Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Realizing the Vision
Improving Interoperability with Z39.50 Profiles

William E. Moen
<wemoen@unt.edu>
School of Library and Information Sciences
University of North Texas
Denton, TX 72603

OCLC SiteSearch User Group Meeting
Dublin, Ohio
May 2, 2000

Presentation Components

The Vision:
Distributed and Integrated Access

The Question of Interoperability

The Z Texas Profile

The Bath Profile

Next Steps

Answer as Many Questions as I Can!

What is Z39.50

- An American National Standard
  - Z39.50 means it was developed by NISO
  - Z39.50 means it was the 50th standard approved by NISO
- Z39.50 is a set of specifications, procedures, and structures for computer communication.
- Z39.50 allows users to search one or more databases from a single interface

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

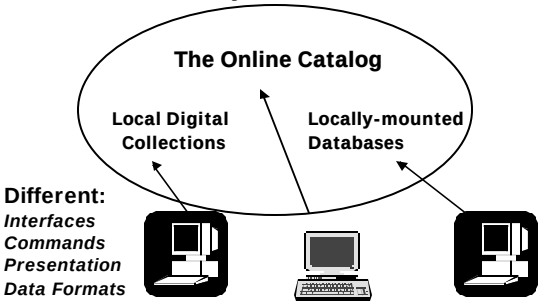
William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Purpose of Z39.50

- To overcome problems of multiple database searching
- To allow users of one system:
  - to search databases on another system
  - to retrieve records from that system
  - without knowing the unique features of each database system
- Search and retrieve bibliographic and non-bibliographic resources (integrated access)
- Lay a technical foundation for resource sharing

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 4

Traditional Online Access to Library Information



Different:
Interfaces
Commands
Presentation
Data Formats

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 5

Information Access From a Single Resource

- Our library catalogs epitomize the power of this approach
  - User interface is optimized
  - Interface tightly linked with database search engine
  - Knowledge of available access points
  - Pre-determined display formats
  - System-specific user options and functions

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 6

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

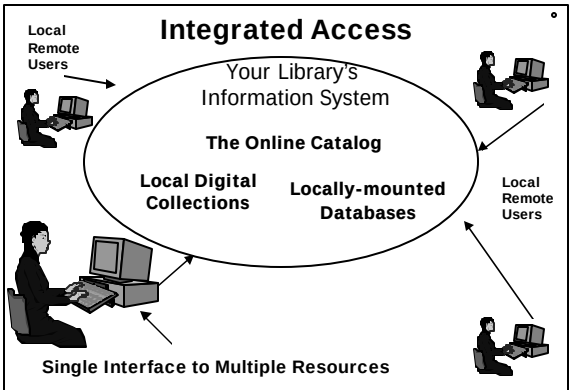
Integrated Information Access Option 1

- Learn to search each resource using native interface
- To achieve: burden is on user to learn each on
- Providing web interface does not reduce burden

Integrated Information Access Option 2

- Develop mechanism for querying separate resources in a standardized manner
- To achieve: burden is on information systems to communicate in ways that hide the difference between them
- Providing Z39.50 access is a solution

Integrated Access



Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

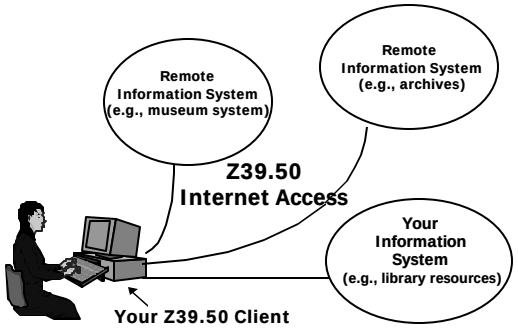
William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Distributed and Integrated Access

- Resources are geographically dispersed (possibly around the globe)
- Resources may hold similar or dissimilar types/formats of information
- Single user interface to distributed resources

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 10

Distributed and Integrated Access



William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 11

Challenges to Distributed and Integrated Access

- Modeling search and retrieval across a variety of information resources
- Providing vocabulary and grammar to express searches (e.g., access points and other query criteria)
- Providing mechanisms for retrieving different formats/types of resources (e.g., bibliographic records, full-text resources, digital objects, or components of database records

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 12

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Z39.50 Addresses the Challenges

- Two systems share a common "understanding"
  - of access points
  - of databases records
- Standardized language -- vocabulary (semantics) and grammar (syntax) for
  - initiating and closing information retrieval sessions
  - expressing searches and their meaning
  - delivering result set records
- The common understanding allows interaction with a remote system

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 13

Centralized vs. Distributed IR

- Two models for networked information retrieval
- Centralized
  - Web search engine model
  - IR from one server is not too difficult
- Distributed
  - Z39.50 model
  - IR from multiple databases across multiple servers is more problematic

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 14

Interoperability: The Fundamental Problem

- Ability of client to successfully search and retrieve information in a meaningful way
- Key issue when searching multiple databases containing similar resources
- Difficult problem when searching databases containing diverse types of resources
- Z39.50 profiles provide a solution

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 15

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Interoperability

- **Working definition:**
  User's ability to successfully search and retrieve information in a meaningful way and have confidence in the results
- **Major issue in networked resources & digital library environment**
- **Z39.50 interoperability**
  - **Theoretical vs practical interoperability**
  - **Conformance to standard vs demonstrable interop**

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 16

Threats to Interoperability

- **The functionality supported by the standard**
- **Differences in implementation of the standard**
- **Differences in local IR systems**
  - **Z39.50 cannot improve searchability of resources**
  - **It can only support what local and remote systems offer**

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 17

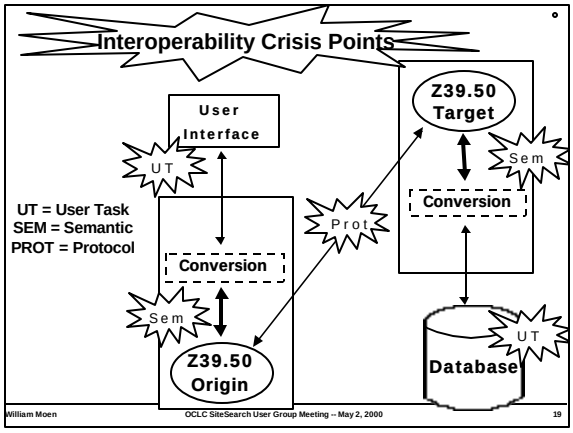
Levels of Interoperability

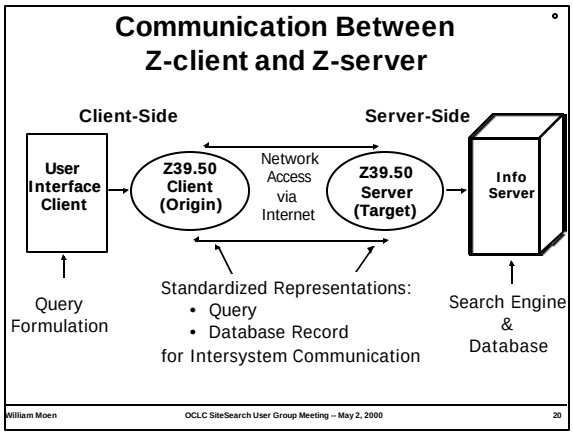
- **User Task level**
  - **do systems support IR tasks of one or more user groups?**
- **Semantic level**
  - **can Z-clients and Z-servers and local IR systems preserve and act on meaning of IR tasks**
- **High-level protocol (functional)**
  - **do Z-client and Z-servers support appropriate Z39.50 services for user tasks**
- **Low-level protocol (syntactic)**
  - **do Z-client and Z-servers interchange PDUs according to standard?**

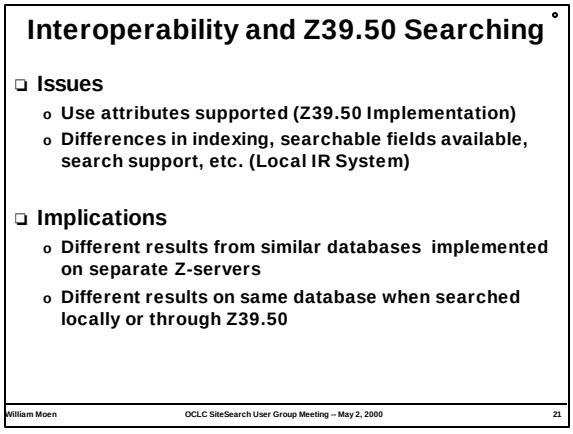
William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 18

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas







Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Interoperability and Z39.50 Retrieval

- **Issues**
  - **Z39.50 Record Syntaxes supported (Z39.50 Implementation)**
  - **Capability of local IR system to prepare records in one or more formats for interchange (Local IR System)**
- **Implications**
  - **Clients and servers may or may not be able to interchange records**

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 22

Problems We Face

- **Strategic use of Z39.50 for integrating access to information resources (local and global)**
- **Dissatisfaction with Z39.50 given the Web**
- **Uncertainty about functional requirements for Z39.50 features and specifications for procuring Z39.50 products**
- **Lack of vendor information/support**
- **Implementation complexity and interoperability problems**

Z39.50 profiles provide solutions

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 23

Z39.50 Profiles

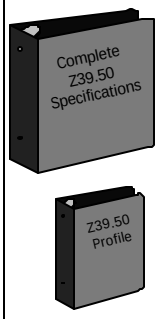
- **Represents community consensus on requirements for Z39.50**
- **Identifies Z39.50 specifications to support those requirements**
- **Improves search and retrieval results**
- **Aids in purchasing decisions**
- **Provides specifications for vendors to build Z39.50 products**

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 24

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Profiles -- Z39.50 Specifications



Z39.50 Profiles

ATS-1

GILS

CIMI

GEO

****

Z Texas

Bath

- ✓ **Represents community consensus on requirements**
- ✓ **Identifies Z39.50 specifications to support those requirements**
- ✓ **Improves search and retrieval results**
- ✓ **Aids in purchasing decisions**
- ✓ **Provides specifications for vendors to build Z39.50 products**

The Z Texas Profile

- **Reflects Texas library requirements & specifications**
- **Who's involved?**
  - **Libraries with Z39.50 implementations**
  - **Libraries acquiring Z39.50 products**
  - **Libraries considering a Z39.50 implementation**
- **Currently representatives from:**
  - **Texas libraries -- public, school, academic, special**
  - **State Institutions/Organizations**
  - **Z39.50 Experts**
- **The Texas Z39.50 Implementors Group (TZIG)**
- **First Initiative: Z Texas: The Z39.50 Profile**

Step 1: Identifying Requirements

- **Providing access -- To what resources?**
- **Types of searching**
- **Interchanging MARC records -- which flavor?**
- **Interlibrary loan, item ordering, document delivery**
- **Full text search and retrieval**

*An **individual effort** to determine requirements **for your** library*

*A **collaborative community effort** to identify **common** requirements for developing Z39.50 specifications*

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Step 2: Specifying Z39.50

- Common specifications for procurement and implementation documented in a profile
- Profile addresses Z specifications such as:
  - Protocol version
  - Searching:
    - Query, Attributes, Attribute Combinations, Default Values, Expected Server Behavior, etc.
  - Retrieval:
    - Record syntax, element set names, etc.
  - Other Z services desired or required
- In synch with other existing/emerging profiles

Gain community (including vendors) support

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 28

Scope and Structure of the Profile

- Scope
  - Current version: search & retrieval from OPACs, including holdings information
  - Subsequent versions: other networked electronic resources
- Structure
  - Modular
  - Current Functional Areas Defined
    - Functional Area -- Library Catalog Search & Retrieval
    - Functional Area -- Retrieval of Bibliographic Holdings Information
  - Potential Functional Areas
    - Cross Domain Searching

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 29

Addressing Interoperability

- Identify searching requirements (tasks)
  - Several categories of searching: basic, specialized
    - Basic: Author, Title, Subject, Keyword, Boolean
    - Specialized: ISBN, Call Number, Specific Controlled Vocabulary
- Defining the searches (semantics and behavior)
  - Exact Title Search -- exact match beginning with first character of first word appearing in the title of an item, in the series title, or the uniform title
- Specifying Z39.50 query to represent the search
  - Standard combination of Z39.50 Attribute Types and Values
- Suggested mapping and indexing of local systems (e.g., MARC fields to index, etc.)

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 30

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Benefits of the Z Texas Profile

- Focuses time and resources to build consensus among Texas librarians
- Ensures Texas libraries' needs are communicated and addressed to Z39.50 vendors
- Educates and informs Texas librarians
- Ensures effective Z39.50 implementation for Texas libraries
- Provides technical infrastructure for statewide resource sharing

The Bath Profile

- ***The Bath Profile: An International Z39.50 Specification for Library Applications and Resource Discovery*** (Draft available)
- International in use and scope
- Based on structure and requirements of Z Texas Profile
- Specifies three Functional Areas:
  - Basic Bibliographic Search & Retrieval
  - Bibliographic Holdings Search & Retrieval
  - Cross-Domain Search & Retrieval
- Specifies several Conformance Levels

Bath Profile Functional Area

- Basic Bibliographic Search and Retrieval
  - Basic Bibliographic Search
    - Level 0 (4 searches)
    - Level 1 (15 searches)
  - Browsing Indexes
    - 6 Scans specified
  - Basic Bibliographic Retrieval
    - Combination of UNIMARC or MARC21 and SUTRS and XML.
- Bibliographic Holdings Search and Retrieval
  - 3 Levels
- Cross-Domain Search and Retrieval
  - Level 0 (4 searches) Level 1 (9 searches)
  - SUTRS and XML as record syntaxes

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Level 0 Title Keyword Search

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

Attribute Type	Attribute Value	Attribute Names
Use (1)	4	Title
Relation (2)	3	Equal
Position (3)	3	Any
Structure (4)	2	Word
Truncation (5)	100	Do not truncate
Completeness (6)	1	Incomplete subfield

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 34

Level 1 Title First Words in Field

Uses: Searches for complete word(s) in the order specified in fields that contain a title of a resource. The field must begin with the specified character string. This search is useful when the beginning words in a title are known to the user.

Attribute Type	Attribute Value	Attribute Names
Use (1)	4	Title
Relation (2)	3	Equal
Position (3)	3	Any
Structure (4)	1	Phrase
Truncation (5)	100	Do not truncate
Completeness (6)	1	Incomplete subfield

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 35

Z39.50 & Dublin Core

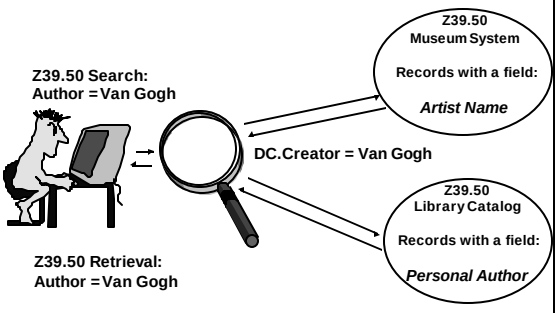
- **Z39.50 enables a common "understanding" by reference to shared vocabulary and structure**
- **Dublin Core can provide the standardized vocabulary for access points in database**
- **Dublin Core can provide the standardized grammar for labeling retrieved data elements**
- **Extensible Markup Language (XML) serves as an interchange structure within Z39.50**

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 36

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Dublin Core as Lens



William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 37

Z39.50 & Dublin Core at Work

- The Bath Profile
  - uses Dublin Core for Resource Discovery to represent access points and label retrieval elements
  - uses XML as the interchange format
- Z39.50 Application Profile for Cultural Heritage Information
  - <http://www.cimi.org/standards/index.html#THREE>
  - uses Dublin Core to represent access points and label retrieval elements

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 38

Synergistic Solutions: Z39.50 and Metadata

- Z39.50 developed to interact with structured metadata in bibliographic databases
- Metadata is now key approach for networked information organization
- Networked environment will see varieties of metadata schemes and implementation models
- Structured metadata + Z39.50 = integrated access to distributed global resources

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 39

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

Z Texas and Bath

- **Z Texas Profile**
  - Z Texas as companion profile to Bath
  - Z Texas as compatible superset of Bath
  - Gaining expert advice from development of Bath
- **Bath Profile**
  - Bath development has slowed work on Z Texas
  - Bath is now in a stable draft version
  - Bath is under review as Internationally Registered Profile

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 40

The Next Steps: Harmonizing Z Texas with Bath

- Upon IRP decision, Bath will be released for use
- Final revisions to Z Texas to harmonize with Bath
- Publication of Z Texas Profile, Release 2.0
- Continuing development on Texas-specific requirements:
  - for searching
  - for retrieval
  - for addressing other types of information
- Developing guidance for implementing Z Texas
- Developing indexing and mapping guidance

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 41

The Next Steps: Assuring Goodness of Products

- **Approaches**
  - Taking the vendor's word for it
  - Conformance testing
  - Interoperability testing
- **Challenges of interoperability testing**
  - Rigorous methodology
  - Tests for different levels of interoperability
  - Metrics for benchmarks and comparison
  - Setting up an interop testbed

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 42

Realizing the Vision:
Improving Interoperability with Z39.50 Profiles

William E. Moen <wemoen@unt.edu>
School of Library and Information Sciences, University of North Texas

An Interoperability Testbed

- **Technical (e.g., for semantic interoperability)**
  - **Reference implementation (trusted)**
  - **Test database (known contents)**
  - **Search and retrieval scenarios**
  - **Known results of searches**
- **Organizational**
  - **Neutral sponsor**
- **Political**
  - **Motivating participation**

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 43

Final Thoughts

- **Convergence of interest, need, motivation**
- **Collaboration within and among communities**
- **Education and marketing**

Window of opportunity for Z39.50 and it's up to us to take advantage of it!

We can realize the vision for our users!

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 44

Z Texas and Bath Profile Resources

- **Z Texas Profile**
  - ***Project Website***
    http://www.tsl.state.tx.us/LD/z3950/
  - ***TZIG Member List***
    http://www.tsl.state.tx.us/LD/z3950/committee.htm
  - ***Texas Z39.50 Listserv***
    http://www.tsl.state.tx.us/LD/z3950/list.htm
- **Bath Profile**
  - ***Copy of draft profile***
    http://www.ukoln.ac.uk/interop-focus/activities/z3950/int_profile/bath/draft/
  - ***Listserv for Discussion***
    ZIP-PIZ-L (subscribe by sending message to:
    LISTSERV@INFOSERV.NLC-BNC.CA)

William Moen OCLC SiteSearch User Group Meeting -- May 2, 2000 45