67th IFLA Council and General Conference
August 16-25, 2001

Code Number: **050-203(WS)-E**
Division Number: **0**
Professional Group: Universal Dataflow and Telecommunications **Workshop**
Joint Meeting with: -
Meeting Number: 203
Simultaneous Interpretation: -

Improving Z39.50 Interoperability: Z39.50 Profiles and Testbeds for Library Applications

William E. Moen
School of Library and Information Sciences
Texas Center for Digital Knowledge
University of North Texas
Denton, TX USA
E-mail: wemoen@unt.edu

Introduction

An operating assumption for the networked environment is that many different information systems need to interoperate for users to successfully discover and retrieve distributed resources. Meaningful interoperability is often elusive. In the library community, the Z39.50 standard protocol (ISO 23950/ANSI/NISO Z39.50) for information retrieval promised seamless and transparent networked access to library resources. Too often, the reality has not lived up to the promise. This paper discusses two efforts that offer solution paths to Z39.50 interoperability.

Interoperability and Z39.50

From its inception in the early 1980s, the goal of the Z39.50 protocol was to enable interoperable system-to-system communication for purposes of information retrieval. (See Moen [1998] for a history of Z39.50 development.) The developers of the standard envisioned its use as a point-to-point protocol where an origin/client would interact with a target/server to search databases of MARC records. That model is still the foundation of the protocol, but current Z39.50 implementations and applications now provide for searching multiple targets and databases concurrently. And the use of Z39.50 is no longer limited to searching bibliographic databases of MARC records. The protocol is a key technology for networked information retrieval.

Interoperability between implementations of Z39.50 has been, and continues to be, a challenge. One of the reasons for the slow progress towards meaningful interoperability is that interoperability is not just one thing. Often, interoperability is presented as a monolithic entity, but in fact, there are many types and levels of interoperability. Miller (2000) provides an expansive definition of interoperability and indicates some of the types. There are multiple factors that can affect information system interoperability (Moen, 2001b). Our understanding of the complexity of interoperability has increased, and we are able to better pinpoint our efforts in addressing and resolving interoperability.

For example, issues of Z39.50 interoperability can be categorized into syntactic protocol interoperability, functional protocol interoperability, and semantic interoperability (Moen, 2001a). By identifying these categories, we are better able to highlight specific interoperability issues and begin to propose solutions.

Implementors and users of Z39.50 have recognized that two "Z39.50 compliant" systems don't necessarily interoperate well (or at all). This happens in part because of the rich functionality available in the optional services and specifications available in the protocol. Unless both systems are configured similarly in terms of choices from the protocol, interoperability suffers. The emergence of Z39.50 profiles provides the solution path to address more detailed specifications for Z39.50 clients and servers to improve interoperability.

Z39.50 Profiles

Profiles are an auxiliary standards mechanism and define a subset of specifications from one or more standards to improve interoperability. One objective of a profile is to detail a set of specifications from those options and choices in a base standard. Implementations conforming to a profile have an improved likelihood of interoperability.

Since the mid-1990s, groups using Z39.50 have developed a number of Z39.50 profiles. 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).

Examples of the first category are the Application Profile for the Government Information Locator Service (GILS) and the CIMI Profile: A Z39.50 Profile for Cultural Heritage Information. (See Z39.50 Maintenance Agency [2001] for a complete listing of Z39.50 profiles.)

This paper focuses on the use of profiles in the second category, and an important example of those Z39.50 profiles for library applications is the Bath Profile: An International Z39.50 Specification for Library Applications and Resource Discovery (The Bath Group, 2001).

The Bath Profile

During the late 1990s, many different national, state, and project groups developed Z39.50 library application profiles to address very similar problems. Unfortunately, these profiles required different Z39.50 specifications for similar client and server behaviour. In 1999, a group of those profile writers and other Z39.50 implementors met in Bath, England to agree on a core set of specifications that addressed international requirements and thus serve as a common core for other national, state, and project library application profiles (Lunau, n.d.).

The Bath Profile is an Internationally Registered Profile first released in June 2000. The profile's purpose is to:

> … identify those features of the Z39.50 standard that are required to allow effective use of Z39.50 software in a range of library applications, including search and retrieval of bibliographic data from library catalogues; transfer of holdings information; cross-domain searches between libraries, museums and archives; updating union catalogues; item ordering and document delivery.

The current version of the Profile identifies specifications for three applications:

- Basic bibliographic search and retrieval, with primary focus on library catalogues
- Bibliographic holdings search and retrieval
- Cross-domain search and retrieval.

To accommodate future extensions to the profile, and to allow an implementor to choose to support one or more applications specified in the Profile, the structure of the document is modular. Functional Areas respond to a specific set of requirements. The three applications listed above are three separate Functional Areas in the Profile. An implementor could choose to support only Functional Area A: Basic Bibliographic Search and Retrieval, with Primary Focus on Library Catalogues, and be compliant with the Profile. To reduce barriers to entry, the Profile also specifies different Levels of Conformance. Level 0 specifies minimum but meaningful Z39.50 specifications to improve interoperability. Higher levels of conformance have additional requirements.

A discussion of Functional Area A can illustrate the structure and content of the Profile. Functional Area A addresses the need to improve interoperability when doing basic bibliographic searching across library catalogues. Level 0 requires clients and servers to support four specific searches:

- Author Search – Precision Match for Established Name Heading
- Title Search – Keyword
- Subject Search – Keyword
- Any Search – Keyword.

Level 1 adds 15 searches. Thus an implementation claiming conformance at Level 1 must support at a minimum the 19 searches identified in the Profile.

For each search, the Profile indicates the behaviour of the Z39.50 client and server and the functionality needed in the underlying information retrieval system. It then specifies the exact Z39.50 attribute combination to express a particular search. The following two searches illustrate these specifications:

Level 0 Title Search — Keyword

Uses: Searches for complete word in a title of a resource.

Attribute Type	Attribute Values	Attribute Names
Use (1)	4	title
Relation (2)	3	equal
Position (3)	3	any position in field
Structure (4)	2	word
Truncation (5)	**100**	**do not truncate**
Completeness (6)	1	incomplete subfield

Level 1 Title Search — Keyword with Right Truncation

Uses: Searches for complete word beginning with the specified character string in fields that contain a title of a resource.

Attribute Type	Attribute Values	Attribute Names
Use (1)	1	title
Relation (2)	3	equal
Position (3)	3	any position in field
Structure (4)	4	word
Truncation (5)	**5**	**right truncation**
Completeness (6)	1	incomplete subfield

The differences in the semantics of the expected behaviour are reflected in the use of different attribute combinations. Specifying an exact attribute combination for every search in the Profile reduces the ambiguity when a client sends a query to a server. A Bath conformant server knows exactly the type of search being requested when it receives a query with a specific attribute combination. This is a vast improvement over current practices of Z39.50 clients and servers, and has the potential to substantially improve the semantic interoperability for searching.

The Profile also addresses the retrieval aspect of Z39.50. A client and server must interchange a record in a syntax recognized by both. Therefore the Profile specifies which Z39.50 Record Syntax(es) must be supported. In Functional Area A, the Profile requires clients and servers to support a combination of MARC 21 and UNIMARC along with at least one non-MARC syntax (e.g., XML).

Since browsing indexes is an associated requirement for searching, the Profile specifies the use of the Z39.50 Scan service in Level 1. Six Scans are specified and must be supported by clients and servers that claim conformance at Level 1 of Functional Area A.

These examples give a sense of how the Profile specifies the use of Z39.50 for search and retrieval across library catalogues. The Profile offers a similar level of specification for the other Functional Areas related to bibliographic holdings information and cross-domain search and retrieval.

In summary, the Bath Profile addresses requirements for international interoperability and identifies Z39.50 specifications to support those requirements. The Bath Profile provides choices from the Z39.50 standard and specifies exactly what and how Z39.50 clients and servers must support those choices. Similarly configured clients and servers will lead to improved interoperability. For additional information about the Bath Profile, see the Bath Profile Maintenance Agency (2001).

Extending the Bath Profile

The core specifications in the Bath Profile provide a common ground for international interoperability. The Bath Profile assumes that other national, state, and project profiles will respond to more specific requirements and extend the core specifications to address those requirements. Companion profiles use the Bath Profile as a foundation and add appropriate specifications to respond to their local needs. Two examples of this are the ONE-2 Profile and the NISO Profile.

The ONE-2 Profile (1999) is being developed as part of the OPAC Network in Europe 2 project. The ONE-2 Profile uses the Bath Profile as a subset of its specifications. In addition to the Bath Profile

Functional Areas, it adds specifications for Ordering and Interlibrary Loans, Union Catalogue Update, and Server Information (ExplainLite). The ONE-2 Profile is to be completed in Summer 2001.

The U.S. National Information Standards Organization (NISO) established a standards committee in Fall 2000 to develop a U.S. National Z39.50 Profile for Library Applications (National Information Standards Organization, 2001). As of Summer 2001, work on this profile is underway. The first version of the U.S. national profile includes the two Bath Profile Functional Areas related to searching library catalogs and interchanging holdings information. The profile extends the number of searches defined in the Bath Profile Functional Area A. The U.S. national profile is being developed through NISO's normal standards development procedures. The result will be a U.S. National Standard that addresses national requirements while using the Bath Profile core specifications as its foundation.

Implementation experience in recent years has indicated the need for more detailed Z39.50 configuration specification to improve interoperability. A framework for improving interoperability in library applications is now available. The Bath Profile provides a core set of specifications that responds to international Z39.50 requirements. Companion profiles that use Bath as their foundation preserve interoperability at the international level while addressing specific national, state, and project requirements.

Now that we have more detailed specifications for configuring Z39.50 clients and servers, we have an opportunity for establishing testing scenarios to assess conformance to the profiles.

A Z39.50 Interoperability Testbed

Except for an early interoperability testbed in 1992 when Z39.50 was in its infancy, there has been no formal testing environment to assist implementors in improving interoperability (see Lynch [1992] for a description of the 1992 testbed). In Fall 2000, the U.S. federal Institute of Museum and Library Services (IMLS) awarded a National Leadership Grant to the University of North Texas for a research and demonstration project to establish a Z39.50 interoperability testbed. Partners in the project include OCLC, SIRSI, and Sea Change/Bookwhere who are contributing data, software, and expertise (for information on the project, see Moen, 2001c).

Realizing the Vision of Networked Access to Library Resources: An Applied Research and Demonstration Project to Establish and Operate a Z39.50 Interoperability Testbed, is a 20-month project to design and demonstrate test methods and metrics to assess interoperability between systems using Z39.50. The overall goal for the project is to improve Z39.50 semantic interoperability among libraries for information access and resource sharing.

There are currently no accepted testing methodologies, formal processes, and interoperability benchmarks by which customers and vendors can assess conformance or demonstrate effective interoperability between Z39.50–accessible systems. The Z39.50 interoperability project focuses on:

- Developing and demonstrating rigorous methodologies, test scenarios and procedures to measure and assess the extent of interoperability between Z39.50 implementations
- Producing a model for interoperability testbeds that can be used by other communities.

The testbed provides a vehicle for assessing the degree of interoperability achieved between a vendor's implementation of Z39.50 clients and/or servers and the testbed's Z39.50 reference implementations.

The testbed is establishing a test dataset of 400,000 MARC 21 records from OCLC's WorldCat databse. Reference Z39.50 client and server implementations will be configured to specifications in the Bath and

U.S. National profiles. A set of test searches against the test dataset through the reference implementations will establish benchmarks for subsequent interoperability testing.

The testbed will be available to vendors and organizations that have Z39.50 client and server implementations. Z39.50 server implementations will mount the test dataset on their systems, and project staff will send test searches from the reference Z39.50 client. Results from the searches will be compared to the established benchmarks. Analysis of variance from the benchmarks will assist the testbed participants in fine-tuning their implementations to improve interoperability.

Conclusion

Interoperability between diverse information systems in the networked environment presents complex and at times confounding challenges. Recent studies evaluating Z39.50 implementations have found that the same searches done on different Z39.50 compliant systems can yield different results (Blue Angel Technologies, Inc., 1998; Hinnebusch, 1998; Lunau,1998). The studies documented librarians' concern about the reliability of Z39.50 to provide effective search and retrieval across library catalogs. Improvements in interoperability between systems can substantially increase users'–especially librarians and other information professionals–confidence that Z39.50 products provide reliable results when searching multiple resources.

The Z39.50 profiles discussed in this paper address the interoperability challenges for library applications. By specifying detailed configuration choices for Z39.50 clients and servers, the likelihood of interoperability is increased. The Bath Profile provides a foundation of core specifications for international interoperability, and national, state, and project profiles can build on the Bath foundation to address national or other requirements while safeguarding international interoperability.

Further these profiles provide a set of specifications that can be used to configure and test implementations. The Z39.50 interoperability testbed at the University of North Texas will provide a testing environment to assist implementors in improving their products. Another expected benefit of the testbed is that additional factors affecting interoperability will surface and can be address (e.g., local indexing practices).

The Z39.50 protocol offers an important, if not strategic, tool for networked information retrieval. After 20 years of work and development, we are close to solving critical interoperability problems. Ultimately, it is the users of our libraries who will benefit from these efforts. Unparalleled access to networked resources may finally become a reality.

References

Bath Group. (2001). Bath profile: An international z39.50 specification for library applications and resource discovery.
Available URL: <http://www.nlc-bnc.ca/bath/bp-current.htm>.

Bath Profile Maintenance Agency. (2001). Bath profile maintenance agency [website]. Available URL: <http://www.nlc-bnc.ca/bath>.

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

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

Lunau, Carrol. (n.d.). The bath profile: What it is and why should I care.
Available URL: <http://www.nlc-bnc.ca/bath/bathfaq.pdf>.

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

Lynch, Clifford A. (1992). The CNI Z39.50 interoperability testbed project. *Information Standards Quarterly, 4*(2): 4.

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. (2001a). Assessing interoperability in the networked environment: Standards, evaluation, and testbeds in the context of Z39.50. In Charles R. McClure and John Carlo Bertot (Eds), *Evaluating Networked Information Services: Techniques, Policy, and Issues*. Silver Spring, MD: American Society of Information Science and Technology through Information Technology, Inc.

Moen, William E. (2001b, June). Mapping the interoperability landscape for networked information retrieval. Proceedings of the First ACM+IEEE Joint Conference on Digital Libraries, Roanoke, VA, June 2001.

Moen, William E. (2001c). Realizing the vision of networked access to library resources [website].
Available URL: <http://www.unt.edu/zinterop>.

Moen, William E. (1998, August). The development of ANSI/NISO Z39.50: A case study in standards evolution. Unpublished dissertation. Syracuse, NY: School of Information Studies, Syracuse University.
Available: UMI Publication Number, AAT 9904698.
Available URL: <http://www.unt.edu/wmoen/dissertation/DissertationIndex.htm>.

National Information Standards Organization, Standards Committee SC AV. (2001). Developing a U.S. national Z39.50 profile for library applications [website].
Available URL: <http://www.unt.edu/zprofile/index.html>.

ONE-2. (1999). OPAC network in Europe 2 [website].
Available URL: <http://www.dbc.dk/one2/>.

Z39.50 Maintenance Agency. (2001). Z39.50 profiles.
Available URL: <http://lcweb.loc.gov/z3950/agency/profiles/profiles.html>.