

A Comparison of Livelink 8 with DOE/LSN Requirements

ISRI Staff

Technical Report 99-02
Information Science Research Institute
University of Nevada, Las Vegas

May 3, 1999

1 Introduction

The following document reviews Opentext's Knowledge Management System, Livelink 8, with respect to the Department of Energy's (DOE) search and retrieval requirements for the Licensing Support Network (LSN). We used a "hands-on" evaluation to determine how closely this system meets each requirement.

1.1 Company and Product Information

Livelink is a product of Open Text Corporation which is publicly traded on NASDAQ. The company was founded in 1991 and has made several acquisitions to establish itself as one of the major providers of application software for document management. Since the acquisition of Information Dimensions in 1998, the company secures 42.7% of the market share of worldwide enterprise document management installations. Livelink applies a web-based interface and modular architecture to provide a collaborative working environment for enterprise document management, information retrieval, workflow and project collaboration, and enterprise group scheduling capabilities.

Livelink software requirements include:

- Netscape Enterprise Server 3.51
- Netscape Navigator 4.05 and/or Internet Explorer 4.01 with SP1
- One of the following relational database systems: Oracle, Sybase, or Microsoft SQL

2 Collection Preparation

Livelink depends on the Internet and the World Wide Web (WWW) for general and public access to its document management system. The most common interface for the user of the Licensing Support Network will be a *web browser*. Browser capabilities together with the flexibility of manipulating the system's interface and the resourceful use of HTML will dictate how the LSN collection should be prepared for loading.

System administrators should thoroughly review the LSN requirements prior to collection preparation. In particular, query, display, and print requirements affect how the collection should be prepared. The three data components of the LSN that should be considered carefully are the header record, the document text, and the page images. The LSN collection needs to be prepared with the objective of page by page display, text/image correspondence and header inclusion. By converting the text files to *Hypertext Markup Language* (HTML), ISRI was able to define document text pages, include document images and incorporate header fields for searching. Using this composite document format, we were able to meet several of the requirements of the LSN in Livelink. See Figure 1 to view the HTML source for the document collection. Although by

```

<html>
<head>
<TITLE>Proceedings of Workshop 28 on the Borah Peak, Idaho, Earthquake, Oct 3-6, 1984</TITLE>

<DocumentType>
  Reports
</DocumentType>

<DocumentSubtype>
  Technical Reports
</DocumentSubtype>

<PublicationDate>
  19850000
</PublicationDate>

<Author>
  Stein, RS Bucknam, RC
</Author>

<Keywords>
  EARTHQUAKES Hydrology Faulting Geodesy Seismology Geophysics
  Idaho
</Keywords>
</head>

<body bgcolor="#ffffff">
<h1>Proceedings of Workshop 28 on the Borah Peak, Idaho, Earthquake, Oct 3-6, 1984</h1>

<hr>

```

Figure 1: HTML source for inserted header record fields

using HTML tags in the HEAD of the document, ISRI was able to make the header fields searchable (see 1), we did not force the inclusion of these fields for display or printing in Livelink.

Through the use of HTML several of the requirements of the LSN were met but having a flexible *Application Program Interface* (API) could make this composite document an inherent component of the system's user interface. Also, several requirements that could not be easily met by manipulating HTML could become an intrinsic part of the document system. We were unable to evaluate the usefulness of Livelink's API for meeting the LSN requirements since the Livelink API is not available for the Solaris platform.

Since Livelink is fully web-based, several of the display and printing requirements of the LSN are dependent on the web browser and/or plugins in place. With several browsers and plugins available, it's difficult to know what capabilities are available to the user and what requirements of the LSN are being met. It is our recommendation that the DOE select and distribute a web browser after it has been tested thoroughly with the Livelink system.

3 Requirements Proficiency

3.1 General Requirements

(R) Year 2K Compliance. DOE-LSN is to be Year 2000 compliant.

1. Livelink passed minimum conformance tests conducted by ISRI.

2. Open Text states , “Livelink *Version 8.1* is British Standards Institution DISC YEAR 2000 conformant for the years 1970 to 2037 with the following exceptions:
 - This determination includes no statement regarding the operating systems or network operating systems that Livelink components run on.
 - This determination includes no statement regarding any third party software that Livelink may use or interact with in any way,
 - Provided user represents dates in the manner required by the Software as set out in the Livelink Software documentation.”

3. Overall Impression

ISRI tested Livelink Version 8.0.4 so the statement published on their website and quoted here may not apply to the version we tested. We assume though that if Livelink is selected for the LSN, the most recent version (and the one to which this statement applies) would be selected. For more information on Livelink’s Y2K compliance, please contact year2000@opentext.com.

- (R) Collection Size.** DOE-LSN must accommodate the anticipated size of 1,000,000+ documents containing 10,000,000+ text pages and images.

REQUIREMENT NOT TESTED

- (R) Internet Accessible.** DOE-LSN must be accessible on the Internet.

1. Livelink satisfies this requirement.
2. Livelink uses the paradigm of a web browser client interacting with a web server as its way of providing a user interface to the system. Since standardized Internet protocols and conventions are used in communications between the client and server, there are virtually no restrictions on where the client can be located on a network to potentially have access to a server running the Livelink system.
3. There are factors that may affect accessibility of the system, all of which are considerations for any Internet communications, not just for Livelink. Here are a few:
 - Internet firewalls. If communications between the client and server must pass through a firewall device, it is possible that communications may be restricted due to security policies.
 - Web proxy servers. Widely used especially by large ISPs, these devices help to reduce bandwidth demands by caching local copies of frequently accessed web pages. Since Livelink is such a complex system, it is possible that minor problems could arise for users who are forced to go through proxy servers.
 - Network bandwidth. The transfer of document images and other types of data will make considerable demands on available network bandwidth. Adequate bandwidth on the server side must be available for Livelink to serve requests from all users who wish to access the system.
 - Network latency. Latency is the apparent delay between the time some data is transmitted from one point on the network to when it is received at another point. With an interactive system such as Livelink, it is important to consider the effect of network latency on the usability of the system. For example, if the average delay is too great between when a user clicks something and a response is received, many users may feel that the system is too frustrating to use.
 - Number of licensed users. To ensure accessibility on the Internet, care must be taken to ensure an adequate number of licenses for the system exist to be able to handle simultaneous requests from many users. System load will need to be carefully monitored and licenses increased, if necessary.

4. Overall Impression

Livelink is fully Internet accessible since this method of access is intrinsic to the system. With the Internet being a fundamental element of Livelink's framework, Internet interfacing concerns are a nonissue. Exclusion of all other forms of access may not necessarily be advantageous.

(R) Windows/Windows NT. DOE-LSN must be usable by clients on Windows and Windows NT operating systems.

1. Livelink meets this requirement. Since the user interface to the Livelink system is through a web browser, and well-known and well-supported web browsers (such as Netscape Communicator and Microsoft Explorer) are available on the Windows/Windows NT platforms, the system is necessarily usable by clients on those platforms.
2. However, there are several requirements to be placed on that usability, given the vendor-supplied user interface to the system:
 - Clients must use recent versions of web browsers. Older browsers may not implement some functionalities that Livelink uses, or their implementations may be buggy. Of particular concern is the browser's support for and correct implementation of the Javascript language. Corruption of or unusability of the user interface to the Livelink system could result due to deficiencies in the browser's Javascript implementation.
 - Windows/Windows NT client machines must have sufficient processing power and memory to be able to run modern versions of the popular web browsers. It is recommended that clients follow the browser vendor's instructions for at least the minimum system requirements. Client browsers will use significant system resources even while simply navigating around Livelink, and especially while retrieving documents and their images.
 - Clients must enable Javascript in their browsers for much of the Livelink user interface to function, particularly navigation aspects. Some users might view this as a security concern. Enabling Javascript may open the client machine up to security holes even in the most recent versions of popular browsers. This could be an unacceptable risk for some users. In addition, browsers as of this writing do not implement the ability to turn Javascript on only when accessing user-specified websites. And most users would find it cumbersome to turn Javascript on when accessing the Livelink user interface and then turn it back off when finished. So the result would be that everyone wanting to access the Livelink system would enable Javascript and leave it enabled all the time. This would further increase the security risk, since now the user's machine is opened up to potential threats from websites all over the Internet, not just those that might exist while accessing the Livelink system.
 - **Overall Impression** Livelink is a completely web-based product so meeting this requirement is not difficult. With this browser dependence comes the inability to control the viewers interface to the system. In particular, Livelink has opted for heavy use of Javascript in their browser interface design which will affect some users' ability to employ all of Livelink's functionality.
Most systems will use the standard web browser interface for client access so these issues are not specific to Livelink.

(B) Platforms/Operating System. DOE-LSN should run on one of the following platforms: Windows NT, Sun Solaris, Alpha Unix.

1. Livelink meets this requirement. Livelink runs on Sun Solaris and Windows NT although Sun Solaris is the only platform evaluated in this review. Specifications for the Sun platform used in this evaluation can be found in the platform description section of the summary report.

2. There do not appear to be any specific issues which affect the performance of Livelink on the selected platform, although the Livelink SDK (Software Development Kit) is not available for platforms other than Windows NT. We did not evaluate the SDK in this review.
3. Livelink installation was aided by Opentext technical staff but the *Installation Guide* seems very complete. System Installation without vendor technical support is achievable.

4. Overall Impression

The installation was generally very smooth. There was a problem with the Oracle installation that surfaced after Livelink was installed but Livelink staff was able to diagnose and solve it quickly. Reinstallation of Oracle after Livelink to correct the problem did not have any adverse effect on the Livelink installation.

(R) Concurrent Users. DOE-LSN shall support up to 150 concurrent users. [LSS2-064]

REQUIREMENT NOT TESTED

3.2 Querying Requirements [LSS2-011]

(R) Query for Document. The DOE-LSN shall provide the capability to query the system for a list of all documents that meet the query criteria and sort the displayed list on the basis of selected displayed fields or relevancy to the query. [LSS2-011]

1. Livelink returns a list of documents that meet the query criteria.
2. *Query for Document* is an integral part of the Livelink interface. There are three different ways a user can enter queries: They can select *search* from the “Go to” pull down menu and 1) enter a quick search natural language query; 2) select advanced search and use the Livelink advanced query language; 3) enter a quick search query directly from the current web page. Figures 2 and 3 show each of these interfaces. If a user does not specify how the query result list is to be presented, then the results are sorted in the following order:
 - (a) Relevancy: The relevance of the document to the query.
 - (b) Type: The MIME type of an item.
 - (c) Name: The name of an item.
 - (d) Date: The date an item was last modified.
 - (e) Size: The size of an item in kilobytes (KB).
 - (f) Location: The Livelink Workspace or higher-level item containing an item of interest.

where 2b through 2f are system parameters, not header fields. The user may adjust which field is to be given higher precedence when sorting, but is not able to add displayable *regions* for the documents in the collection and is not able to sort on them. The user is given a choice to sort in either ascending or descending order on the 6 aforementioned regions as well as select a desired weighting scheme.

3. These capabilities are well documented and searchable online in Livelink.

4. Overall Impression

Livelink definitely has the capability to return a list of relevant documents that satisfy a given query. Furthermore, Livelink is able to sort the returned list on some of the more commonly desired fields such as name, date, and size and does so on relevance by default. With the tools provided, we were unable to reconfigure Livelink’s interface to allow the display and sorting of header fields.

Note: Sort order is assigned to a particular query rather than as a user preference. This makes it necessary to reset the *Display Options* each time a new query is run to provide the selected display order.

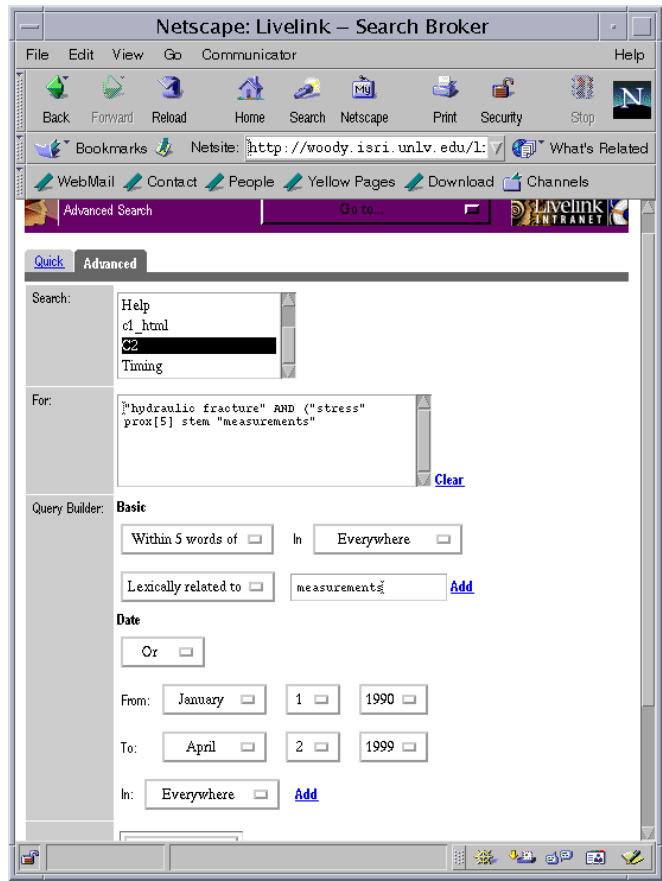


Figure 2: Advanced query interface in Livelink

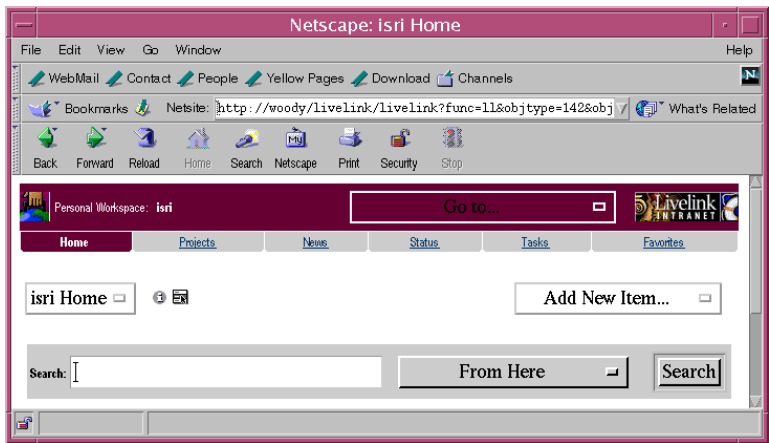


Figure 3: Quick search interface in Livelink

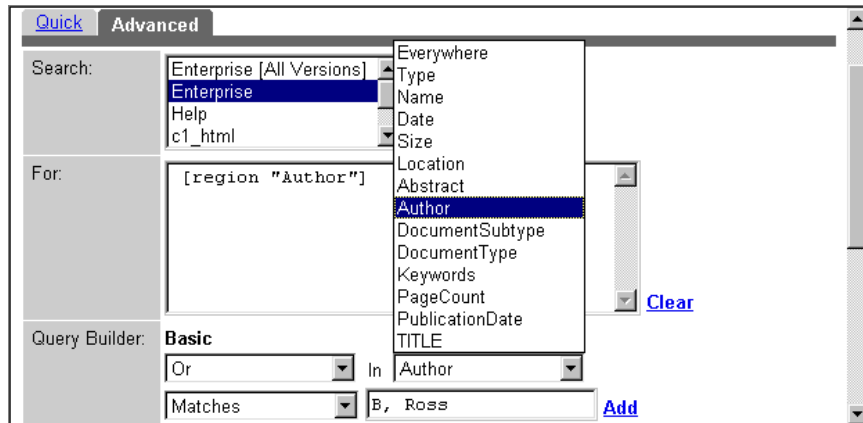


Figure 4: Header search interface

- (R) **Query Header.** The DOE-LSN shall provide the capability to query the system by specifying the content of one or more header fields to obtain a list of all documents that satisfy the query. [LSS2-011-1]

Comment: The search will be able to sort appropriate fields, such as date, accession number, etc. in ascending or descending order. It is anticipated that the DOE-LSN will allow the user to select and search multiple bibliographic header fields.

1. Livelink has the ability to return relevant documents when queried on specific header fields or on combinations of header fields. But since these regions are not displayable, display options, such as sorting, are a moot point. We were unable to reconfigure Livelink's interface to allow the display and sorting capabilities for header record fields.
2. After some investigation, we determined that the most natural way to include header fields and associate them with a document was to include them as HTML-tagged regions as shown in Figure 1. This use of regions makes the ability to "Query Header" an integral part of the Livelink system. Figure 4 shows the header query interface.
3. Inclusion of header fields within documents was implemented by ISRI (see Section 2). It is possible though that header inclusion could have been developed using Livelink's SDK if it were available for Solaris.
4. The user should be provided with some information of data format (e.g., Author field: last name, first initial) since the header fields are relational and must match exactly to retrieve requested documents.

5. Overall Impression

In order to handle header field searching, ISRI was forced to reverse-engineer the Livelink indexing procedure and enter tags into the HTML file (thus creating an out-of-spec HTML document). The correct way would have been to represent metadata with the HTML <META> tag, but it is not clear that it would be possible to do this with Livelink. ISRI received virtually no support from Opentext during this process.

Searching of header fields seems to work quite well, but this is little consolation if the searched fields are not displayable. One inelegant option would be to include the header fields in the body of the text, but this would still not solve the issue of being able to sort using these fields.

Note: Field names are case sensitive, so if the user enters a query directly into the query box, he must note the case of each field name.

- (R) **Query Text.** The DOE-LSN shall 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 that satisfy the query. [LSS2-011-2]

Comment: Describe any query optimization techniques used by your system.

1. Livelink provides the capability to specify character strings to search the full text of the document collection. Text searching is an integral part of the Livelink system.
2. Text searching can be accessed in three different ways in Livelink by: 1) entering a quick search natural language query; 2) selecting advanced search and using the Livelink advanced query language; 3) entering a quick search query directly on the current web page.
3. Text searching capabilities are well-documented online.

4. Overall Impression

The Livelink advanced query language is comprehensive with respect to its *full text* searching capabilities. It provides several *operators* and *modifiers* to construct complex queries. These will be listed and described in detail in the “Text Query Parameters” requirement in 3.2. A query builder is available to aid the user in the advanced query language.

No optimization techniques were disclosed from Opentext support.

- (B) Text Query Parameters.** The DOE-LSN shall provide the capability to specify single and multiple character wildcards, to utilize proximity searching, and root searching as part of a full-text query and to combine multiple result sets. [LSS2-011-3]

Comment: The intent of this requirement is to allow users to combine any of the previously executed searches that were performed during the same search and retrieval connection.

1. Livelink provides the query parameters listed, standard operators such as AND, OR, BUT NOT and has several more operators and modifiers (described in 1). Livelink does not however, allow the capability to combine query results sets. The operators and modifiers are an integral part of the advanced search language.
2. Livelink provides a point-and-click query builder interface that allows simplified construction of complex queries. Figure 2 shows the advanced search page.

3. Overall Impression

Livelink provides an extensive set of operators and modifiers for searching the full text of the documents in the LSN. As in any query language, a good understanding of both the syntax and the semantics of the operators is required to construct meaningful expressions. Livelink does provide a “quick search” interface as well for users who are unfamiliar or are uncomfortable with the advanced query language. Queries or results can be saved in Livelink, but there is no method for combining queries.

- (R) Query Header and Text.** The DOE-LSN shall provide the capability to query the system by specifying a combination of header field values and the text query parameters from the full text of the document to obtain a list of all documents that satisfy the query. [LSS2-011-4]

1. Livelink does have the ability to combine searches of header fields with parameters of the full text of the document. No obvious restrictions or problems were encountered when combining these types of parameters and any possible combination seems allowable.
2. Querying both the header and the text of a document can only be done on the *advanced search page* in Livelink but Livelink provides an easy to use interface for including header field searching into the user’s query.
3. This capability is obvious to the user and system regions are documented online. Customized regions, like the LSN header record fields, are not documented in the system and the inclusion of these were not obvious at installation time.

4. Overall Impression

Operator/Modifier in Menu	Keyword	Effect
Synonym or	SOR	same as OR, except that in addition the phrases to the left and right are treated as synonyms of each other. Matching scores may differ affecting document rank
Exclusive or	XOR	results in finding all items containing exactly one of the two specified phrases, not both
Within n	PROX[n]	results in finding all items containing both specified phrases within n words of each other
Lexically related to	stem	broadens the query to include words derived from the main part of the specified term
Sounds like	phonetic	broadens the query to include words that sound like the specified term
Synonyms of	thesaurus	broadens the query to include words from the thesaurus entry for the specified term
Begins with	right-truncation	broadens the query to include words whose beginning matches the specified term
Ends with	left-truncation	broadens the query to include words whose ending matches the specified term

Table 1: Operators and Modifiers in Livelink

Livelink provides the ability to search both the full text and header fields in a single query. The problems associated with displaying the results of header searches can be found in the response to the **Query Header** requirement in Requirement 3.2.

Note: When a user wishes to query the full text of a document he/she is inadvertently querying all regions. That is, the default region is set to “Everywhere.” It is possible to query just the full text region of the document collection however, since Livelink supplies a negation operator that excludes specified regions from consideration.

- (R) **Provide Query Status.** The DOE-LSN shall provide the user an indication of the query status during a query and allow the user to terminate queries in process without terminating the session or losing previous result sets. [LSS2-011-5]

Comment: It is always possible to construct a query so broad that it results in an unmanageable results list. Users should be able to determine that an ongoing query is too-broad and terminate the query in process. An indication that the session is still connected and that the query is working is adequate.

1. Livelink does not provide any specific query status indicator while the system is searching other than what’s displayed by the browser. The browser indicator is specific to the browser (i.e. Active “N” icon in Netscape Navigator, and spinning “e” icon in Internet Explorer) while the request is in progress, the status line at the bottom of the browser gives information about page download in bytes retrieved.

2. The download of any page transfer can be terminated by selecting the “stop” icon on the browser, but again, this is a browser feature and may differ from browser to browser. This does not interrupt the current session and does not affect any *previously saved* result sets.
3. If the user is familiar with the browser in use, then the browser indicator and the “stop” icon can aid in query status and query termination.
4. Since these are browser features rather than an integral part of Livelink, no documentation is provided.
5. **Overall Impression**

Since Livelink is a web-based system, the only interface to the DOE-LSN collection will be a web browser. Most users will be familiar with their browser and will know how to observe and use these features. The user can easily determine whether the query is still in progress or stalled, but there is no other query status information.

Note: Livelink does not provide the total number of documents retrieved.

(B) Query Assistance. The DOE-LSN shall 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. [LSS2-011-6]

Comment: Examples might include synonym processing, thesaurus, natural language queries, or other search aids. Because a variety of approaches are used in the commercial market, no one approach is specified.

1. Livelink provides a natural language interface on its “Quick Search” page. All other listed search aids are a part of Livelink’s “Advanced Search” query language. These capabilities are an integral part of the Livelink system.
2. A natural language query can be submitted from almost any page in Livelink by entering the query in the search box as shown in Figure 3. The other features can be included into the user’s search on the “Advanced Search” page with selections from drop down menus as shown in Figure 2. These capabilities are easy to apply.
3. These capabilities are briefly described online but the documentation could have been more developed.
4. Livelink meets this requirement in that it gives the user the capacity to apply query assistance parameters to his query. There are some shortcomings of the query assistance parameters that are addressed below:

Thesaurus: Livelink uses the Inso thesaurus technology (version if231e) and the Houghton Mifflin Dictionary (version ETMR00512). This thesaurus is not viewable by the user making it impossible to know if the expansion being applied is what the user had in mind. Under the current Livelink configuration, it is not possible to utilize a different thesaurus. It might be possible in the future, but not now.

Feedback: Called *similarity searching* in Livelink, it can be applied to either a pre-run query or selected text. Unfortunately, the user is unable to select the documents the system uses as feedback; Livelink selects the top ten words from the top ten documents for the default feedback terms. The numbers can be changed by the user, but the method is fixed: A user can only change the number of documents (to say 20) and the number of words selected from those documents (to say 15). Livelink would then take the top 15 words from the top 20 documents. The user is unable to tell the system which documents he believes are relevant to his query.

Natural Language: As stated, natural language queries can be run from the “Quick Search” page. Advanced queries can also be entered here although with no interface assistance. Natural language queries cannot be run with dependable results from the “Advanced Search” page.

Sounds Like: broadens the query to include words that sound like the specified term. Useful for phonetically misspelled words but doesn't show much promise for OCR errors.

5. Overall Impression

Livelihood meets this requirement. Livelihood though does not allow the user to apply these search aids to their fullest:

- A thesaurus is helpful if it can be browsed and the relationship between a term and its synonym is known. Also some collections, like the LSN, have their own nomenclature. This may require modification or augmentation of the default thesaurus which is currently not possible in Livelihood.
- Feedback should incorporate the sense of the query writer. He should have the ability to tell the system what documents are most related to his query. Livelihood does not allow this ability in its "find similar" mechanism.
- The "sounds like" approximation doesn't encompass OCR errors.

3.3 Display Capabilities [LSS2-012]

NOTE: All display capabilities in Livelihood are dependent on the user's browser. These requirements were tested in several versions of both Netscape and Microsoft Explorer but exceptions may occur.

(R) Display Document. The DOE-LSN shall provide the capability to display a document. [LSS2-012]

1. Livelihood meets only a part of this requirement. A document can be selected for display only after a query is run. From the returned results list, a document can be selected and displayed. Livelihood returns a web-page based document with the ability to view document text within the user's browser; images are viewable if the user has an appropriate viewer with capabilities to display TIFF images, (see Requirement 3.3) installed on their computer. The ability to view the header fields of a document is not provided.
2. Displaying the header fields, the document images, and the document text as a single unit is not an integral part of Livelihood.
3. The concept of a DOE-LSN document in Livelihood was integrated into the HTML pages as described in 3.2. Although not difficult to implement, it may not be the most natural representation.
4. **Overall Impression** Livelihood does not meet all the expectations of this requirement first, since the header fields cannot be displayed and second, because the association between an image and text page was implemented using HTML image tags embedded in the document text. See Section 2 for more information on Livelihood collection information.

(R) Display Header. The DOE-LSN shall provide the capability to display the header of a document. [LSS2-012-1]

1. Livelihood does not meet this requirement. Although we were able to incorporate the header fields as regions into the *head* of the HTML document, the header fields are not displayable. Some default system regions in Livelihood can be selected for display, but user-defined regions cannot. This requirement is not an integral part of the Livelihood system.
2. **Overall Impression**
The Display Header requirement is not met in Livelihood; the notion of user-defined document meta data is evidently not a concept that was considered by the system designers.

(R) Display Text. The DOE-LSN shall provide the capability to display a page of text of a document. [LSS2-012-2]

Text Format: The text representation of material in DOE-LSN shall be page-delimited ASCII text. [LSS2-056]

1. Livelink can display document text through the use of a browser. The notion of page-delimited text is defined by the documents original format and needs to be implemented at the time of document loading.
2. Once the method of loading is well defined, viewing text pages of a document is obvious. Moving from page-to-page within a document is accomplished with HTML tagging included by ISRI staff.

3. Overall Impression

Displaying ASCII text is an integral part of the Livelink system, but incorporating single pages into a complete document while keeping these pages distinct is not. HTML though is flexible enough to enhance documents to improve display characteristics.

- (R) Locate Search Terms in Document.** The DOE-LSN shall provide the capability to locate the terms in the document text that satisfy a full-text query and to move from one term to the next or previous term without displaying intermediate text. [LSS2-012-3]

Comment: This function is performed as the user is viewing the document. It is typically implemented by highlighting the search terms in the document and providing a "go to next term" function that places a cursor at the line or word of the search term.

1. Livelink does not directly support this requirement. The only way to highlight search terms is by using the browser's find command. This is hardly acceptable because each search term in the query would have to be searched separately by the user giving little continuity to the document's relevance.
2. This capability is not an integral part of the system but Livelink user support claims term highlighting may be incorporated into the next release of Livelink.
3. **Overall Impression** Locating search terms in returned documents is important feedback to the user. It gives them a clear understanding of why a document is returned by the retrieval system. By not highlighting search terms, Livelink degrades the usefulness of the system even if the relevance of the documents returned is high.

- (R) Display Image.** The DOE-LSN shall provide the capability to display-the images of a document, page by page, including full page views of the images-of 8-1/2 by 11 inch pages up to E-size pages. [LSS2-012-4]

Image Formats:The electronic image of documentary material in DOE-LSN shall use *Aldus Tagged Image File Format* (TIFF) Group 4 for bitonal images and *Joint Photographic Experts Group* (JPEG) for color and grey scale images. These formats are part of the Adobe TIFF I Version 6.0 representation. Adobe TIFF is an industry standard developed and put into the public domain by Adobe. [LSS2-057]

1. Livelink does not itself support this requirement. Since Livelink is browser-based, this capability is dependent on the specific web browser being used. This feature most often comes as a plugin. Later releases of Internet Explorer and Netscape Navigator for Windows 95/98/NT and Netscape for UNIX come equipped with the necessary viewing plugin. Earlier versions may require the user to download the appropriate plugin.
2. Image display is not an integral part of Livelink. The ability to view a document image requires the applicable HTML links to be added to the document and a separate plugin for image viewing.
3. If the user is familiar with web browsing then viewing a document image is apparent.

4. Overall Impression

Image display is a capability of the browser and plugin, not Livelink itself. It is also dependent on the way the collection is generated for Livelink.¹ Since Livelink is browser-based, the user will have the ability to view an image using tools outside of the Livelink system.

(R) Image Viewing. The DOE-LSN shall provide image viewing capabilities for image enlargement, reduction, scrolling, and rotation. [LSS2-012-5]

1. Livelink does not itself support this requirement. Since Livelink is browser-based, these capabilities are dependent on the specific web browser and plugin being used.
2. The listed requirements are in particular plugin specific. The user must procure a plugin that provides these abilities. Both Internet Explorer and Netscape Navigator for Windows 95/98/NT come with an appropriate plugin for image enlargement, reduction, scrolling, and rotation. On the UNIX systems, in most cases, a plugin must be downloaded that supports image enlargement, reduction, and rotation.
3. **Overall Impression** Again, Image Viewing is not directly supported by Livelink. Because it is browser-based and because plugins can easily be acquired, the user will have the ability to manipulate an image using tools outside of the Livelink system.

(R) Display Image and Text. The DOE-LSN shall provide the capability to concurrently display an image page of a document and its associated text. [LSS2-012-6]

Comment: There must be a one-to-one correspondence between each page of text and its corresponding page image. This assumes each page will be tagged in the text version.

1. Displaying an image page and its corresponding text page is achievable in Livelink but is dependent on collection preparation (see Section 2) and use of an appropriate plugin.
2. Displaying image and text simultaneously is not an integral part of the Livelink system.² collection preparation is an important stage for implementing image/text correspondence and for including this ability into the document. Also, since Livelink doesn't directly provide this capability, what the user actually views is not controlled by Livelink. The browser window displays the text of a page while the designated plugin displays the page image.
3. Ease of using this capability depends heavily on initial collection preparation (refer to Section 2) and the plugin being used.

4. Overall Impression

For Livelink to meet this requirement, the collection needs to be prepared with the objective of page by page display and text/image correspondence.

(R) Viewing Options. The DOE-LSN shall allow the user to view the following combinations: 1) header, 2) image, 3) text, 4) header and text, 5) header and image, and 6) text and image. [LSS2-012-7]

Comment: The user must be able to look at text and headers without looking at images.

1. Livelink does not support all of the listed combinations. The most obvious drawback is Livelink's inability to display the header fields.
2. Image viewing is dependent on collection preparation and plugins; text is presented in the browser; the combination of these two are described completely in Section 2.
3. **Overall Impression** Due to Livelink's inability to support header fields display, this requirement is not met. There may be ways to modify the interface to include header field display, but it is not apparent from Livelink documentation or from responses received from technical support. Also, display of any kind will always be dependent on the browser and plugin in use. This may or may not be suitable for the DOE-LSN.

¹ See Section 2 to view the HTML generated to include an image into a text page of a document.

3.4 Printing Requirements [LSS2-013]

NOTE: All printing capabilities in Livelink are dependent on the user's browser. These requirements were tested in several versions of both Netscape and Microsoft Explorer but exceptions may occur.

(R) Print Document. The DOE-LSN shall provide the capability to print a document at a local printer. [LSS2-013]

Comment: It is assumed that the local printer is capable of printing the requested document type.

1. Livelink, through the use of a web browser, gives the user the ability to print what is presented in the browser window (an HTML document) to any default printer.
2. Exhibit A shows a printed document from the Livelink system. It includes all document pages, image links, and text links. Other document components, like image pages, would have to be printed separately.
3. If the user is familiar with their web browser, printing a document is available and simple to use.
4. **Overall Impression** Livelink depends completely on the user's browser capabilities to print a document from the DOE-LSN collection. What the user receives in printed form will depend on the collection's preparation in HTML. See Section 2 for more details.

(R) Print Header. The DOE-LSN shall provide the capability to print a document header at a local printer. [LSS2-013-1]

1. Livelink does not meet this requirement since there is no method to display or access the header fields.
2. The only interface given to the user of the DOE-LSN collection is what is provided by the browser. This makes collection preparation crucial for the display and the printing of document components.
3. **Overall Impression** Printing the header fields, like displaying the header fields, is not realized in the Livelink system because the notion of a "header field" associated with each document is not inherent in Livelink. These requirements could possibly be met by changing the HTML document structure or by modifying the Livelink interface.

(R) Print Text. The DOE-LSN shall provide a user selectable capability to print from one page to all of the text of a document, and any selected ranges of pages, at a local printer. [LSS2-013-2]

Comment: The system must be able to discern pages within a document for printing.

1. Livelink depends on the browser's printing capabilities to print the text of a document. The browser can print an "HTML page," which in this case is the entire document; a range of pages can be printed in some browsers but they will not be equivalent to the pages in the document.
2. Livelink can print all the text of a document, but there is no concept of "pages" that corresponds directly with the printed pages of an LSN document.
3. Printing document text is an integral part of the browser interface; printing one page to a selected range of pages within a document is not.
4. **Overall Impression**
Livelink's dependence on the user's browser governs document text printing. The "HTML page" is what gets printed, which in this case, is the complete document. The document structure could be modified to adhere to this requirement but not without affecting other document access issues.

(B) Report Generation. The DOE-LSN should provide report generation capabilities for several of the above listed tasks.

1. Livelink meets this requirement with two specific report generating facilities: 1) Change Agents; 2) LiveReports.
2. Change Agents are reports that are created and by Livelink users to track changes to the collection based on user specification. LiveReports are created by the Livelink administrator and disseminate information to users about changes in the system or the collection.

3. Overall Impression

Report Generation can be defined in several ways. With these two reporting facilities Livelink provides a method for users to request information about the collection based on user-defined criteria. The system administrator is also given the capability to query the underlying RDBMS to track statistical information about database contents. These kinds of reports provide what traditional report generation would provide but can be tailored to individual requests.

(R) Print Standard Image. The DOE-LSN shall provide a user selectable capability to print from one to all images, and any selected ranges of images, of 8-1/2 by 11-inch (or smaller) pages of a document, at a local printer, reduced to a single 8-1/2 by 11-inch paper. This includes the capability of printing an oversized page image , up to E-sized, on a single 8-1/2 by 11-inch sheet of paper. [LSS2-013-3]

1. Livelink depends on the browsers capabilities as well as an appropriate plugin to satisfy this requirement. This capability is not an integral part of the Livelink system.
2. Printing images, in particular TIFF images, may not be an inherent component of the default browser in use. In many cases, the user will need to download a plugin before they can view or print a standard image.
3. Most viewer plugins have the ability to print the current image to a local printer. Viewers *usually* include the ability to reduce, enlarge, and rotate an image page, so these capabilities should be met in general.
4. The ability to print a *selected range* or *all* page images of a document is usually not a feature of viewer plugins and was not built into the structure of the HTML documents in our test collection. Apart from modifying the Livelink interface, a plugin viewer may need to be adapted as well. Some modification to the HTML document structure may also be required.

5. Overall Impression

Adding this feature to the Livelink interface is only one consideration of this requirement. Since Livelink's browser interface depends on a separate plugin, modification to this plugin must also be considered. Livelink's closed environment and dependence on outside software make this requirement more difficult to realize.

(R) Print Oversized Image. The DOE-LSN shall provide the capability to print an oversized page image, up to E-sized, on a single sheet of paper at 100 percent of the size of the original image. [LSS2-013-4]

1. Livelink depends on the browsers capabilities as well as the appropriate plugin to satisfy this requirement. This capability is not an integral part of the Livelink system.
2. Printing oversized images may not be an inherent component of the default viewer plugin.
3. Many viewer plugins do not have the capability to display or print oversized images even if the default printer has this ability.

4. Overall Impression

Since Livelink's browser interface depends on the plugin's capabilities, the selection and recommendation of the plugin to LSN users is an important consideration. See Section 2 for more information.

(R) Print Results List. DOE-LSN shall provide the capability to print some or all of the summary lines of a results list. [LSS2-013-5]

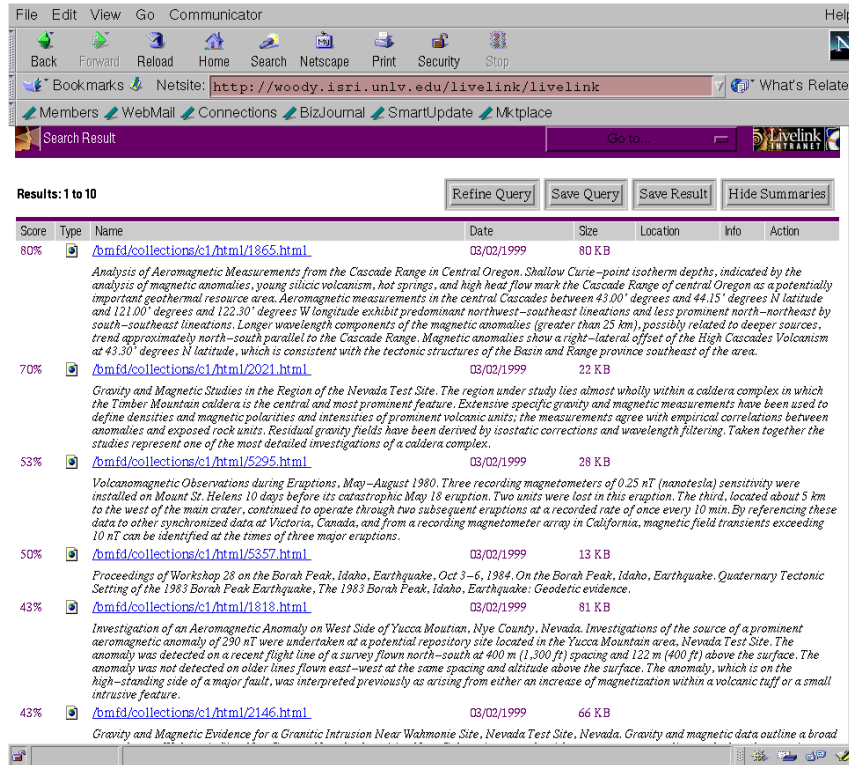


Figure 5: Livelink query results list page

1. Using the print capabilities of the browser, Livelink can print a page of results which includes the document title (directory path and file name) and a short summary for each document (see Figure 5).
2. The number of documents per page (with restrictions) and whether summaries are displayed and printed can be selected by the user. Each page of results must be printed separately to print all documents returned in the results list.
3. The total number of documents returned for a query is not disclosed making it difficult to determine if the query submitted is too broad.
4. Printing the content of any browser page is straight forward as long as the user is familiar with the browser in use.

5. Overall Impression

Livelink's dependence on the user's browser governs results list printing. The "HTML page" is what gets printed, which in this case, is the number of "hits" returned per page. The only control the user has over results list printing is in the "Display Options" selections which includes: 1) Number of Results 2) Show/Hide Summaries.

Note: The "Display Options" settings are tied to a query rather than to a user, making it necessary to reselect these options for each new query entered.

(R) Print Screen. DOE-LSN shall provide the capability of printing the screen display. [LSS2-013-6]

1. Using the print capabilities of the browser, Livelink can print the "HTML page" displayed. This may or may not define the "screen display." If the "screen" is defined as just the viewable area within the browser, printing just this portion is not possible unless the full HTML page is in view.
2. Printing the content of any browser page is straight forward as long as the user is familiar with the browser in use.

3. Overall Impression

Printing the screen display is a keyboard function that is rarely used in graphical user interface (GUI) environments commonly used today. This requirement seems to be a carry-over from line-based interfaces and is probably not a strong requirement.

(R) Request Paper Copy. DOE-LSN shall provide the capability to submit an electronic request for a paper copy of the header, images, or text of a document or of an entire results set, including oversized and color images. [LSS2-014]

1. “Request Paper Copy” would typically not be a standard feature of any search system. Evaluation of this requirement is based on the simplicity with which this feature could be added seamlessly.
2. “Request Paper Copy” could possibly be met by “embedding” this ability within the HTML document structure but would probably be more serviceable as a separate feature of the Livelink interface.
3. It is possible that this requirement could be added to Livelink using the the Livelink SDK but since Livelink does not currently support their SDK for Sun Solaris we were unable to implement it in this evaluation.

4. Overall Impression

“Request Paper Copy” is not a typical out-of-the-box feature of any search engine but with a flexible interface (especially a browser interface), implementation should not be difficult. Short of OpenText’s professional services, there was no simple way on the Sun Solaris platform to include this feature into Livelink.

(R) Process Paper Copy Requests. DOE-LSN shall provide the capability to receive and read an electronic request for a paper copy of a document and print the requested copy.

Comment: This is not anticipated to be a highly automated function. The requested body must be able to receive requests and print out the requested document. The rest of this function may be procedurally implemented.

1. “Process Paper Copy Requests” is the receiving end of “Request Paper Copy” and would not be a standard feature of any search system. Evaluation of this requirement is based on the simplicity with which this feature could be added seamlessly to Livelink.
2. This requirement should be incorporated into the administrative component of Livelink.
3. It is possible that this requirement could be added to Livelink using the the Livelink SDK but since Livelink does not currently support their SDK for Sun Solaris we were unable to implement or test it for this evaluation.

4. Overall Impression

“Process Paper Copy Requests” is not a typical out-of-the-box feature of any search engine but with a flexible interface (especially a browser interface), implementation should not be difficult. Short of OpenText’s professional services, there was no simple way on the Sun Solaris platform to include this feature into Livelink.

3.5 System Administration Requirements

(R) Monitor System Status. DOE-LSN shall provide authorized users the capability to monitor the status of the system and communication components and to interrupt, restrict, or disable capabilities in order to optimize use of system resources. [LSS2-033]

1. This requirement is marginally satisfied by the Livelink product.
2. The Livelink system allows the *administrative user* to monitor the indexing data flow for all collections in the system. An authorized user can schedule, enable, or disable the File Walker, HTML conversion, and Indexing processes. This allows the Livelink

administrator to schedule the processes such that system resources are used efficiently. Additionally, the search capability for each collection in the Livelink system can be independently enabled or disabled.

3. The Livelink retrieval system does not provide extensive system monitoring capabilities. It provides limited feedback as to which components of the system are currently running, scheduled to run, or disabled.
4. Livelink does not provide information about the percentage of system resources being utilized by each component.
5. **Overall Impression**
The Livelink system allows the administrative user to control system activities but does not provide extensive monitoring of system resources.

(R) Monitor Session Activity. DOE-LSN shall provide the capability for an authorized user to monitor user session activity levels and identify and cancel queries or other system activities. [LSS2-033-1]

1. This requirement is not satisfied by the Livelink system.
2. Since the Livelink document management product is a web-based system, it does not provide session level auditing of user activity, nor does it provide the ability to cancel active queries.
3. These features could be implemented through the web server capabilities, but would require a non-trivial level of effort to extend the Livelink interface.
4. The Livelink product does provide an automated Change Agent that notifies the administrator via email about changes in the system status. The information available in the Change Agent report is of minimal use for tuning system performance. These reports only contain information about objects created by users.
5. **Overall Impression**
The only user level control available to the Livelink administrator is the ability to create and delete users, disable and enable user login, and the ability to grant permissions to users. There is no concept of “session control” granted to the Livelink administrator.

(R) Database Administration Tools. DOE-LSN shall provide authorized users the capability to assess the availability, integrity, and performance of the databases of the the DOE-LSN, including those pertaining to the storage of document header fields, text, and image data, and adjust database performance parameters or restrict or disable database features in order to optimize system performance.

1. This requirement is not satisfied by the Livelink system.
2. The Livelink system allows the administrator to determine the availability of the databases in the system. It does not, however, provide the ability to determine the integrity of the databases, or adjust performance parameters for optimizing the overall system performance.
3. Livelink provides the administrator with a limited ability to manage the Livelink tablespaces in the underlying DBMS.
4. **Overall Impression**
Livelink is not designed as a database repository for document header fields, text, and images. There is no concept of storing these components in databases, rather, these components are left in their original files in their original formats. The index of the collection is stored in Livelink but no information about the index is available to the Livelink administrator.

3.6 Internet Requirements

(R) Web Server Interface. DOE-LSN must interface with a Web Server for querying the system and returning query results.

1. The Livelink system satisfies this requirement.
2. The Livelink system requires the use of a web server in order to function. Interacting with the Livelink system through the use of a web browser on the client side and the web server on the server side is the only supported method of interacting with the system.
3. No command line interface or other user interface exists. All user connections with Livelink go from the client's browser exclusively to the web server—never to any Livelink server software directly.
4. **Overall Impression** The *only* method of accessing Livelink is through its web server. Although Livelink fulfills this requirement, exclusion of all other methods of access may not be advantageous.

(B) CGI and Perl5. DOE-LSN should use the CGI Standard and be accessible from the Perl5 programming language.

1. Livelink uses the CGI standard exclusively in servicing requests from client browsers. Opentext has stated that, “all interaction with Livelink is through HTTP GETs and POSTs.”[1]
2. Opentext has further indicated that there are at least two ways for their customers to build their own custom interface to Livelink:
 - By writing HTML forms using existing Livelink pages as examples and make requests of the Livelink software run by the web server using HTTP GET and POST operations. This method would be somewhat limited since the client would have little or no control over the format of the page and data that is returned from the server, but nevertheless this might be sufficient if the client is a small program designed to accept information, parse the results, and process it.
 - By building a custom interface to Livelink using the Livelink API which the company states is currently available in C/C++, Java, and Visual Basic. The company did not indicate it was possible to use the Perl 5 language with the API. They state that the API is available “for platforms that support those languages.”[1]
3. Although the Livelink API was not part of this evaluation, the company seems to indicate that, using it, it is possible to build a completely customized interface to Livelink. Finally, the company says:

There are no real limitations to an interface written using CGI, the API's or HTML except for the fact that you would lose the benefit of our work on the interface—which may or may not be a problem.[1]
4. **Overall Impression** Livelink provides the CGI standard for making some customization and enhancements to its basic interface. The API is currently not available for the Sun Solaris platform and therefore, we were unable to evaluate its potential.

(B) ODBC/JDBC Compatibility. DOE-LSN should be accessible using Open Database Connectivity (ODBC) and Java Database Connectivity (JDBC).

1. Livelink is itself not ODBC/JDBC compatible. Opentext states that, “for performance, reliability and scalability reasons there is no ODBC interface to Livelink.”[1]
2. Livelink uses Oracle to store some types of information. Depending on the structure of the document collection and how it was loaded, some or all of the information in that collection may actually be stored in an Oracle database, not within Livelink. It is possible to directly access an Oracle database through ODBC and JDBC and run queries on data stored there.
3. Livelink states that it is possible to “write a Java application that incorporates our Java API and JDBC to access the RDBMS directly.”[1] Again, this capability would depend on the Oracle database server and related ODBC and JDBC drivers.

Query	Average response time (seconds)
INJD-T3-Q1	8.5
TEJA-T3-Q2	8

Table 2: Timing for Retrieval of Results List

4. Overall Impression

Since ODBC and JDBC are standard ways of accessing relational databases and since Livelink is not a relational database itself, this requirement does not apply directly to Livelink. The selection of the underlying RDBMS will determine ODBC/JDBC compatibility.

3.7 Timing Requirements

(R) Timing Strings. The DOE-LSN shall meet the average response times shown in Table 4. The performance shall be achieved with 15 concurrent DOE-LSN users active on the system. [LSS-065]

The Livelink system cannot be evaluated at this time under the minimum required load of 15 users, or with the required 5 million pages of document data. However, the minimum timing requirements were analyzed with the smaller test collection of about 50,000 pages of document data.

These tests were all performed on a remote Windows NT 4.0 client machine. The client PC is a 450Mhz Pentium II running Netscape Communicator 4.5 to connect to the Livelink server. The machines are connected via a 10Mbit LAN. Since we are not testing Livelink under operational conditions, the load on the server is relatively low compared to its capabilities.

1. Retrieval of query results list. LSS2-065-2
 - The DOE-LSN requires that the query results list be retrieved in 45 seconds for UNLV test queries INJD-T3-Q1 and TEJA-T3-Q2. Each query was made five times, the time to retrieve each result list was measured, and the average computed for each query over all trials computed. Table 2 summarizes the Livelink average response time for each queries.
 - **Overall Impression** These average times are for the retrieval of the first ten relevant documents according to the Livelink system. Given the size of the current collection and the light load on the system, it is unknown if Livelink will satisfy this requirement.
2. Retrieval of header fields for document identified in query results list. LSS2-065-3
 - The Livelink system does not allow for retrieval of header fields information for documents returned in a query results list.
3. Retrieval of text data for document identified in results list. LSS2-065-4
 - The LSN requires that the first page of text be retrieved in five seconds, and each subsequent page in one second. Under our implementation of the LSS prototype collection in the Livelink system, when a document is selected from a query results list, the entire document is returned. Note that thumbnail images for each page in the document are returned as well as the corresponding text. Also, the text for the page is returned and displayed almost instantaneously while the thumbnails continue loading.
 - From a sample of 10 documents retrieved from a query results list, the time to retrieve the entire document from the Livelink server was measured. These results are shown in Table 3.
 - **Overall Impression** From Table 3 it is shown that the average retrieval time per page is approximately one second. From the results of this experiment, it is believed that Livelink will satisfy this timing requirement.

Document	Page Count	Retrieval time (seconds)
1	3	2
2	2	1.5
3	1	1.5
4	2	1.5
5	2	4
6	2	4
7	16	13
8	14	18
9	7	4
10	4	4
Total:	53	53.5

Table 3: Timing for Retrieval of Document Text

4. Retrieval of image data for documents identified in results list. LSS2-065-5

- The DOE-LSN requires the first page image to be retrieved in ten seconds, with each subsequent image retrieved in two seconds.
- From a sample of ten document pages, each page image was retrieved and displayed in approximately seven seconds (on average). It is believed that much of this time is wasted due to network traffic. On the local machine running the Livelink system, each page is returned in under one second (on average).

- **Overall Impression**

On a network designated for LSN traffic, it is believed that the Livelink system will satisfy the requirement for the retrieval of the first document page image. It is unknown if the two second retrieval time requirement can be satisfied under operational conditions.

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-T3-Q2* Database contains headers for at least 5 million pages of documents. A total of 10 documents found.	45 seconds/70 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.	5 seconds/8 seconds
LSS2-065-4	Retrieval of text data for document identified in query results list.	Database contains at least 5 million pages of documents.	First page: 5 seconds/8 seconds Each subsequent page: 1 second at Main Facility 2 seconds at supported sites
LSS2-065-5	Retrieval of image data for documents identified in query results list.	Database contains at least 5 million pages of documents.	First page: 10 seconds/15 seconds Each subsequent page: 2 seconds at site 3 seconds at other sites

Table 4: Response Time Requirements

EXHIBIT A:
Printed Livelink Document