This chapter is forthcoming in *Evaluating Networked Information Services: Techniques, Policy, and Issues* edited by Charles R. McClure and John Carlo Bertot. Publisher: American Society for Information Science thru Information Today, Inc. Expected publication date, June 2001.

DRAFT PREPRINT

Assessing Interoperability in the Networked Environment: Standards, Evaluation, and Testbeds in the Context of Z39.50

William E. Moen, Ph.D.
Texas Center for Digital Knowledge
School of Library and Information Sciences
University of North Texas
Email: <wemoen@unt.edu>
Voice: 940-565-3563
Fax: 940-565-3101

Introduction

An underlying assumption of any network is that various components and processes will work together to produce desired results (e.g., data transmission, data interchange, reliability of services, etc.). The term interoperability has been used to characterize this working together, especially, the workings of lower level data communication components. Usage of the term has evolved to refer more generally to the extent to which different types of computers, networks, operating systems, and applications work together effectively to exchange information in a useful and meaningful manner. Miller (2000) suggests a perspective that is even more encompassing: he says that to be interoperable means “one should actively be engaged in the ongoing process of ensuring that the systems, procedures and culture of an organisation are managed in such a way as to maximise opportunities for exchange and re-use of information, whether internally or externally.”

The author is grateful to several colleagues who read earlier drafts of this chapter and suggested improvements, in particular, Clifford Lynch and Teresa Lepchenske.

This broader use of the term brings about new challenges in addressing and assessing the facets of interoperability. In this chapter, we explore the concept of interoperability and the roles of standards, evaluation, and testbeds in addressing and assessing the extent to which diverse systems interoperate. Specifically, this chapter explores issues of interoperability within the context of, and from experience with, the national and international standard *Information Retrieval (Z39.50): Application Service Definition and Protocol Specification (ANSI/NISO Z39.50-1995/ISO 23950)* (National Information Standards Organization, 1995). ANSI/NISO Z39.50 defines a computer–to–computer communications protocol used by systems for purposes of information retrieval. It is increasingly deployed in a variety of information communities including libraries, museums, geospatial and other data centers, and state and national governments. (For background information and a selected list of basic resources about Z39.50, see Moen, 1995; Moen and Lepchenske, 2000).

We use the knowledge gained from identifying Z39.50 interoperability issues to propose a multi–level framework for investigating and evaluating interoperability and to explain the limited role of the standard in assuring interoperability because of other critical factors that affect the “exchange and reuse of information.” The framework allows us to identify a range of threats to interoperability and to explore problems of interoperability both from implementor/vendor and user perspectives. This framework is being developed to address Z39.50 interoperability specifically in the context of cross–bibliographic database searching. We suggest, however, that this framework can assist in targeting assessment methods and metrics appropriate to different kinds data and applications where interoperability is a key factor. We exercise the proposed framework by reviewing and discussing several evaluations of Z39.50 projects in terms of the framework.

The networked environment is heterogeneous in that it hosts many different technologies, various data formats, multiple applications, and other networked life–forms. A functional goal in this environment is to hide the heterogeneity from users so they may effectively do business, search for information, communicate, and perform other tasks. For example, the centralized web search engine services appear to hide this complexity from the users. This is done by virtue of centralizing indexes to web resources and making them searchable through a single system. Contrast this to the way Z39.50 performs searching and retrieval across multiple resources of varying data types, on entirely different information retrieval systems—each with its own peculiar functionality and search structure; this presents a challenge at an entirely different level of magnitude when it comes to hiding the heterogeneity.

Hiding or masking the networked environment’s heterogeneity, or at least specific information spaces within the networked environment, and enabling users carry out specific tasks successfully is desired. The concept of interoperability can assist us in characterizing the issues and challenges in reaching the goal of transparent access to information. The experience with Z39.50 has taught us that getting diverse information systems to communicate and interoperate is a challenge. But that experience has also brought us to a point where we can articulate more clearly what is at stake with interoperability and to develop the methods and procedures for evaluating that interoperability.

Interoperability

There is little doubt interoperability is a key issue in the networked environment (see for example, Lynch, 1993; Lynch and Garcia-Molina, 1995; Miller, 2000; and Payette, et al., 1999). Interoperability or its absence can affect information access. Technical interoperability can raise important policy and organizational issues (Moen, 2000). Interoperability, however, suffers from a lack of a clear definition (Miller, 2000). Webster's Third International Unabridged Dictionary provides no definition of the term. The Oxford English Dictionary (OED) defines the component parts as: inter–"mutually, reciprocally, together; between or among themselves; with each other" and operable–"capable of being accomplished; capable of being actually used." Interoperability can be defined simply as the capability of two entities to work together to accomplish some process or task.

Within the information technology and networking context, a number of definitions surface:

- the ability of one machine . . . to interact usefully with other machines on a casual, ad hoc basis, without the prior planning or negotiation between the organizations operating those machines (Lynch, 1993 p. 3).
- components of a system…communicate with one another effectively, correctly, and provide the expected services to the user" (Preston & Lynch, 1994)
- the ability of different types of computers, networks, operating systems, and applications to work together effectively, without prior communication, in order to exchange information in a useful and meaningful manner (Abbas, et al., 1999).

The context for interoperability in this chapter is the networked environment in which numerous components work together to accomplish some goal. The components can take the form of hardware, software, data interchange formats, protocols, etc. To examine the complexity of interoperability, the focus of this chapter is on the use of Z39.50 to accomplish information retrieval tasks.

We can begin unbundling the concept of interoperability by viewing it in terms of levels and types and arrive at a multi–level perspective on it. Lynch and Garcia-Molina (1995) suggest a continuum of levels of interoperability which vary with individual implementations:

- from common tools and interfaces (navigation and access)
- to syntactic interoperability or the interchange of metadata or the diverse uses of digital objects
- to deep semantic interoperability or the ability to access consistently and coherently similar classes of digital objects and services.

They stated, "Deep semantic interoperability is a 'grand challenge' research problem; it is extraordinarily difficult, but of transcendent importance, if digital libraries are to live up to their long-term potential."

Based on experience with Z39.50 implementations, several types of interoperability can be articulated.

- **Low–level protocol (syntactic):** do two implementations interchange protocol messages according to specifications or a standard?
- **High–level protocol (functional):** do two implementations support the common services/functions?
- **Semantic level:** do two implementations preserve and act on meaning of information retrieval tasks?
- **User Task level:** do two systems support the information retrieval tasks of one or more user groups?

Additional types or levels will likely surface as we look at other network tools, technologies, and protocols, and our understanding of the scope and facets of interoperability problems increases.

In the case of Z39.50, low-level interoperability relates exclusively to syntactic protocol specifications defined in the standard. The other levels relate not only to the Z39.50 standard and its implementation in a system but also to the local information retrieval system and its capabilities. Preston and Lynch (1994) highlight the complexity of interoperable systems and how choices and deployment affect the ability of two systems to communicate meaningfully. The User Task Level goes further and focuses not only on the protocol and its function but considers those things that affect an information system user's sense of that system's interoperability with another information system. To further lay the groundwork for this discussion, we review briefly how Z39.50 works and point out the potential threats to interoperability in Z39.50 implementations.

Z39.50 In Brief

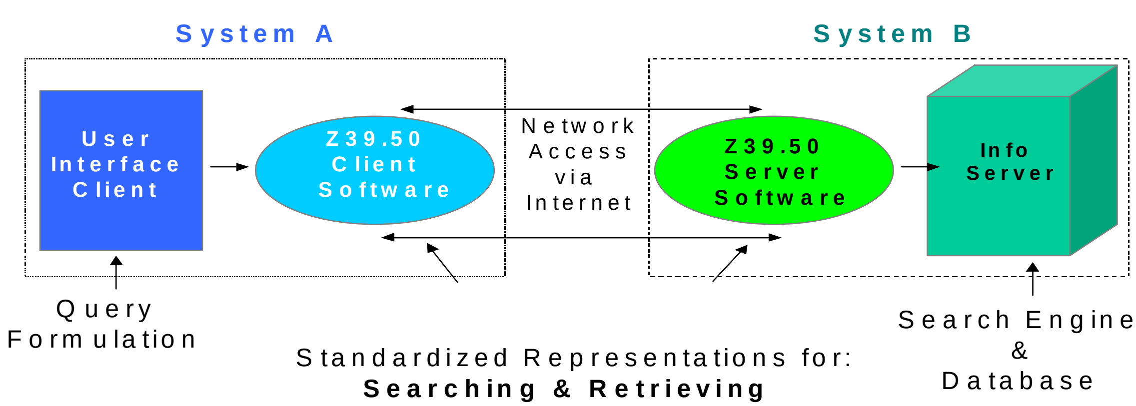
The purpose of Z39.50 is to allow a user on one information system to carry out search and retrieval transactions on another information system without prior knowledge of the details of the system. A user is able to interact with the familiar user interface of a local system to access one or more information systems and resources; this reduces the training burden for the user to learn many different search languages. The results of the search are returned and displayed in the familiar user interface of the user's local system.

The Z39.50 protocol provides an intersystem language that allows this communication between systems for the purposes of information retrieval. (Other communications protocols, such as File Transfer Protocol, Simple Mail Transport Protocol, etc. serve different purposes.)

Z39.50 is an application layer protocol within the seven–layer Open Systems Interconnection (OSI) Basic Reference Model (see Piscitello & Chapin, 1993). Implementations of Z39.50 have used various networking and transport mechanisms, but the majority of implementations currently use the Internet's TCP/IP protocols for. (The perceived effectiveness and interoperability of Z39.50 systems can be affected by the lower–level transport and telecommunications mechanisms used, but discussion of the impact of these mechanisms is beyond the scope of this paper.)

Z39.50 can be viewed as a protocol and a language two computers use to communicate for the purpose of information retrieval. The protocol provides the rules for using the language (e.g., the sequence of messages). Like any language, Z39.50 provides both syntax and semantics, structure and vocabulary, that comprises the language. This standardized, intersystem language is used to express searches, request information to be returned, and a range of other information retrieval transactions. Figure 1 illustrates Z39.50 client and server components for two communicating information systems.

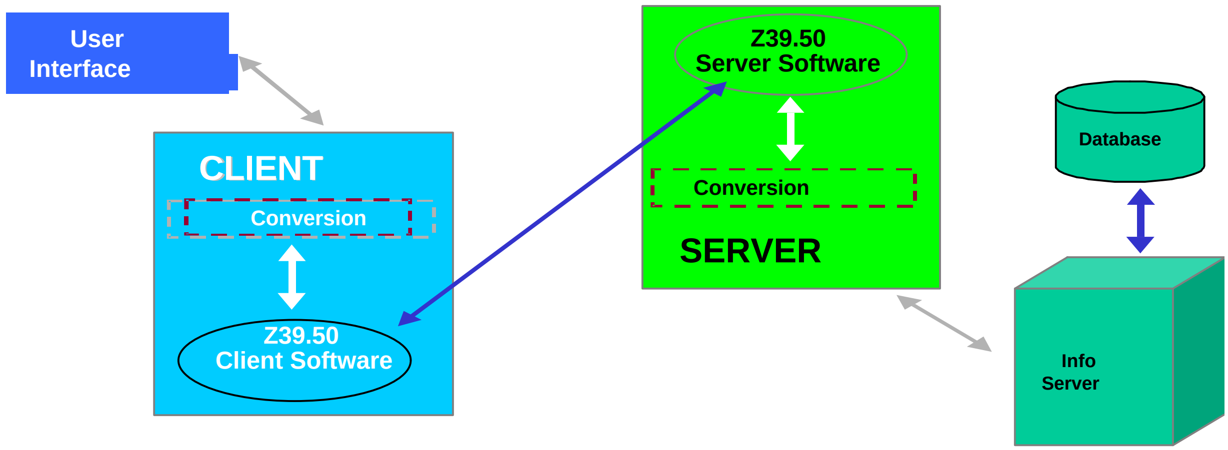
Figure 1
Information Systems Accessible via Z39.50



Z39.50 client and server software provide interfaces or front ends to local information retrieval systems. When a Z39.50 server is implemented for a local information retrieval system, a critical programming and configuration task is to provide for the translation or mapping from the local systems languages into the standardized Z39.50 language. An analogy for Z39.50 is Esperanto, which serves as a universal language that allows two people to communicate when they don't share a common language. Figure 2 illustrates where in Z39.50 implementations this conversion between local languages and Z39.50 standardized language occurs.

The analogy of Z39.50 as a standard language into which local systems convert their messages into helps illustrates the potential problems of interoperability. One must accurately translate from one language into another if meaning (semantics) is to be retained. Current implementations of Z39.50 for online library catalogs (and for other applications) cause users problems when searching and retrieving information because of different interpretations (i.e., mappings) of the standardized language into disparate system schemas and indexing policies. Interoperability is jeopardized by these differences.

Figure 2
Use of Conversion in Z39.50 Implementation



The information retrieval systems (and their underlying database search engines) on which Z39.50 is implemented are just as critical in addressing interoperability as the Z39.50 client and server software implementations. Interoperability can be jeopardized by differences in information retrieval systems, their configuration, and their functionality; all of which affect their capability to support information retrieval tasks. For example, System A may support truncation of search terms, while System B does not have that functionality. In this case, System A can use Z39.50 to send a truncation search to System B, but System B will not be able to execute the intended search of the user. Interoperability is jeopardized by these differences. We can refer to the issues and problems in the above example as constituting fundamental interoperability problems within the Z3950 context of search and retrieval.

Threats to Interoperability

Several critical threats to interoperability exist. A threat to interoperability can be defined as any organizational decision, specification, configuration or other implementation decision that reduces a user's ability to successfully search and retrieve information in a meaningful way and have confidence in the results. This frames interoperability no longer as a binary "yes, two systems are interoperable" or "no, two systems are not interoperable" but rather proposes a continuum of interoperability. Two systems may be more or less interoperable. In this context, the starting point for evaluating interoperability is the user and the capability provided to his/her by the systems to complete—successfully or adequately—information retrieval tasks.

Three focal technical areas can delineate Z39.50 interoperability:

- The standard
- Implementations of the standard
- Local information retrieval systems.

We will look briefly at each of these.

The early development of Z39.50 initially addressed the needs of libraries that wanted their systems to communicate for purposes of information retrieval and resource sharing. From the approval of the first version of Z39.50 in 1988 through 1992, Version 2, and finally with Version 3 approval in 1995, standards developers enhanced and expanded functionality available through the protocol. Additional functionality supported by the protocol provided the basis for broader interoperability.

For example, in the first two versions, the protocol machinery for exchanging anything other than data in the structure of MARC records was very limited. With Version 3, came the definition of a Generic Record Syntax (GRS-1) that enables the retrieval of arbitrarily structured data (e.g., data not represented in MARC format). Inadequacies of available record syntaxes in earlier versions of the standard would have precluded MARC and non–MARC systems from interoperating robustly. Version 3, however, is a fully featured and robust information retrieval protocol.

Current interoperability issues typically relate more to Z39.50 implementation and local information retrieval systems functionality rather than the adequacy of the standard, the protocol it defines, and the features supported by the protocol. Yet Z39.50 has been the object of criticism because to some it represents an outdated technology since it prescribes semantic level interaction between systems and also low-level protocol encoding mechanism. Some think that other available protocols could be used for passing the semantic content of the Z39.50 messages between systems.

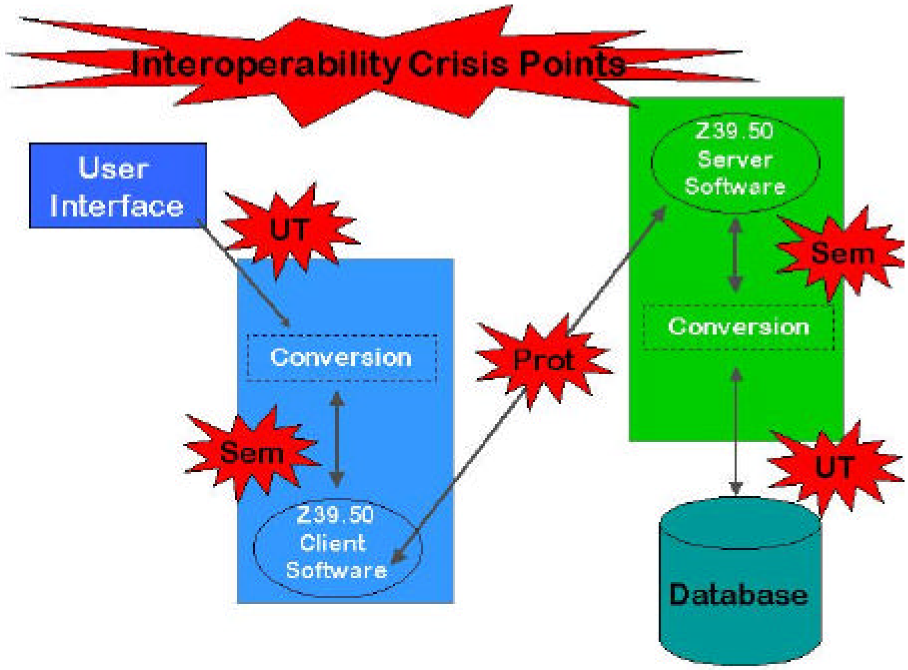
Z39.50 implementation is an involved, sometimes modular process. Because Z39.50 is as complex as it is fully–featured, an implementor wanting to develop a Z39.50 product needs to make a large number of decisions and choices from the options and features available in the standard. Two implementors independently can develop Z39.50 client and server products that are conformant with the standard; yet the result may be that one implementor's client does not interoperate very well (or at all) with the other's server. Precisely because of the flexibility in the standard (represented by the options and features), interoperability can be reduced or non–existent depending on how it is implemented. Thus, the flexibility provides both benefits and barriers to implementors and interoperability.

Finally, local information retrieval (IR) systems and their functionality are often the deciding factors in the extent of interoperability available. Z39.50 serves as an intermediary between two IR systems, sitting above the local IR system. Each local IR system has its own search and retrieval functionality (e.g., some systems support truncated searches, others don't). Each IR system makes the data in the records searchable according to local needs (i.e., level and policies for indexing). While this third area is beyond the jurisdiction of the Z39.50 standard, differences between local IR systems threaten interoperability. Z39.50 cannot make a local IR system execute a particular type of search if the underlying IR system doesn't support it.

For example, assume that two IR systems utilize Z39.50 for intersystem communication. A searcher on System A sends a search via Z39.50 for a subject term that it wants truncated (e.g., search is for the subject "comput*") and requests System B execute the search against its database. While the Z39.50 software on System B may understand the request, the local IR system does not support truncation searches. Thus, the user is not able to carry out the IR task as desired. This points to an important fact; the existence and use of a standard is not the only factor in achieving interoperability.

Tying this together in the context of a multi–level perspective on interoperability and with the four types of interoperability introduced above (i.e., Low–level protocol, High–level protocol, Semantic, and User), we can see more clearly the potential crisis points when two systems communicate. Figure 3 illustrates these crisis points.

Figure 3
Crisis Points in Z39.50 Interoperability



Key: UT = User Task Level; Sem = Semantic Level; Prot = Low– and High–Level Protocol

Given the maturity of many of the Z39.50 toolkits and source code, the Low–level protocol area usually does not cause much in the way of interoperability problems. The focus at the Low-level is on the ability of Z39.50 implementations (clients and servers) to construct the appropriately structure protocol messages (Application Protocol Data Units) and send them in the correct sequence. In 1992, the Coalition for Networked Information (CNI) sponsored the first Z39.50 Interoperability Testbed. That testbed resolved many low–level protocol interoperability issues; Clifford Lynch, the coordinator of that testbed, noted that the focus was on getting Z39.50 systems to talk to one another and achieving mechanical interoperability over the Internet (Personal Communication, January 2000).

Z39.50 Version 2 (1992) offered fewer services and less functionality than the current Version 3, and the 1992 Z39.50 Interoperability Testbed focused on the three core Z39.50 services: Initialization, Search, and Present. Conformance to the standard was not the focus of the testbed; it was demonstrable interoperability between systems. Without the foundation of low–level protocol interoperability, however, none of the other levels of interoperability are possible.

The area of High–level protocol interoperability focuses also on the implementation of the standard, but the issues here relate to the specific features or services of Z39.50 supported by communicating systems. At a minimum, Z39.50 implementations will support:

- Initialization Service (for the client and server to set up the communications session)
- Search Service (for the client to express a search in the standardized Z39.50 vocabulary and the server respond with a report of the search)
- Present Service (for the client to request records be returned in a suitable syntax and the server complying).

But the standard defines other protocol services, including:

- Delete: to delete result sets on a server
- Access Control: to support various forms of user authentication
- Sort: to order the records in a result set
- Scan: to browse an ordered list of terms (e.g., an index) to identify search terms.

The standard does not specify that any of these newer protocol services must be implemented for a Z39.50 client or server to be conformant. This results in a situation where two systems might each support Initialization, Search, and Present, but only one of the two supports Sort and Scan. In that case a request from System A that supports Sort and Scan to System B that does not will fail the test of interoperability—at least in terms of the Sort and Scan services. A Z39.50 client may be more or less robust that a Z39.50 server, and therefore we need to examine if a specific Z39.5 client or Z39.50 server threatens interoperability. It stands to reason, though, that two systems (i.e., Z39.50 client and server software) must support the same services from the standard to provide a basis for interoperability using those services. This higher–level of protocol interoperability relates to functionality provided by the Z39.50 client and server, but this level also intersects with the local IR system functionality. In the above example, System B may have chosen not to support Sort and Scan precisely because its local IR system does not have the capability to sort result sets and provide a browseable list of its indexes. Although limitations of local systems directly impact interoperability, newer information retrieval systems will likely support the common functions defined in Z39.50, and this issue will begin to resolve itself.

Semantic level interoperability is less amenable to easy solutions in part because of the range of factors affecting it. In the report from the U.S. Government's Information Infrastructure Technology and Applications (IITA) Digital Libraries Workshop, Lynch and Garcia-Molina (1995) noted that the "precise definition of deep semantic interoperability was the subject of some debate, but deals with the ability of a user to access, consistently and coherently, similar (though autonomously defined and managed) classes of digital objects and services, distributed across heterogeneous repositories, with federating or mediating software compensating for site-by-site variations." As noted earlier, they characterized deep semantic interoperability as a "grand challenge" for digital library research.

Going back to the analogy of Z39.50 as a language, the meaning (semantics) of the protocol messages needs to be clear if two systems are to share an "understanding" of the message. Z39.50 provides standardized "vocabularies" to express queries using registered sets of attributes (where attributes are used in the Z39.50 query to characterize a search term). The attribute sets provide the "words" in the vocabulary for searching. Using the vocabulary, we can express—in a standardized Z39.50 query—a query for books ***by*** Mark Twain (where Twain is the author) as well as a query for books ***about*** Mark Twain (where Twain is the subject). Compare this to a typical web search engine where it is not possible to differentiate a search for Mark Twain as author versus Mark Twain as subject (the reasons for this go beyond the scope of this paper).

Z39.50 implementations, however, do not always support (i.e., understand and act on) the same "words" from the standardized vocabulary for searching. Taking an example from searching library catalogs, System A wants to search System B for a corporate author and uses the correct Z39.50 attribute to characterize its search term as a corporate author. But System B does not support that particular Z39.50 attribute. The semantic intention of the user and his/her search cannot be acted upon. However, the System B does support a name search, and in an attempt to be helpful, processes the corporate author search as a name search; the results, however, may include records that are not relevant to the original corporate author search; semantic loss has occurred. In both these cases, semantic interoperability is reduced or does not exist.

The Semantic level of interoperability is strongly affected by the local IR systems functionality and indexing policies. Although the standard provides mechanisms for clearly—if not unambiguously—expressing search requests, retrieval requests, and other IR functional requests, the differences in local IR systems can jeopardize semantic interoperability. In the example above, the two systems are online library catalog (i.e., bibliographic databases) built using standard MARC records. However, System A allows specific MARC fields to be searched for corporate author names while System B, with the same basic set of records, has chosen not to create indexes or is incapable of creating indexes to support the access point of corporate author. There is likely a strong relationship of the search capabilities of the underlying IR system and the Z39.50 attributes it supports in its Z39.50 server software.

The final level of interoperability focuses on the user and his/her tasks. The levels discussed above have become fairly well–understood, at least in the library community, through Z39.50 implementation experience. The User task level of interoperability has yet to be investigated and only a tentative outline of this level can be offered here. Users use information systems for a variety of reasons, but one can assume users are interested in searching for information to solve a particular problem, answer a question, or complete a task. The question for users is whether two systems are interoperable to the extent they meet the users objectives. While the Z39.50 community has already and continues to work out protocol (syntactic and functional) and semantic interoperability, those efforts have not been well–informed by an understanding of users' expectations for interoperability. (The evaluation studies and the recent profiling efforts discussed below are implicitly acknowledging these expectations.) When are two systems interoperable enough for users to complete their tasks? The fact is that there are different user groups with differing expectations for interoperability. Take the examples of an undergraduate student searching remote library catalogs via Z39.50 to see if a neighboring library has a copy of a particular book versus a cataloger searching remote libraries for purposes of cataloging a particular edition of a book; the user tasks are different, and the likelihood is that expectations of interoperability, especially at the semantic level, are quite different. Acceptable interoperability for one user group may be insufficient interoperability for another.

Achieving interoperability is a complex challenge. The diversity of information retrieval systems, the variation in conformant implementations of a standard, the numerous levels where threats to interoperability can occur, and the lack of empirical data on users' expectations for interoperability between systems all need to be taken into account when addressing and assessing interoperability. Once we can identify the threats to interoperability, there is a major issue of the cost to solve the interoperability problem. In the case of Z39.50, we now have hundreds, if not thousands of Z39.50 client and server implementations on top of existing online library catalogs. This installed base can be considered part of the interoperability problem. What level of interoperability can we achieve and at what cost?

The multi–level perspective offered above provides a way to understand the challenges to interoperability and can point to levels of interoperability evaluation that can be undertaken. The next section discusses several evaluation reports of Z39.50 projects that can be understood in the context of this multi–level perspective.

Recent Evaluation Studies of Z39.50 Implementations

In 1998, four studies of Z39.50 use by libraries were released:

- State of Iowa Libraries Online (SILO) initiative (Blue Angel Technologies, 1998)
- Virtual Canadian Union Catalogue project (Lunau, 1998)
- Virtual Electronic Library of the Consortium on Institutional Cooperation (CIC) (Hinnebusch, 1998)
- California Linked Systems Project (Williams, 1998)

These studies were the first published empirical investigations and resulted in important findings about the problems of Z39.50 interoperability. More recently, Williams (1999) provides an update to her 1998 California Linked Systems Project, and Coyle (2000) reports on an evaluation to compare a virtual union catalog using Z39.50 with the University of California's Melvyl catalog. One can surmise from the studies that a standard can provide a context for evaluation, but complex standards such as Z39.50 with its many choices and options, and implemented in conjunction with different information retrieval systems, cannot serve as an evaluation metric by virtue of being a standard.

The common thread among these studies is that they examined and evaluated the use of Z39.50 to search across library catalogs. Although interoperability problems exist when a Z39.50 client communicates with a single Z39.50 server, some relatively straightforward solutions are available to improve interoperability such as gaining knowledge about the server implementation and configuring the client specifically for communicating with that server. Customizing client configuration for each server—an approach used in these projects for improving interoperability—is an option, but whether or not it is an effective approach is an open question. For example, the Virtual Canadian Union Catalogue (vCuc) report indicated that the total process of configuring a client for a specific server—including identifying the server, acquiring details about the server implementation, configuring the client, and testing—took on average 3.5 to 7 hours (Lunau, 1998, p. 15).

Designers of the Z39.50 protocol envisioned it for point–to–point communications. Some current Z39.50 clients allow the user to search more than one resource at a time, and in these cases, the Z39.50 client establishes multiple point–to–point sessions with remote Z39.50 servers. These resources may be relatively homogeneous (e.g., searching multiple online library catalogs) or diverse (e.g., searching online catalogs and museum collection management systems). The use of Z39.50 for searching multiple resources at the same time compounds the complexity of the interoperability problems especially when those resources are heterogeneous (e.g., searching library catalogs and museum collection management systems at the same time). The Z39.50 client will have the responsibility for attempting to hide the diversity among the resources being searched (i.e., making this work requires more robust client software). Look again at Figure 3 and imagine communication is occurring not between two systems but twenty or more, with an associated multiplier effect on the interoperability crisis points.

Each of the studies had specific purposes and foci, and approached their evaluations differently. For example, the SILO evaluation used questionnaires to librarians and vendors as well as follow–up interviews to collect data on the SILO implementations. The CIC evaluation was an in–depth analysis on the CIC's libraries implementations of Z39.50 (e.g., the Z39.50 services supported, the attribute supported, etc.) and the underlying information retrieval systems indexing of MARC records. The vCuc report contained evaluation of Z39.50 implementation but also served as a final report on the entire vCuc Pilot Project. The study of the California Linked Systems Project focused on an assessment of what had been achieved so far and identified challenges and successes of using Z39.50 for searching multiple remote library catalogs. Only the SILO report is called an "evaluation" where the others include evaluative assessments of Z39.50 implementations as part of the larger studies. These differences make cross–study comparability difficult, yet several important common themes emerged and can be discussed in terms if the multi–level perspective on interoperability presented above.

- **Low-level protocol interoperability:** The studies did not indicate interoperability problems at this level. This reflects the relative maturity of the protocol and implementations at this level.
- **High-level protocol interoperability:** The implementations discussed by the reports (in 1998) were primarily Version 2 protocol implementations and generally only supported the three core Z39.50 services of Init, Search, and Present. However, the CIC analysis showed a trend toward diversity in other Z39.50 services supported by various servers. Of the 11 Z39.50 servers analyzed, 7 support the Scan service and 1 supports the Sort

service. On the client side, 5 of the 6 clients supported Scan, and 2 of the 6 supported Sort (Hinnebusch, 1998). The other studies did not detail the services supported but the clients and servers. If a client and server do not support the same Z39.50 service, overall interoperability is reduced.

- **Semantic interoperability:** Most of the interoperability problems identified in the studies fall into this category. This category spans Z39.50 implementation and local IR system functionality and configuration. Two key aspects of this set of problems are differences in support for Z39.50 attribute types and values (Z3.50 implementation), and local indexing practices and polices (local IR system functionality and configuration).

  <u>Z39.50 Attribute Support:</u> Although there was common support for the Z39.50 Bib-1 Attribute Set, there were wide differences in support for specific attribute types and values. In Z39.50, the Use attribute type indicates an access point. The vCuc project analyzed server support for Z39.50 Use attribute revealed that only two access points were supported in common by 12 separate implementations: Title and Subject (Lunau, 1997). Hinnebusch's analysis of CIC implementations revealed more access points in common with all Z39.50 implementations supporting the following Use attributes: Title, ISBN, ISSN, Subject, and Author. In practice, this means that in the vCuc implementations, a user could only be assured that Title and Subject searches would work across all implementations. CIC implementations provided more extensive interoperability for searching give the additional common access points. Whether or not either the vCuc or CIC implementations provide sufficient search interoperability cannot be determined unless we identify specific user task requirements.

Another aspect of interoperability at the level of attribute support is the combination of attribute types and values supported by server implementations and how they interpret the attribute types and values. Servers interpret the attributes to determine how to execute a particular query (e.g., is this a truncated author keyword search or some other kind of author search). Hinnebusch noted "it is most assuredly this mismatch of client configuration and server support [for attribute combinations] that lies at the heart of many of the difficulties that CIC libraries have experienced with Z39.50 searching" (1998, p. 8). The SILO report also noted that differences in Use attribute and attribute combination support had an impact on the quality of the search results.

<u>Local Indexing Policy:</u> The common support for limited Z39.50 Use attributes in server implementations likely reflects the differences in local IR access points or indexes available. Each index is populated with data from one or more fields/subfields in a MARC record. The indexing policy for local IR system determines which fields/subfields will be indexed to support searches. For example, the MARC record has any number of fields/subfields that contain title data. If different IR systems do not index the same fields/subfields, title searches across different systems (even if they hold the same MARC records) may yield different results. The variation of indexing policies of different local IR systems affected the overall quality of searching in the projects. Only the CIC report provides an analysis of systems' indexing policies (see Figures 16 through 23 in Hinnebusch, 1998). The SILO and vCuc reports indicate, however, that this is a key issue to address to improve interoperability.

- **User task interoperability:** None of the report specifically discussed interoperability in terms of users. As noted above, this type of interoperability has not been well

> articulated. However, one can glean from the approach in the vCuc project the importance of taking into account user tasks when evaluating interoperability. The vCuc evaluation included controlled searching using a set of test searches (Lunau, 1998, p. 9 and p. 25+). Participants were also asked to do uncontrolled searches that would reflect searches that are typical for your institutions. Types of searches suggested included: ILL requests, reference questions, searches for cataloging copy, searches to verify bibliographic information prior to ordering a new item, etc. Summary results from the testing indicated that a confluence of factors affected search results including particular characteristics of the client, the database, and the interpretation of attributes. As a result, the report stated, "Users were not confident in the results they received…. Only 37% felt they were receiving accurate and relevant results" (Lunau 1998, p. 13).

The importance of these evaluations cannot be underestimated. Empirical data confirmed existing anecdotal evidence of Z39.50 interoperability problems. Further, the studies clarified both Z39.50 and non–Z39.50 variables affecting interoperability. As the SILO report indicated, "a number of weaknesses listed by library respondents are not problems associated with the Z39.50 protocol itself, but rather in the inconsistent way in which vendors have implemented many of the open–ended items" (Blue Angel Technologies, 1998). Additionally, including users' perceptions in the evaluation indicate that even more factors may subvert a user's judgment about interoperability, such as quality of a database and system response time. These factors are not related necessarily to the standard or its implementation but can affect the users perception of adequate interoperability to support their tasks.

These studies provided empirical evidence of problems with Z39.50 interoperability, and they also suggested possible solution paths. One key solution path is the use of Z39.50 profiles to improve Z39.50 interoperability.

Profiles

Profiles are an auxiliary standards mechanism that defines a subset of specifications from one or more standards to improve interoperability. Their use with computer communications protocols stems from the original context of Z39.50 development, namely the Open Systems Interconnection (OSI) framework of the 1980s. OSI protocols typically included many options and choices from which implementors chose. One objective of a profile is to detail a set of specifications from those options and choices in a base standard(s). Implementors conforming to a profile would produce systems that had an improved likelihood of interoperability.

The motivations for developing Z39.50 profiles can be categorized as either:

- Prescribing how Z39.50 should be used in a particular application environment (e.g., government information, cultural heritage museums, etc.)
- Solving interoperability problems with existing Z39.50 implementations within a community (e.g., libraries) or across two or more communities (e.g., library and museums).

The latter will be the focus of this discussion. (A review of Z39.50 profiles is beyond the scope of this chapter; for a list of Z39.50 profiles see the Library of Congress Z39.50 Maintenance Agency website, specifically <http://lcweb.loc.gov/z3950/agency/profiles/profiles.html>.)

Although the evaluation studies discussed above appeared in 1998, Z39.50 implementors, particularly those within the Z39.50 Implementors Group (ZIG), had acknowledged interoperability problems of searching across library catalogs at least since 1995. In early 1996, the ATS-1 Profile (specifying author-title-subject searching across library catalogs) provided the first attempt to solve some of the problems. The ATS-1 Profile was only a stop gap measure, and problems left unresolved in that profile were addressed by several subsequent profiles built upon ATS-1. These profiles added more specifications to resolve problems not adequately addressed by ATS-1. (For a brief discussion of these profiles and the issues they addressed, see Moen, 1998a).

Between 1999 and 2000, an international effort produced *The Bath Profile: An International Z39.50 Specification for Library Applications and Resource Discovery* (Bath Profile Group, 2000; for background on the Bath Profile, see Lunau, 2000). The Bath Profile itself was informed by several previous profiles, but most importantly by the *Z Texas Profile: A Z39.50 Profile for Library Systems Applications in Texas* (Texas Z39.50 Implementors Group, 1999; for background on the Z Texas Profile, see Moen, 1998b). These two profiles focused effort on resolving semantic interoperability problems for cross–catalog information retrieval and prescribing specific Z39.50 services required to support various user tasks (e.g., Init, Search, Present, Scan).

As an example, the Bath Profile addresses semantic interoperability for searching by defining a core set of 19 searches; requirements for these cross–catalog searches resulted from discussions among librarians. Defining the searches included naming a search, prescribing IR system behavior to process the query, and prescribing the Z39.50 query vocabulary to unambiguously express the query. For example, the Profile defines an Author Keyword Search with Right Truncation. The semantics (i.e., prescribed IR system behavior) for that search is: “Searches for complete word beginning with the specified character string in fields that contain the name of a person or entity responsible for a resource." The specification of the query using Z39.50 Attributes is:

- Use Attribute (1) = author (1003)
- Relation Attribute (2) = equal (3)
- Position Attribute (3) = any position in field (3)
- Structure Attribute (4) = word (2)
- Truncation Attribute (5) = right truncation (1)
- Completeness Attribute (6) = incomplete subfield (1).

This combination of attribute types and attribute values expresses this and only this search.

The Profile also defines an Author Keyword Search (without truncation). The semantics of this search is: "Searches for complete word in fields that contain the name of a person or entity responsible for a resource." And the specification of the query using Z39.50 Attributes is:

- Use Attribute (1) = author (1003)
- Relation Attribute (2) = equal (3)
- Position Attribute (3) = any position in field (3)
- Structure Attribute (4) = word (2)
- Truncation Attribute (5) = do not truncate (100)
- Completeness Attribute (6) = incomplete subfield (1).

The only difference between the two Z39.50 queries is the difference in the Truncation Attribute value (1 or 100), but this semantically differentiates one search from the other. Thus, there should not be any ambiguity when the server receives either of these queries as to the search behavior being requested. The Bath and Z Texas Profiles address two kinds of interoperability, High–level protocol and Semantic.

Even though the profiles address the Z39.50 aspect of semantic interoperability, the Semantic level is also affected by the indexing policies and search functionality in the local IR system. To address variations in indexing in different systems, the Texas Z39.50 Implementors Group (TZIG) is developing recommendations and guidelines for common indexing policies to support the searches specified in the Profile. However, Hinnebusch (1998, p. 3) warns, "the integrated library system vendor community has long used indexing as a market determinant. Attempts at standardizing indexing of bibliographic data have been firmly rejected by the vendors and there is little reason to expect this situation to change." Also, local institutions define indexing policies that address the needs of their users, and these local decisions result in wide variation in available access points across library catalogs. The TZIG believes that the networked environment changes many business models, and the demand for reliable cross–catalog searching may influence or change vendors' business models when it comes to indexing policies and may encourage libraries that want to participate in virtual library catalogs to consider common indexing policies to support certain searches.

Profiles can improve interoperability, but like a standard, they cannot by themselves assure interoperability. Yet, by reducing the diversity of options and choices from the base standard, profiles can provide a more effective evaluation baseline. This leads to the possibility of formal interoperability testing as an evaluation method for assessing compatibility and appropriateness of products through demonstrable interoperability.

Interoperability Testing and Interoperability Testbeds

The evaluation studies discussed previously highlighted interoperability as a key problem with Z39.50 implementations. Simply putting in a Request for Proposal (RFP) that a system must be Z39.50 compliant did not result in interoperable systems for many of the reasons already discussed. Profiles provide a solution path to address High–level protocol and Semantic level interoperability by specifying configuration of Z39.50 client and server implementations. Requiring compliance to a profile's specification in a RFP may yield systems that are more interoperable. To adequately assure goodness of products, however, another level of evaluation is essential.

Interoperability testing is an appropriate evaluation approach for complex standards in the networked environment. It is a "procedure in which two or more implementations are tested against each other, with the standard [or profile] used primarily as a reference to judge problems and incompatibilities, and secondarily as a guide to the functions that should be tested and the general behavior to be expected (Preston and Lynch, 1994). Conformance testing is another approach to evaluating a single implementation where it is "compared to the standard [or profile] to be sure that the implementation does what the standard [or profile] specifies. If the implementation conforms to the specifications set out by the standard [or profile], then it is considered to be interoperable" (Preston and Lynch, 1994). We pointed out earlier, however, that two developers of Z39.50 products could produce conformant implementations that may not be interoperable. Demonstrable interoperability rather than conformance to a standard is critical. Lynch (1993, p. 42) stated unequivocally "an interoperable system in the real world is far more valuable to the mission of a library than a system that achieves abstract virtue through the implementation of some specific set of politically or philosophically correct standards."

There are currently, however, no accepted testing methodologies, formal processes, and interoperability benchmarks by which customers and vendors can assess conformance to profile specifications or demonstrate effective interoperability between systems that claim conformance to a standard or profile. An interoperability testbed is an accepted approach for interoperability testing. Preston and Lynch (1994) noted that:

> Because the emphasis is on implementations, testbeds lead to a "whole system" approach to testing rather than one focused on individual standards conformance or interoperability and can be very useful not only in dealing with problems directly related to a given standard but in identifying problems that arise from the interaction between different standards or at the boundaries between standards and implementor agreements often needed to produce real-world interoperating systems.

This holistic approach for interoperability assessment is informing the creation of a formal interoperability testbed at the Texas Center for Digital Knowledge and the School of Library and Information Sciences, University of North Texas.

[Author's note: If/When we get the final decision on the IMLS grant award for this interop testbed, I'll likely want to insert some wording to that effect and cite the proposal which I'll put up online. They are saying Sept 19 on the award announcement.]

The testbed will be a technically– and organizationally–trusted environment providing producers and consumers a forum in which to demonstrate and evaluate Z39.50 products. The foundation of the testbed will be reliable, valid testing methods and scenarios using Z39.50 client and Z39.50 server reference implementations. Z39.50 producers and consumers will use the testbed for assessing products/systems that claim conformance to a specific set of Z39.50 specifications (e.g., as represented in a profile). A primary focus will be semantic interoperability testing.

Building upon the success and experience in the 1992 Z39.50 Interoperability Testbed and with the focal goal of assessing semantic interoperability, we have identified a number of components for the proposed testbed. Although this project will focus initially on Z39.50 semantic interoperability, the following testbed components are likely critical in interoperability testing no matter the networked system under evaluation:

- **Technical:** Hardware and software components of the communicating systems; there is a requirement for flexibility in configuration of the components.
- **Reference implementation:** A trusted implementation, configured to the extent possible to support a particular set of specifications, and open to inspection and vetting by trusted experts; reference implementations provide the benchmarks for search and retrieval results against which other systems will be evaluated.
- **Content:** The test corpus or data used by the system (e.g., for a library catalog, a set of bibliographic records); should be of sufficient size to provide a real–world environment for testing.
- **Test Scenarios and Benchmarks:** Reliable and validated test scenarios to provide comparability between systems participating in the testbed; benchmarks are established by using the test scenarios with the reference implementations; metrics will assess the levels of interoperability achieved by systems.
- **Procedural:** Procedures of the testbed must be clear and transparent; they should address participation guidelines; security of the test corpus; feedback and challenges; and privacy, conflict–of–interest, and trade secret safeguards.
- **Organizational:** Organization and governance of the testbed must be articulated; the sponsoring organization should be a neutral and trusted party free from conflicts of interest with organizations that might be using the testbed for product testing and evaluation; stakeholder should be represented on an advisory group for the testbed.

- **Personnel:** Sufficient personnel with suitable expertise to develop and manage the ongoing operation; an advisory group representing a range of the stakeholders that will monitor, assess, challenge, and otherwise help make the testbed a success.

The interoperability testbed will be a vehicle for assessing the extent of interoperability that is achieved between a vendor's implementation of Z39.50 clients and/or servers and the testbed's reference implementations of Z39.50. It must be formal, rigorous, and trusted. But to be successful, the testbed must be used.

Motivating Participation in Interoperability Testing

Evaluation in the networked environment can lead to improved services or abandonment of services and products that no longer provide value to users and organizations. Product developers want their systems to succeed in the marketplace. Rigorous interoperability testing discussed above may dissuade testbed participation if product developers and vendors think that evaluating the extent of their products' interoperability with other vendors' products may adversely affect market share. What will motivate participation in evaluations of networked services and products?

In the context of Z39.50, reference to the standard in RFPs has induced vendors to develop Z39.50 products. Profiles provide additional configuration requirements for vendors. Referencing a profile in a RFP will motivate vendors to configure Z39.50 clients and servers accordingly. But an interoperability testbed offers both producers and consumers a mechanism to demonstrate the appropriateness and compatibility of products. Including in a RFP that a vendor must demonstrate interoperability of their products in the testbed can serve as a motivation for participation in the testbed.

The testbed itself needs to be attractive to potential participants. Concerns discussed above such as neutral sponsorship, trade secret protection, trustworthiness, and rigor can make the testbed attractive. Add to that an organizational culture of the testbed that emphasizes constructive assistance for product improvement rather than "gotcha!" type evaluations, and the testbed can become a support for product development as well as product evaluation.

All the stakeholders—the consumers as well as the producers—need to perceive the testbed as open, trustworthy, and fair. In developing the testbed, stakeholder input is vital. Participation in an interoperability testbed may hinge on acceptance of the testing methodologies and metrics. The process of developing specific aspects of the testbed such as metrics and test scenarios should involve input, if not consensus, of the stakeholders.

Finally, the ongoing operation of a testbed needs to be supported. For the proposed Z39.50 testbed, initial development and operation of the testbed will rely on grants and contributions. Over the longer term, the stakeholders—consumers and producers—may be the ones to fund its operation. That outcome assumes that the testbed addresses a pragmatic need for effective evaluation.

Based on these considerations, the organization and management of the testbed rather than the technical and evaluative components may offer the larger challenges to its success. Yet, the need for interoperability testing and testbeds appears as a fundamental evaluative need in the networked environment if seamless access to distributed information is to be realized.

Conclusion

Interoperability is a term that means different things to many people. Yet, the concept is central in discussions of how components in the networked environment will work together. Assessments of interoperability have become more problematic as the concept has evolved from reference to interchanging data packets to retaining semantic intention of users in networked information retrieval transactions.

This chapter provided a preliminary framework for discussing and analyzing levels of interoperability in the context of Z39.50. It analyzed specific areas of system–to–system communication that threaten interoperability. Some of these areas are better understood than others (e.g., the Low– and High–level protocol areas) but others, such as the Semantic level, have not been fully explored from the perspective of evaluation. The interoperability testbed described here will address and explore evaluation methods for semantic interoperability.

We also introduced the User Task level of interoperability. Although this topic has yet to be understood and framed properly for research, we suggest that no matter the extent of interoperability at the other levels, it will be the users' assessment of perceived interoperability of networked systems about which we will be most concerned. We can assume that the User Task Level may be an intersection of concepts related to interoperability and concepts related to usability.

There is work to be done since in this dynamic environment, evaluation strategies and methods continually need to evolve to address the critical issues of interoperability in its many guises. The proposed multi–level framework for interoperability will inform a rigorous testbed to be established at the Texas Center for Digital Knowledge. The testbed initially will allow the investigation of kinds of interoperability and methods for assessing semantic interoperability within the context of Z39.50 implementations and library catalogs. We anticipate, however, new insights about interoperability issues that may have broader utility in evaluation strategies for the networked environment. There is, however, a pragmatic aspect to interoperability that must be addressed, and that is facing up to the cost of solving interoperability deficiencies. Yet, without an understanding of what can be considered acceptable levels of interoperability for different users groups, there can be little in the way of reasoned discussion on the cost effectiveness of various solutions to improve interoperability.

Works Cited

Abbas, June, Monika Antonelli, Mark Gilman, Pamiela Hight, Valli Hoski, Jodi Kearns, Teresa Lepchenske, Martha Peet, Mike Pullin, and Amy Stults. (1999). *An Overview of Z39.50, supplemented by a case study of implementing the ZebraServer under the Linux operating system.* Denton, TX: School of Library and Information Sciences, University of North Texas. Available: <http://www.unt.edu/wmoen/Z3950/GIZMO/contents.htm>

Bath Profile Group. (2000, June). *The bath profile: An international Z39.50 specification for library applications and resource discovery, Release 1.1. An internationally registered profile*. Available: <http://www.ukoln.ac.uk/interop-focus/bath/current/>

Blue Angel Technologies, Inc. (1998, May). *An evaluation of Z39.50 within the SILO project*. Chester, PA: Blue Angel Technologies, Inc. Available: <http://www.silo.lib.ia.us/bluang.html>.

Coyle, Karen. (2000, March). The virtual union catalog: A comparative study. *DLib Magazine, 6*(3). Available: < http://www.dlib.org/dlib/march00/coyle/03coyle.html>.

Hinnebusch, Mark. (1998, August). *Report to the CIC on the state of Z39.50 within the consortium*. Champaign, IL: Committee on Institutional Cooperation. Available: <http://ntx2.cso.uiuc.edu/cic/cli/z39-50report.htm>.

Lunau, Carrol. (2000, March). *The Bath profile: What is it and why should I care?* Ottawa, Canada: National Library of Canada. Available: <http://www.nlc-bnc.ca/bath/prof.pdf>

Lunau, Carrol. (1998, June). *Virtual Canadian union catalogue pilot project: Final report*. Ottawa: National Library of Canada. Available: <http://www.nlc-bnc.ca/resource/vcuc/vcfinrep.pdf>

Lunau, Carrol. (1997, September). *Bib-1 attributes supported by selected vendor servers.* Ottawa: National Library of Canada. Available: <http://www.nlc-bnc.ca/resource/vcuc/ebib-1.pdf>

Lynch, Clifford. (1993, March) Interoperability: The standards challenge for the 1990s. *Wilson Library Bulletin, 67*(7): 38-42.

Lynch, Clifford and Garcia-Molina, Hector. (1995). *Interoperability, scaling, and the digital libraries research agenda: A report on the May 18-19, 1995 IITA digital libraries workshop.* (U.S. Government's Information Infrastructure Technology and Applications (IITA) Working Group, Reston, Virginia). Available: <http://www-diglib.stanford.edu/diglib/pub/reports/iita-dlw/>

Miller, Paul. (2000, June). Interoperability: What is it and why should I want it? *Ariadne* 24 Available: < http://www.ariadne.ac.uk/issue24/interoperability/intro.html>

Moen, William E. (2000). Interoperability for information access: Technical standards and policy considerations. *The Journal of Academic Librarianship, 26*(2): 129-132.

Moen, William E. (1998a). *Texas Z: The Texas Z39.50 requirements and specifications project. A discussion paper*. Prepared for the Texas State Library and Archives Commission. Available: <http://www.unt.edu/wmoen/Z3950/TexasZDPAug98.htm>

Moen, William E. (1998b, Winter). The Z Texas project: Implementing Z39.50 to improve statewide access to Library catalogs and other online resources. *Texas Library Journal, 7*(4)*:* Available: <http://www.txla.org/pubs/tlj74_4/moen.html>

Moen, William E. (1995). *A guide to the ANSI/NISO Z39.50 protocol: Information retrieval in the information infrastructure*. Bethesda, MD: NISO Press. Available: <http://www.cni.org/pub/NISO/docs/Z39.50-brochure/>

Moen, William E. and Lepchenske, Teresa. (2000, May). *Z39.50: Selected list of resources*. Denton, TX: School of Library and Information Sciences, University of North Texas. Available: <http://www.unt.edu/wmoen/Z3950/BasicZReferences.htm>

National Information Standards Organization. (1995). *Information retrieval (Z39.50): Application service definition and protocol specification (ANSI/NISO Z39.50-1995).* Bethesda, MD: NISO Press. Available: URL: < http://lcweb.loc.gov/z3950/agency/document.html >

Payette, Sandra, Blanchi, Christophe, Lagoze, Lagoze, Overly, Edward A. (1999, May). Interoperability for digital objects and repositories. *D-Lib Magazine,* 5(5). Available: <http://www.dlib.org/dlib/may99/payette/05payette.html>

Piscitello, David M. and Chapin, A. Lyman. (1993). *Open systems networking: TCP/IP and OSI*. Reading, MA: Addison–Wesley.

Preston, Cecilia M. and Clifford A. Lynch. (1994). Interoperability and conformance issues in the development and implementation of the government information locator service (GILS). In William E. Moen and Charles R. McClure, *The Government Information Locator Service (GILS): Expanding Research and Development on the ANSI/NISO Z39.50 Information Retrieval Standard*, *Final Report*. Syracuse, NY: School of Information Studies, Syracuse University.

Texas Z39.50 Implementors Group. (1999, April). *Z Texas profile: A Z39.50 profile for library systems applications in Texas, Release 1.0.* Available: <http://www.tsl.state.tx.us/ld/projects/z3950/TZIGProfile99Apr20.htm>.

Williams, Joan Frye. (1999, November). *A study of linked systems with recommendations for the future of the CLSA statewide data base*. Sacramento, CA: Information Technology Planning, Management + Marketing. Available: <http://ferguson.library.ca.gov:80/lds/linksys.pdf>

Williams, Joan Frye. (1998, May). *Linked systems study update and proposed technology infrastructure for the library of California.* Sacramento, CA: Information Technology Consultant.
Available: <http://http://ferguson.library.ca.gov:80/loc/telecom/t_studies/index.cfm>