of verifying substantial and timely compliance with the requirements in §2.1003 regarding availability of material in electronic form. Additionally, the LSNA is responsible for the ongoing integrity of the data that has been made available. To accomplish this mission, the LSN will have an audit and reporting capability to independently monitor system activities of the LSN site and the underlying participant collections.

4.1 Availability of electronic data

To ensure the ongoing availability of data, component subsystems such as the participants' servers and the LSN server site itself must be made available to the LSN audit system for collection of server activity associated with posting, modification and deletion activities; and for server performance data in responding to requests for files. The audit system will provide authorized individuals the ability to monitor participants' document server performance in providing requested object files to requests made through the portal. (Such as for images associated with previously retrieved text.)

For the LSN audit site itself, the system administration staff will require network monitoring tools needed to aid in the discovery or diagnosis of network faults and suboptimal configurations.

System administration staff will require tools to monitor usage and access.

System administration staff will require tools to monitor and aid recovery from attempted or successful security incidents.

4.2 Integrity of data

To ensure the ongoing integrity of data, component subsystems such as the participants' servers and the LSN server site itself must be made accessible to the audit systems' recording of all activity (accounting) associated with posting, modification and deletion activities conducted on each underlying document collection's bibliographic, text, and image files.

4.3 Audit Tools

The audit system must provide software tools to:

- collect audit data on all aspects of LSN performance;
- store audit data as a record;

- aggregate, compile, crosscut, and otherwise analyze audit data by providing a software capability (package) which can be used to perform the analyses and provide mechanisms to generate output and reports;
- create user-defined reports which include descriptive statistical computations and sub-grouping (ad hoc);
- store user-defined, statistical reports on participant, portal, and docket system activities; **and**
- print statistical reports on participant, LSN server, and docket system **activities.**

The audit system should allow only authorized individuals to enable/disable the audit functions and to backup and remove audit files from the system. Additionally, the audit server should adhere to the same standards, where applicable, as the LSN search and retrieval server in order to ensure compatibility.

TAB A

Minimum Header Fields for LSS Documentary Material
OLD Header Structure
LSN Candidate Deletions marked “(DELETE?)”

LSS Field	Mandatory or Req'd by Participant	Mandatory or Req'd by LSSA	Multivalued (max # of entries)	Max Field Length	Authority Table	Required Format
LSS Accession Number	N	M	N	13	N	Must be a unique number 3 digit alpha code w/ submitting organization, 7 numerics and check digit
Participant Accession Number	M	N	Y (50)	25	N	Alpha numeric, no required format
Title	M	N	N	1000	N	N
Author Name	M	Y	Y (200)	50	Y	Last Name, First Initial, Middle Initial
Author Organization	M	N	Y (200)	65	Y	
Document Date	M	N	N	8	N	YYYYMMDD
Document Number	R	N	Y (5)	30	N	N
Version	R	N	Y (5)	50	N	N
Access Control Information	R	N	Y (10)	3	Y	N
Related Records Number	R	Y	Y (500)	25	Y	Alpha numeric

LSS Field	Mandatory or Req'd by Participant	Mandatory or Req'd by LSSA	Multivalued (max # of entries)	Max Field Length	Authority Table	Required Format
Related Record Code	R	Y	Y (500)	7	Y	N
Special Class	R	N	Y (10)	50	Y	N
Abstract (Field under consideration to be deleted)	N	N	N	5000	N	N
Package Identifier	R	Y	Y (500)	50	N	N
Document Type (Includes package types)	M	N	Y (3)	40	Y	N
Identifiers	N	N	Y (100)	80	N	N
Comments	N	N	N	1000	N	N
Media	R	N	Y (5)	7	Y	N
QA Record	M	N	N	1	Y	N
Traceability Number	R	R	Y (10)	50	N	N
Traceability Code	R	R	Y (10)	5	Y	N
Keywords	N	N	N	5000	N	Y - Separate terms and phrases by punctuation
Number of Images	N	M	N	6	N	N
Physical Location Reference Information	R	R	N	1000	N	N
Address Name	R	N	Y (500)	50	Y	Last Name, FI MI

LSS Field	Mandatory or Req'd by Participant	Mandatory or Req'd by LSSA	Multivalued (max # of entries)	Max Field Length	Authority Table	Required Format
Addressee Organization	R	N	Y (500)	65	Y	
LSS Record Housekeeping Info: Date Received at LSS Date Available in LSS Date/Time Loaded into LSS Date/Time of Last Modification LSS Indexer ID (DELETE?) Station ID (DELETE?) QC ID (DELETE?) Subject and Abstract Cataloger ID (DELETE?) Cataloging QC ID (DELETE?) Processing Stage Status (DELETE?) Verification ID (DELETE?) Change Tracking Log Electronic Signature Verification Electronic Image Location Searchable Text Reference Info.	N	Y	TBD	TBD	TBD	TBD
LSS Audit Info	N	R	TBD	TBD	TBD	TBD

Symbols: Y = Yes; N = No; R = Required; M = Mandatory; TBD = To Be Determined For date fields, Y = year, M = month, D = day
Table definitions:

Data submitted by participant: This field will be submitted by the participant (Mandatory = must be provided for each unit [record]; Required = must be provided if applicable; Optional = provided at discretion of participant.)

Provided by LSS System or LSSA: This field will be provided by LSS. (Mandatory = must be provided for each unit [record]; Required = must be provided if applicable)

Multi-valued: Multiple entries allowed in a field.

Controlled Authority List: List of accepted entries to be used by all participants, such as document types or specific forms of an organization name.

Text searchable: The ability to perform phrase or single-word searches of the field entries.

Comments/Issues: Any additional comments or outstanding issues.

Note: Detailed definitions of header fields are provided in TAB C.

TAB B

Old LSS Response Time Requirements (New LSN Benchmarks Required)

Requirement Identifier	Function/Event	Conditions	Response Time 15/50 concurrent users
LSS2-065-2	Retrieval of query results list.	UNLV test query INJD-T3-Q1 or TEJA-T#-Q2.* Database contains headers for at least 5 million pages of documents. A total of 10 documents found.	90 seconds/140 seconds
LSS2-065-3	Retrieval of header data for document identified in query results list.	Database contains headers for at least 5 million pages of documents.	10 seconds/15 seconds
LSS2-065-4	Retrieval of text data for document identified in query results list.	Database contains at least 5 million pages of document.	First page: 10 seconds/15 seconds Each subsequent page: 3 seconds at the Main Facility, 6 seconds at the Supported Sites
LSS2-065-5	Retrieval of image data for document identified in query results list.	Database contains at least 5 million pages of documents.	First page: 30 seconds/45 seconds Each subsequent page: 6 seconds at the Main Facility 9 seconds at the Supported Sites
LSS2-065-6	Document bilevel scanning	8 ½ inch paper. All prior data entry needed for document scanning complete prior to initiation, paper loaded on scanner.	30 pages per minute for single sided, 15 pages per minute for double-sided (30 bilevel images created)

Requirement Identifier	Function/Event	Conditions	Response Time 15/50 concurrent users
LSS2-065-7	New Document Access	Measured from the time a new document (header, text and image) has been captured and stored in the LSS until it is available for retrieval and viewing at the supported sites.	24 hours. Comment: This means that any method of dissemination from the capture site(s) to the retrieval sites must support access to newly entered documents within 24 hours, on an ongoing basis. Weekly or monthly updates to the reading rooms, for example, would not be acceptable.
LSS2-065-8	New Transcripts	Measured from the time that daily transcripts are imported into the LSS until the time they are available for viewing as text.	1 hour.
LSS2-065-9	Prepare Searchable Text	Time to perform <u>Prepare Text for Search</u> function. (See Section 3.1)	Less than 20 seconds, on average, to add a document consisting of 10 full text pages, to an existing text base of 5 million pages.
LSS2-065-10	Backup Data	Time to backup system data of any type. Incremental or full backup.	2 Gigabytes per hour
LSS2-065-11	Restore Data	Time to restore system data of any type.	1.5 Gigabytes per hour

*These queries were selected as representative of typical user queries which are not overly simple or complex. INFD-T3-Q 1: Find documents where text includes phrase like 'repository' & 'seals' or text includes phrase like 'shaft' & 'seal' or text includes phrase like 'borehole' and 'seal', order by document id. TEJA-T3-Q2: Find documents where text includes phrase like 'faults' & 'Basin and Range Province' or text includes phrase like 'faults' & 'Nevada' or text includes phrase like 'faults' & 'Yucca Mountain' order by document id.

TAB C

Abstract

A brief narrative description of the subject content of the document, or a full description of the contents of a document that cannot be imaged and converted to searchable text. The abstract is generally written by the author.

Access Control Information

A code indicating that access to a document is restricted. Access is restricted if the document is privileged, proprietary, or copyrighted.

Addressee Name

The names of all the persons to whom a document is addressed. Each entry in this field is linked to a corresponding entry in the Addressee Organization field.

Addressee Organization

The affiliation of each receiver or the organization to whom the correspondence is addressed if there is no personal receiver. Each entry in this field is linked to a corresponding entry in the Addressee Name field.

Author Name

The name of each person listed on the document as responsible for all or part of its creation. Only personal authors are entered in this field. Corporations as authors are captured in the Author Organization field. Each entry in this field is linked to a corresponding entry in the Author Organization field.

Author Organization

The name of the organization (i.e., company, corporation or group) with which the author is affiliated at the time the document was created, or the name of the organization responsible for creating or originating the document when there is no personal author. If an author works for one organization and is representing another, both affiliations should be captured. Each entry in this field is linked to a corresponding entry in the Author Name field.

Comments

Any information not covered in other fields which the submitter or indexer believes would be of help to identify or retrieve the document, or to further explain any field entry for the document. This field can be used for entries such as the language of the document (if it is not English) or the page numbers that are missing in an incomplete document.

Document Date

The date on which the document was completed, issued, effective or published. If the date is unknown, information in the document will be used to estimate a date.

Document Number

The identifying number(s) assigned to a document that distinguishes it from other documents (e.g., DOE Order No., Public Law number, report number). Document numbers appear (typed or handwritten) on the document itself and are considered to be control numbers. The Document Number is generally assigned by the issuing agency. Examples are report numbers, or public law numbers such as SAND86-1023, PL95-16, or H101-364.

Document Type

The general format or physical presentation of the document. Examples include correspondence, report or procedure.

Identifiers

Words or phrases that the submitter or indexer believes represent the subject content of the document and will assist users in retrieval. These may be acronyms or informal terms or cross-references to alternate nomenclature. The terms in this field may be included in a controlled vocabulary/LSS Thesaurus.

Keywords

Words or phrases that the submitter provides with the document to represent the subject content of the document and to assist the user in retrieval. Keywords are not necessarily contained in the LSS Thesaurus.

LSS Accession Number

A unique identifier assigned to each document entering the system. The capture station at which the document enters the LSS is also identified as part of this number. The LSS Accession Number will also be used as a Related Record Number pointer for documents which have relationships to other documents in the LSS.

Media

The physical material upon which a document is stored.

Number of Images

The number of images of a document that was imaged from a hard-copy.

Package Identifier

An identifier assigned to all components of a group of documents submitted as a single entity. This field enables a package containing many documents which may or may not have relationships among them to be reassembled quickly and easily.

Participant Accession Number

A unique identification number assigned by the participant organization to each document submitted for entry into the LSS. This number assists the organization in locating documents it has submitted. This field should contain a specific alpha code identifying the participant organization, e.g., DOE, NRC, NEV, and any other alphanumeric scheme which the submitting organization might use in accessioning their own documents. The number used may be the accession number used in the submitting organization's records system.

Physical Location Reference Information

Information on the storage location of an item submitted to the LSS as header only because its form does not permit imaging.

QA Record

An indicator of whether the document is a quality assurance record. Quality assurance records are those whose contents have been determined to furnish evidence of the quality and completeness of data, items, and activities related to the safety of the repository program.

Related Record Code

The code that represents the type of relationship between the document being entered and a document to which it is related. Each code in the authority list will have a reciprocal code; for example, the reciprocal of a document (A) that is attached to another document (B) is document (B) has attachments (A). Examples of Related Record Codes include: REV (revises or is a later version of), COR (corrects) or SUPR (supersedes). Each entry in this field is linked to a corresponding entry in the Related Record Number field.

Related Record Number

This field contains the LSS Accession Number(s) of a document that has a particular relationship to the document being entered. There are several types of relationships, such as: parent/child (a document and its attachments); original/subsequent (a document and a later version, comments, corrections, or errata); and whole/part (a book and its chapters, a journal and its articles); and an information package and the cataloging units it contains. The type of relationship is captured in the Related Record Code field. Each entry in this field will be linked to a corresponding entry in the Related Record Code field.

Special Class

This field identifies documents with special characteristics that are not captured in other fields; for example, a document in a foreign language.

Title

An identifying sentence or phrase given to the document that appears on the document, i.e., the actual title. If the actual title is not present for a document, a title must be created.

Traceability Code

A code that indicates the type of traceability number. Examples of this code include: DTN (technical data link), DI (Document Identifier), and WBS (Work Breakdown Structure). Each entry in this field is linked to a corresponding entry in the Traceability Number field.

Traceability Number

An identifier that has been assigned to a document in order to link it to a specific activity. These identifiers will enable searchers to easily retrieve all documents associated with any given activity by providing a special linkage not available through other fields. Examples of traceability numbers include WBS numbers, Data Tracking Number, and configuration item identifiers. Each entry in this field is linked to a corresponding entry in the Traceability Code field.

Version

The version, revision number, or status of a document that has or will have multiple iterations. It will correspond to information contained on the document, e.g., Revision 2, Version 1. Final, or Draft.

TAB D
ESTIMATED LSS DATA VOLUMES
(LSN Data Volume Estimates Need to be Updated
Labat should have new charts ready by 2/14/2000)

Year	OCRWM Pages/Year	OCRWM Cumulative	NRC Pages/Year	NRC Cumulative	Others Pages/Year	Others Cumulative	Total Pages Added Yearly	90% Relevant Cumulative	50% Relevant Cumulative
1994	580 k	6,905 k	59 k	550 k	18 k	18 k	657 k	6,782 k	4,020 k
1995	750 k	7,655 k	59 k	654 k	23 k	41 k	832 k	7,584 k	4,522 k
1996	1,351 k	9,005 k	65 k	760 k	42 k	82 k	1,457 k	8,947 k	5,345 k
1997	1,682 k	10,687 k	71 k	891 k	52 k	134 k	1,804 k	10,644 k	6,369 k
1998	1,970 k	12,657 k	78 k	1,046 k	61 k	195 k	2,109 k	12,632 k	7,569 k
1999	2,013 k	14,670 k	86 k	1,203 k	62 k	257 k	2,161 k	14,663 k	8,795 k
2000	2,276 k	16,946 k	95 k	1,381 k	70 k	327 k	2,440 k	16,959 k	10,181 k
2001	2,371 k	19,317 k	104 k	1,567 k	73 k	400 k	2,548 k	19,351 k	11,625 k
2002	1,628 k	20,945 k	114 k	1,694 k	50 k	450 k	1,793 k	20,994 k	12,616 k
2003	1,584 k	22,529 k	126 k	1,818 k	49 k	498 k	1,759 k	22,593 k	13,581 k
2004	1,756 k	24,285 k	139 k	1,956 k	54 k	552 k	1,949 k	24,365 k	14,651 k
2005	1,708 k	25,993 k	152 k	2,089 k	53 k	605 k	1,913 k	26,088 k	15,691 k
2006	1,514 k	27,506 k	168 k	2,208 k	47 k	652 k	1,728 k	27,615 k	16,613 k
2007	1,674 k	29,181 k	184 k	2,339 k	52 k	703 k	1,910 k	29,305 k	17,632 k
2008	1,756 k	30,937 k	203 k	2,476 k	54 k	757 k	2,013 k	31,077 k	18,702 k
2009	1,247 k	32,184 k	223 k	2,574 k	38 k	795 k	1,509 k	32,335 k	19,461 k
2010	1,124 k	33,308 k	245 k	2,662 k	35 k	830 k	1,404 k	33,469 k	20,146 k