Model Formulation ■

Virtual Shelves in a Digital Library: A Framework for Access to Networked Information Sources

TIMOTHY B. PATRICK, PhD, GORDON K. SPRINGER, PhD, JOYCE A. MITCHELL, PhD, MARY ELLEN SIEVERT, PhD

Abstract **Objective:** Develop a framework for collections-based access to networked information sources that addresses the problem of location-dependent access to information sources.

Design: This framework uses a metaphor of a virtual shelf. A virtual shelf is a general-purpose server that is dedicated to a particular information subject class. The identifier of one of these servers identifies its subject class. Location-independent call numbers are assigned to information sources. Call numbers are based on standard vocabulary codes. The call numbers are first mapped to the location-independent identifiers of virtual shelves. When access to an information resource is required, a location directory provides a second mapping of these location-independent server identifiers to actual network locations.

Results: The framework has been implemented in two different systems. One system is based on the Open System Foundation/Distributed Computing Environment and the other is based on the World Wide Web.

Conclusions: This framework applies in new ways traditional methods of library classification and cataloging. It is compatible with two traditional styles of selecting information sources, searching and browsing. Traditional methods may be combined with new paradigms of information searching that will be able to take advantage of the special properties of digital information. Cooperation between the library-informational science community and the informatics community can provide a means for a continuing application of the knowledge and techniques of library science to the new problems of networked information sources.

■ JAMIA. 1995;2:383-390.

An issue of fundamental importance to the development of large-scale, biomedical digital libraries is the need for efficient and reliable methods of access to networked biomedical information sources. Whatever else these methods do, they must allow the user

to select an information source, to discover its current, up-to-date network location, and to retrieve information from that source. Furthermore, such methods must work well within the confines of the collection of information sources maintained at a particular institution and must be scalable to support mutual sharing of the collections owned by different institutions.

This paper describes a framework for accessing networked biomedical information sources. The basic metaphor of this framework is that of a *virtual shelf*. A virtual shelf is a general-purpose server that is dedicated to a particular information subject class. A virtual shelf may be used to access a specific biomedical information source or to browse for information sources by biomedical subject.

This approach has been implemented in two dem-

Affiliations of the authors: Medical Informatics Group, School of Library and Informational Sciences, and Computer Science Department, University of Missouri-Columbia, Columbia, MO.

Supported in part by grants LM07089 and LM05513 from the National Library of Medicine, and also by Pittsburgh Supercomputing Center grant NCR930001P from the NIH National Center for Research Resources.

Correspondence and reprints: Timothy B. Patrick, PhD, 605 Lewis Hall, Medical Informatics Group, University of Missouri-Columbia, Columbia, MO 65211. email: migtemp@muccmail.missouri.edu, patrick@condor.cs.missouri.edu

Received for publication: 5/17/95; accepted for publication: 7/15/95.

onstration systems: one based on the Open Software Foundation/Distributed Computing Environment (OSF/DCE)¹ and one based on the World Wide Web (WWW).²

Background and Overview

It has long been known that location-*dependent* identifiers, such as WWW Uniform Resource Locators (URLs), have limited use as pointers to information sources in a decentralized, changing network environment. A URL, for example, works fine as a temporary pointer to an information source. However, as soon as the position of that information source in its host's file system changes, or as soon as some other aspect of its location changes, the URL for that information source is defunct. A solution that has been proposed to the problem of location dependence is to use enduring, location-*independent* identifiers of information sources and to map these identifiers to current network locations if, and only if, the need arises.³

The virtual shelf solution to the problem of location dependence is to use two mappings. The first mapping is from the enduring, location-independent call number assigned to an information source to the enduring, location-independent identifier of an appropriate general-purpose server. The second mapping is from the enduring, location-independent identifier of a general-purpose server to a current network location.

An advantage of this dual mapping is that it separates issues of logic from implementation issues. The mapping between call number and server identifier is a matter of logic and is implementation independent. The call number assigned to an information source contains a controlled vocabulary code that specifies its preferred subject class. Controlled vocabulary codes are used as location-independent identifiers of general-purpose servers. Thus, the vocabulary code contained in the call number of an information source also functions as the identifier of a server that will provide access to it. So, for example, a *molecular biology* information source will be accessed at a *molecular biology* server.

The mapping from server identifier to current network location is implementation dependent. Although the details may vary from implementation to implementation, the essential matter is that the controlled vocabulary codes—server identifiers are mapped to current, up-to-date network locations. The viability of this mapping is ensured by using a system of server registration. A location-independent server identifier is mapped to the network locations of servers currently registered in a location directory.

There may be, for any given server identifier, multiple servers currently registered in the location directory. That is, there may be multiple servers available for any given subject class. Such redundancy is useful to support fault tolerance and, combined with a load-leveling mechanism, may be used to provide faster access to desired information sources.

Selecting and Accessing a Specific Information Source

A specific information source is accessed in four steps. In the first step, the call number for an information source is retrieved from the catalog of information sources. In the second step, the call number is mapped to a location-independent server identifier. In the third step, the locations currently associated with the server identifier are retrieved from the location directory. In the fourth step, the information source is accessed via the server at one of the reported locations.

Figure 1 depicts a sample entry from a catalog of information sources. The entry shown is for the molecular sequence analysis software program, FASTA.⁴ The record includes the authoritative name of the information source, *FASTA*; the author of the information source, *William Pearson*; a set of Medical Subject Headings (MeSH)⁵ describing the information content of the information source; the type of information source, *software program*; and a network call number for the information source. The preferred subject heading, *molecular sequence data G5.331.599*, is marked with an asterisk. The call number, *G5.331.599-0001*, was generated from the code associated with the preferred subject heading. The latter portion of

Name	FASTA
Author	William Pearson
Subject	*molecular sequence data G5.331.599 sequence homology, nucleic acid G5.331.599.110.791 Sequence homology, amino acid G5.331.599.56.750
Type	software program
Call Number	G5.331.599-0001

Figure 1 A sample catalog record for the FASTA software program. The call number assigned to the FASTA software program, *G5.331.599-0001*, is based on the Medical Subject Headings (MeSH) code for the preferred subject heading, *molecular sequence data G5.331.599*.

the call number, 0001, is included to make the call number unique for FASTA.

Figure 2 depicts sample entries from a location directory. Each entry associates a server identifier with one or more network locations. The first entry associates the server identifier G5.331.599 with two network locations. It indicates that servers providing access to the *molecular sequence data* G5.331.599 subject class are located at the network locations *condor.cs.missouri.edu* and *csdeca.cs.missouri.edu*. In the MeSH vocabulary, *genetics, biomechanical* G5.331 is a parent of *molecular sequence data* G5.331.599. The second entry indicates that a server providing access to the *genetics, biochemical* G5.331 subject class is located at *biosci.mbp.missouri.edu*.

Once the location of a server has been retrieved from the location directory, the server may be used to access a specific information source. Figure 3 depicts the overall process of selecting and accessing the FASTA information source.

Browsing for Information Sources by Subject

Organizing servers by subject classes allows for an indirect mapping between the location-independent identifiers assigned to information sources and the

Server	G5.331.599
Location	condor.cs.missouri.edu csdeca.cs.missouri.edu
Server	G5.331
Location	biosci.mbp.missouri.edu

Figure 2 Sample location directory entries. The Medical Subject Headings (MeSH) codes G5.331.599 and G5.331 are interpreted as location-independent server identifiers. The location directory associates these server identifiers with the network locations of currently available servers.

current network locations of servers. This indirect mapping supports efficient browsing for information sources by subject classes. A virtual shelf may be used in a manner analogous to the way a physical shelf is used in a traditional library, i.e., a user can go to a virtual shelf server for a particular subject class and leisurely examine the materials available there.

Of course, even if the servers were not organized by subject class, it would still be possible to "browse" for information sources by subject class. A list of information sources assigned to the desired subject class could be retrieved from the catalog and then each source in turn could be accessed at its network location. However, if servers are not organized by subject class, that is, if information sources are scat-

- (1) Search the catalog with a subject heading

sequence homology, nucleic acid → catalog record:

Name	FASTA
Author	William Pearson
Call Number	G5.331.599-0001
Subject	*molecular sequence data sequence homology, nucleic acid sequence homology, amino acid

Figure 3 Use of a catalog and a location directory to access the FASTA software program.

- (2) Extract server identifier from the call number of FASTA

G5.331.599-0001 → G5.331.599

- (3) Search the location directory with the server identifier:

G5.331.599 → location directory entry:

Server	G5.331.599
Location	condor.cs.missouri.edu csdeca.cs.missouri.edu

- (4) Send the request for FASTA access to one of the reported locations:

FASTA request → *molecular sequence data* server at *condor.cs.missouri.edu*

tered arbitrarily across a collection of servers, then it is necessary to perform a separate location directory lookup for each information source examined. Systems based on the virtual shelf framework avoid this overhead. A single location directory lookup will suffice to determine the network locations of all information sources assigned to a desired subject class.

Browsing for information sources by subject class is accomplished in three steps. In the first step, the desired subject heading is mapped to a controlled vocabulary code that is used as a location-independent server identifier. In the second step, the locations currently associated with that vocabulary code-server identifier are retrieved from the location directory. In the third step, a list of available information sources is retrieved from the server at one of the reported locations. Figure 4 depicts the process of browsing for information sources in the subject class *molecular sequence data*.

Collections-based Access

Recently, a proposal has been made to develop procedures to support bibliographic searches across the Internet.⁶ A view of the Internet as a kind of vast, monolithic digital library appears to underlie this proposal. One implication of this view is a logical flattening of the Internet as an information resource. According to this view, the Internet supports a single large set of individual information sources.

The virtual shelf framework for accessing networked biomedical information sources is compatible with this flattened view of the Internet. That is, it could

be used to provide for universal access across multiple institutions. Assuming the existence of an appropriately published standard, it would be possible for anyone to put up a server that would advertise itself in some global directory service, e.g., *Aliweb* (<http://web.nexor.co.uk/public/aliweb/aliweb.html>), as serving a particular subject class. Consequently, for any given subject class there might be many servers available, representing the efforts of many different individuals and institutions. There are, however, good reasons to reject the paradigm of the Internet as a monolithic digital library and to adopt instead a view of the Internet as supporting multiple collections of information sources. According to this collections-based paradigm, information sources are always accessed as elements in the collections of information sources maintained by particular institutions, or multi-institutional cooperatives. The reputation of a given institution is a reasonable basis for trust in the authenticity of the information sources to which it claims to provide access. Thus, an advantage of the collections-based approach is that it provides a way to perform bibliographic searches across *trusted* portions of the Internet. If a desired information source is not available in the collection maintained by a user's home institution, the user may, given appropriate privileges, search the catalog and access the information source in the collection maintained by a trusted, co-operating institution.

Another advantage of the collections-based paradigm, particularly with regard to the virtual shelf method, is that collections-based access can ensure that a given information source is available at a server with a matching subject class. In general, agreement

(1) Map the subject heading for the desired subject class to a controlled vocabulary code.

molecular sequence data → G5.331.599

(2) Search the location directory using the controlled vocabulary code as a server identifier:

G5.331.599 → location directory entry:

Server	G5.331.599
Location	condor.cs.missouri.edu
	csdeca.cs.missouri.edu

(3) Retrieve a list of information sources from the server at one of the reported locations:

Request a list of information sources from the *molecular sequence data* server at *condor.cs.missouri.edu*

Figure 4 Browsing for biomedical information sources in the subject class *molecular sequence data*.

between the subject class assigned to an information source and the subject class assigned to a server is merely a necessary and not a sufficient condition for the availability of the information source at that server. A physical *molecular biology* shelf in a traditional library should not be expected to support all possible *molecular biology* information sources, and neither should a *molecular biology* shelf in a digital library. Of course, only the *molecular biology* information sources listed in the catalog for a collection should be expected to be available on the *molecular biology* shelf of that collection. However, even though the contents of the virtual shelves in a collection are determined by the catalog for that collection, agreement between the subject class of an information source and the subject class of a server is still not a sufficient condition for the availability of the information source at that server. For any number of reasons, including human error, a given information source might at any time be unavailable at a particular server. Yet where the catalog of information sources and the location directory are specific to an institution, bureaucratic measures may be taken to ensure that the availability of the information source at a matching virtual shelf is at least highly probable.

One disadvantage of collections-based access to networked information sources is that it requires some sort of coordination of catalogs between institutions. One solution to this problem of coordination is to use some form of standard, institutional-independent catalog records.⁷ The use of standard, institutional-independent catalog records allows different institutions to use the same catalog records and supports an open environment for cooperation collections. As discussed below, however, standardization of record format is not sufficient to ensure uniformity of cataloging across institutions. This may require central sources of catalog records similar to those used by the traditional library, for example, OCLC.

Beyond Bibliographic Selection

The virtual shelf framework is intended to provide efficient and reliable methods of access to networked biomedical information sources. It is bibliographic-based and is intended to provide methods of access to networked information sources that are analogous to methods of access to materials in the traditional library. The central metaphor used here, that of a *virtual shelf*, is intended to suggest an analogy between the logico-physical organization of materials in the traditional library and the assignment of instances of conceptually related information sources to the same physical server.

The virtual shelf framework is compatible with two traditional styles of selecting information sources, bibliographic searching and browsing. Yet the framework is not intended to provide final, definitive solutions to the many fundamental problems of selection. It has been for some time a subject of discussion whether bibliographic cataloging methods provide a sufficient foundation for selection of materials from a large information resource, digital or otherwise.^{8,9} This issue is most acute for a digital information resource. The bibliographic kinds of searching and browsing described here, though important, surely do not take full advantage of the power of digital libraries. After all, one of the most important properties of digital information is that it can be manipulated, analyzed, processed, and otherwise interpreted to produce further information.¹⁰ This property of digital information may provide the basis for methods of selection that are, for at least some contexts, more flexible, accurate, and informative than bibliographic searching and browsing. Examples of such methods include cooperative data mining¹¹ and intelligent information agents.¹²

The kind of bibliographic method described here may provide a functional foundation for more analytic and dynamic methods of selection. Selection requires *access* to be fruitful, and the bibliographic framework described here provides a solid foundation for efficiently and reliably locating and accessing information sources in a distributed and heterogeneous networked environment. In other words, bibliographic agents based on the virtual shelf framework might be used to provide services in conjunction with higher-level knowledge or information agents. A knowledge agent, perhaps based on collaboration with other peer knowledge agents, may select information; a bibliographic agent, acting in service of the knowledge agent, may locate physical instances of the works that contain that information. The knowledge agent might, in turn, be used to create or update the bibliographic agent.

Sample Implementations

The virtual shelf framework for accessing networked information sources has been implemented in two different systems. One system is based on OSF/DCE¹³⁻¹⁵ and one system is based on the WWW.¹⁶ The first two steps of access are the same for each system. The third and fourth steps are implementation dependent.

Figure 5 depicts an entry from the catalog of information sources used by the two systems. In the example entry shown, the ID field contains the au-

ID FASTA
 AU William Pearson
 DS Use FASTA to determine a family of known sequences
 DS that are related to a newly discovered sequence.
 KW *molecular sequence data G5.331.599
 KW sequence homology, nucleic acid G5.331.599.110.791
 KW sequence homology, amino acid G5.331.599.56.750
 TP software program
 CN G5.331.599-0001

Figure 5 The University of Missouri-Columbia catalog entry for the FASTA software program.

thoritative name of the information source, *FASTA*. The AU field contains the name of the author, *William Pearson*. The DS fields contain a description of the appropriate use of the FASTA analysis tool. The KW fields contain subject information. The preferred subject heading is *molecular sequence data G5.331.599*. The TP field specifies the type of information source. The CN field contains the network call number of the FASTA analysis tool, *G5.331.599-0001*. In steps one and two of accessing FASTA, its call number is retrieved from the catalog and is mapped to the location-independent server identifier, *G5.331.599*.

In step three, the location-independent server identifier is mapped to current network locations. Step three is essentially implementation dependent. For example, the OSF/DCE-based system uses the OSF/DCE Name Service as its location directory. The OSF/DCE Name Service uses a special type of location-independent Universal Unique Identifier (UUID) to

identify servers. This requires a further mapping from the server identifier of step two to one of these special UUIDs. In the WWW-based system, on the other hand, the vocabulary codes are themselves used as server identifiers and so no further mapping is required. Aside from differences due to their location directories, the two systems differ with regard to the kinds of clients and servers they use, the kinds of location information they use, and the kinds of server registrations they use. The OSF/DCE-based system uses Remote Procedure Call clients and servers. The WWW-based system uses a Mosaic client and HyperText Transfer Protocol (HTTP) servers. In the OSF/DCE-based system, location information is provided in the form of *binding information*. In the WWW-based system, location information is provided in the form of URLs. Finally, OSF/DCE supports automatic server registration, while the WWW-based system uses a manual type of server registration. Figure 6 depicts for each system sample mappings from a *molecular sequence data* server identifier to network locations. Note that the OSF/DCE UUID, *07533159-9000-0000-dddd-08002b37598d*, contains the vocabulary code *G5.331.599* with the *G* mapped to its alphabetic position, *07533159-9*.

Future Directions

Five future directions for development and implementation of the virtual shelf framework for accessing networked biomedical information sources are 1) the development of systems that do not always re-

WWW based System

The *molecular sequence data* server identifier *G5.331.599* is mapped to two current URLs

<http://condor.cs.missouri.edu>
<http://csdeca.cs.missouri.edu>

OSF/DCE based system

(1) The *molecular sequence data* server identifier *G5.331.599* is mapped to the OSF/DCE UUID *07533159-9000-0000-dddd-08002b37598d*; *07533159-9* is *G5.331.599* with *G* mapped to its alphabetic position, *07*.

(2) The server UUID *07533159-9000-0000-dddd-08002b37598d* is mapped to the server binding information

`<string binding> ncacn ip tcp:128.206.100.90[]`

Figure 6 Two location directory implementations. The World Wide Web (WWW)-based location directory uses the Medical Subject Headings (MeSH) codes themselves as server identifiers and maps them to Uniform Resource Locators (URLs). The Open Software Foundation/Distributed Computing Environment (OSF/DCE) location directory uses special Universal Unique Identifiers (UUIDs) as server identifiers and maps them to server-binding information.

quire a location directory search in order to access an information source, 2) the development of a central repository of standard-form catalog records, 3) the use of Unified Medical Language System (UMLS) Metathesaurus¹⁷ meta-concept vocabulary to index information sources, 4) a generalization of the framework to nonbiomedical domains, and 5) the development of standards for classifying servers.

It is fairly common for patrons of traditional libraries to bypass the process of using a catalog and shelf directory to locate a desired book, and instead to rely on memories of where the book is physically located in the library. If the situation has changed since the book was last used, if the book has been moved to a new location or if the shelves themselves have been rearranged, the library patron can resort to a catalog and shelf directory search to find the current location of the book in the library. The same common sense strategy may be applied to accessing networked biomedical information sources in the virtual shelf framework. The two implementations described here always require at least a location directory search in order to access a specific information source. A direction for future research is to develop systems that remember, modulo some acceptable N , the last network location at which a specific information source was located. This is somewhat analogous to setting a bookmark on a WWW browser, albeit with an available recourse if the bookmark becomes defunct. When access to the information source is required, the remembered location is tried first. If the access fails, recourse is made to a full-blown catalog and location directory search.

The use of standard-form catalog records supports cooperation among different institutional-based collections of information sources, but adherence to a standard record format alone does not ensure uniformity of cataloging across collections. The development of a central repository of standard-form catalog records can ensure uniformity of cataloging across collections. Moreover, it can relieve the caretakers of individual collections of the burden of producing all of the catalog records required by their respective collections. As is the case with the standard catalog records produced by OCLC, local codes might be appended to the standard call numbers by individual institutions. Yet the semantic core of the call number, the standard vocabulary code, would be constant across institutions and would provide a common reference point. A likely candidate institution to maintain a central repository of standard-form catalog records, at least for the biomedical domain, is the National Library of Medicine (NLM). In fact, the NLM UMLS Information Sources Map¹⁸ is a prototype central re-

pository of standard-form catalog records, although somewhat different from the ones described here.

The sample implementations of the virtual shelf framework described in this paper index information sources with the MeSH vocabulary. A natural experiment is to use the UMLS meta-concept vocabulary, rather than the MeSH vocabulary, to index information sources. Given appropriate gateways to the UMLS meta-concept vocabulary, catalog searches may then be conducted in the UMLS source vocabulary preferred by the user.

Although the two implementations of the virtual shelf framework described here are concerned with biomedical information sources, the framework itself is not intrinsically biomedical. It would be useful to attempt to apply the framework generally across multiple information domains. It would not, in such a generalized catalog, be appropriate to base all indexing and call numbers on MeSH alone. A more general vocabulary, e.g., Library of Congress Subject Headings,¹⁹ might be used alone or in combination with MeSH.

The classification of servers is central to the virtual shelf framework. In order for this framework to be scaled to support cooperation among the collections maintained by different institutions, it will be necessary to develop standards to guide classification of servers. It is, however, somewhat premature to speculate on the character of those standards. Classification of servers using hierarchical vocabulary codes implies a corresponding logical hierarchy of servers. It is reasonable to expect that the best kinds of server hierarchies to support, and so the best kinds of standards for server classification, will in part be determined by empirical tests.

Conclusion

The virtual shelf framework is an attempt to provide organized access to networked information sources by applying in new ways tested, trusted, and generally successful methods of library classification and cataloging. Professional and academic librarians have for years been the information classifiers of our intellectual communities, and they ought to take the opportunity to play a lead role in solutions to the problem of organized access to networked information sources. For this to happen it will be necessary to extend traditional library paradigms of information sources. For example, scientific data banks and scientific analysis programs—in their executable form—must be considered as much a part of a library collection as any book or journal, digital or otherwise.

In addition, to take advantage of the power of digital media, nonbibliographic methods of selection must be explored both in cooperation with and by librarians. Such cooperation between the library-informational science community and the informatics community can provide a means for a continuing application of the knowledge and techniques of library science to the new problems of networked information sources.

The authors thank Dr. Lynn Silipigni Connaway and two referees of this journal for their comments on this paper.

References ■

1. Open Software Foundation (OSF). Introduction to OSF/DCE. Englewood Cliffs, NJ: Prentice-Hall, 1992.
2. Obraczka K, Danzig PB, Shih-Hao LI. Internet resource discovery services. *IEEE Comput.* 1993;26:8-22.
3. Metcalfe ES, Frisse ME, Hassan SW, Schnase JL. Academic networks: Mosaic and World Wide Web. *Acad Med.* 1994;69(4):270-3.
4. Pearson WR, Lipman DJ. Improved tools for biological sequence comparison. In: *Proceedings of the National Academy of Sciences of the United States of America*. Washington, DC: National Academy of Sciences, 1988:2444-8.
5. National Library of Medicine. Medical Subject Headings. Bethesda, MD: National Library of Medicine, 1994.
6. Daniel R Jr. A global distributed directory service for the World Wide Web. <http://www.acl.lanl.gov:80/URI/URC/proposal/>
7. Martin D, Jul E, Burge M, Hickey C. Assessing information on the Internet: toward providing library services for computer-mediated communication, results of an OCLC research project. *Internet Res.* 1993;3(1):54-69.
8. Bush V. As we may think. *Atlantic Monthly.* 1945;176:101-8.
9. Levy DM. Cataloging in the digital order. In: *Digital Libraries '95: The Second Annual Conference on the Theory and Practice of Digital Libraries*, June 11-13, 1995, Austin, Texas.
10. Patrick TB, Springer GK. Data path integration in the digital library. In: *ISDL '95: International Symposium on Digital Libraries 1995*, August 22-25, 1995. Tsukuba Science City, Ibaraki, Japan.
11. Davies WHE, Edwards P. Agent-based knowledge discovery. In: *AAAI 1995 Spring Symposium Series on Information Gathering from Distributed, Heterogeneous Environments*, March 27-29, 1995 Stanford University.
12. Birmingham WP, Durfee EH, Mullen T, Wellman MP. The distributed agent architecture of the University of Michigan digital library. In: *AAAI 1995 Spring Symposium Series on Information Gathering from Distributed, Heterogeneous Environments*, March 27-29, 1995, Stanford University.
13. Patrick TB, Springer GK. Classifying and identifying servers for biomedical information retrieval. *JAMIA*. 1994 Supplement: 93-7.
14. Patrick TB, Springer GK, Nicholas HB, Ropelewski AJ. Information retrieval maps for biomedical research. In: *Medinfo '95, Proceedings of the 8th World Congress on Medical Informatics*, July 23-27, 1995, Vancouver, British Columbia, Canada.
15. Springer, GK, Patrick, TB, Davison SE, Nicholas HB, Ropelewski AJ. Use of a knowledge representation to share expertise in molecular sequence analysis. In: *KARP-95, Second International Symposium on Knowledge Acquisition, Representation, and Processing*, September 27-30, 1995, Auburn, Alabama.
16. Patrick TB. Network name services. In: *Medical Informatics Group Noon Seminar*, September 14, 1994, Medical Informatics Group, University of Missouri-Columbia.
17. Humphreys BL, Lindberg DAB. The Unified Medical Language System Project: a distributed experiment in improving access to biomedical information. In: Lun KC, Degoulet P, Piemme TE, Reinhoff O, eds. *MEDINFO 92*. Amsterdam, The Netherlands: North Holland, 1992:1496-500.
18. Masys DR, Humphreys BL. Structure and function of the UMLS Information Sources Map. In: Lun KC, Degoulet P, Piemme TE, Reinhoff O, eds. *MEDINFO 92*. Amsterdam, The Netherlands: North Holland, 1992:1518-21.
19. Library of Congress. Library of Congress Subject Headings, 14th ed. Washington, DC: Library of Congress, Office for Subject Cataloging Policy, 1991.