Z39.50

Recent Developments and Future Prospects

Presented at the September 30, 1996
Z39.50 Seminar at the
Royal Library of Belgium

Ray Denenberg
Library of Congress
ray@rden.loc.gov

October 1996

Historical Background

A brief historical overview of Z39.50 is provided as context for discussion of *Z39.50 recent developments and future prospects*. Although the historical events leading to the development of Z39.50 are sometimes tracked back to the 1960s, momentum to standardize an information retrieval protocol began to sharpen in the early 1980s with the beginning of the Linked Systems Project, LSP, whose implementation began In 1982, and which became operational in 1985. The participants were the Library of Congress, RLG, and OCLC.

The essence of LSP was the Authorities application: the establishment and maintenance of a nationwide database of name authority records. Two application level protocols were developed: Record Transfer and Information Retrieval. The primary function of the authorities application was the transfer of the authority records between systems. supported by the Record Transfer protocol. A background function was the intersystem searching of authority records, supported by the Information Retrieval protocol.

Both the Record Transfer and information retrieval protocols were developed to support authority record exchange, but were intended to support record exchange and intersystem searching regardless of record type.

In 1983 the LSP participants submitted both protocols, Record Transfer and Information Retrieval, for consideration as American National Standards. For Record Transfer, attempts to standardize were eventually abandoned (and ultimately, the Record Transfer protocol itself was replaced by FTP).

There was however substantial interest within the U.S. in standardizing an information retrieval protocol, and the LSP Information Retrieval protocol was submitted to ANSI/NISO, who formed a committee that prepared it for ballot, in 1984, when it was given the designation "Z39.50", as it is known today. (NISO was formerly named Z39, and continues to use that designation for its standards.) The 1984 ballot failed within NISO, for reasons beyond the scope of this paper (primarily because it was not yet sufficiently well-developed). There was significant further development over the next three years; Z39.50 was re-balloted in 1987, this time successfully, and was approved by ANSI in 1988.

Independently, in 1984, a work item was approved in ISO for a "Search and Retrieve" protocol, called SR. There were several drafts of the SR standard between 1984 and 1991 when it was finally approved. As difficult as it was to achieve consensus on Z39.50 in the U.S., it was more difficult to achieve international consensus on SR, because of the various conflicting national interests represented. Of course the U.S. input was influenced by Z39.50, which was not entirely stable during the period of SR development. The result was that several incompatibilities remained between SR and the 1988 version of Z39.50.

The ZIG and Maintenance Agency

In 1990 the *Z39.50 Implementors Group* (ZIG) was established, initially to develop profiles; its role has evolved and now its primary activity is to develop and recommend enhancements to the standard.

Also in 1990 a *Z39.50 Maintenance Agency* was established, at the Library of Congress. In late 1991 the Maintenance Agency put forth version 2 for ballot; it was approved in 1992. Version 2, developed by the Maintenance Agency in collaboration with the ZIG, replaced and superseded the 1988 version.

There were two categories of change in version 2: changes necessary for alignment with SR, and features deemed necessary by implementors, to provide sufficient functionality so that implementation would be economically justified.

There were differences between Z39.50 version 2 and SR. Z39.50-1988 had included two services not initially in SR: access control and resource control. Both were retained in version 2 (resource control underwent substantial revision). These, as well as some other optional features, were carefully incorporated so that an SR and Z39.50 origin/target pair interwork transparently (in other words, the SR implementation is not even aware that its partner is a Z39.50 implementation).

Development of Z39.50-1995 (Version 3)

Many enhancements had been proposed by ZIG members for the 1992 version, to support a wide range of information retrieval capabilities. But those features were not yet fully developed, and their incorporation into the 1992 standard would have caused significant delay. The Z39.50 Maintenance Agency had been assigned, as top priority, to revise Z39.50-1988 to achieve bit-compatibility with the SR. The proposed new features were deferred with a commitment to implementors that development of the required features would proceed, and that the resultant subsequent version would be a compatible superset of the 1992 standard.

Development of Z39.50-1995 began in late 1991. For each meeting of the ZIG, from December 1991 through April 1994, a revised draft was developed by the Z39.50 Maintenance Agency. Each draft underwent careful scrutiny by implementors, and was discussed at length both over the ZIG Internet mail list, and at the ZIG meeting. Comments and discussion for each draft, and agreements reached at each ZIG meeting, were incorporated into the subsequent draft. In April 1994, the ZIG recommended that the draft be finalized.

The 1992 version came to be known as *version 2*, and the 1995 version, *version 3*. (These version designations have specific protocol significance, they do not refer to versions of the *standard*. Z39.50-1992 specifies protocol version 2; Z39.50-1995 specifies protocol versions 2 and 3.)

Although Z39.50-1992 replaced and superseded Z39.50-1988 (and Z39.50-1988 is obsolete) the relationship between Z39.50-1992 and Z39.50-1995 is quite different: Z39.50-1995 is a compatible superset of the 1992 version. An implementor may obtain complete details of version 2 from the Z39.50-1995 document, and build an implementation compatible with Z39.50-1992.

Z39.50-1995 represents a consensus of the ZIG, which has in effect acted in an advisory role to the maintenance agency, in the effort to develop both Z39.50-1992 and the Z39.50-1995.

Progression of SR

Thus the development of Z39.50, until 1992, was strongly influenced by SR. Z39.50-1992, although compatible, was a superset of SR, and the features developed for Z39.50-1995 were not in SR. Beginning in 1992, ISO advanced several proposals to bring SR into compatibility with Z39.50. These included amendments to include Access Control, Resource Control, and Proximity searching, all three of which were Z39.50-1992 features, as well as several Z39.50-1995 features, for example, Sort, Scan, Explain, Segmentation, Concurrent Operations, and Extended Services. In 1994, however, ISO decided that the process of alignment by individual amendments was extremely burdensome, and decided instead to try to adopt the text of Z39.50 verbatim as an ISO standard. Currently, "ISO 23950" is in the process of fast track ballot, ending November 1995.

The Maintenance Agency Web Page

The Library of Congress maintains a Web page (http://lcweb.loc.gov/z3950/agency) for matters pertaining to the maintenance, ongoing development, and implementation of Z39.50. Available via this page are the texts of Z39.50-1992 and Z39.50-1995, definitions of registered objects, extensions, interpretations, and clarifications of the standard, implementor agreements, the Register of Implementors, and Z39.50 profiles. The latter two are described below.

Z39.50 Register of Implementors

The Z39.50 Register of Implementors is maintained by the Z39.50 Maintenance Agency. It is available in HTML and there also is a print version available in postScript.

Currently, nearly 100 companies, institutions, and organizations are registered. These include universities, vendors, consortiums, consultants, government agencies (U.S. federal as well as non-U.S. government agencies, and state government agencies), Information service, national libraries, publishers, research institutions, corporate libraries, and manufacturers.

The registered universities number about fourteen; there are nearly 35 vendors and consultants. Information services include RLG; LEXIS-NEXIS; OCLC, Knight-Ridder, and Chemical Abstracts Service. National libraries of Canada, Sweden, Denmark, UK, and US are registered.

The register is truly international, including participants from Australia, Austria, Canada, Denmark, Germany, Italy, Netherlands, Norway, Sweden, U.K., and U.S.

The register is keyed by organization (i.e. it is not a register of individuals) Each entry includes a contact person, addressing information, and most importantly, a description of the company's Z39.50 product or implementation, either existing or planned. The only criteria for inclusion in the register are: a company must supply some such description; and every registered organization must either update or verify their entry in the register at least once a year.

Z39.50 Profiles

A Z39.50 profile is a set of implementor agreements specifying the use of Z39.50 to support a particular application (for example GILS or WAIS), function (for example author/title/subject searching), community (examples: the museum community, chemists, musicians); or environment (examples: the Internet, North America, Europe).

By "specifying the use" of Z39.50, we mean to select options, subsets, and values of parameters, where these choices are left open in the Z39.50 standard.

Z39.50-1995 is (of necessity) a large standard, rich in functionality. In general an implementation does not support the complete standard, but rather a conforming subset corresponding to specific relevant requirements.

The two primary reasons for Z39.50 profiles are to provide a specification for vendors to build to so that the resulting products will interoperate, and to provide a specification that a customer may reference for procurement.

There are a number of well-developed profiles, as well as several currently under development.

Developed Profiles

GILS profile

GILS, the Government Information Locator Service, is a response to the need for users to identify and locate publicly available Federal information resources. The GILS Profile provides the specifications for the overall GILS application, including the GILS "Core" data elements that comprise a GILS record describing an information resource, and the use of Z39.50 to search and retrieve GILS records.

ATS Profile

The "Author-Title-Subject" profile aims to improve the reliability of Z39.50 search results. When a client requests, for example, an author search, the intent of the ATS profile is that the server will execute the search based on its concept of author. If the server does not support an author search, it should not re-cast the search, substituting some attribute other than author, without the client's knowledge and consent. Neither should the server treat the inability to perform a search as a successful search with no results.

The profile specifies the use of bib-1 within a type-1 query for searching by author, title, or subject, to provide basic search access to bibliographic databases.

WAIS Profile

The WAIS (Wide Area Information Servers) profile specifies rules for access to WAIS servers supporting Z39.50 version 2.

Profiles Under Development

Collections Profile

In August, 1995, the Library of Congress convened a team of representatives from several institutions to develop a Z39.50 profile for access to digital libraries. Participating organizations included Getty, Berkeley, University of Michigan, University of California, OCLC, LC, RLG, Chemical Abstracts Service, IBM, FCLA, TRW, Knight Ridder, SilverPlatter, as well as consultants and liaisons.

The scope was narrowed to apply to navigation of digital collections, and was named the *Z39.50 Profile for Access to Digital Collections* (Collections Profile). The larger problem of access to digital libraries was left to the province of other profiling efforts, including CIMI and the Digital Library Object profiles described below. Other groups were initiating independent efforts to develop profiles aimed at specific types of objects and collections. The intention was to coordinate these efforts and that these latter profiles would be developed as compatible extensions or subsets of the Collections profile.

The profile aims to address the problem faced by libraries and other institutions who create collections, organized thematically -- by subject, creator, historical period, etc.-- with numerous, diverse objects, both digital and physical. These collections are often organized hierarchically and distributed across servers. Significant resources may be invested in digitization and in the intellectual efforts of aggregation, organization, and description of the information in a collection. Yet to a remote user or client, the collection may appear to be simply an accumulation of objects and undifferentiated data, because there is no agreed-upon semantics for navigating the collection, to locate and retrieve objects of interest. Coherent organizational structures, imposed on the data, are necessary to provide a view that supports navigation.

A key obstacle to effective navigation is the inability to distinguish content from description. A primary goal of navigation is to locate and retrieve objects of interest; a vital step in that process is to locate relevant descriptive information. Thus it is useful to navigate among descriptive information as well as content, and consequently, to be able to distinguish content from description.

The profile exploits organizational structures to allow a client to navigate through structured information. A coherently defined set of descriptive data is used to manage and navigate collections of otherwise undifferentiated data. These organizational structures allow the data to be viewed as distributed, hierarchical collections.

The objectives of the profile are to:

- Provide the means for a server to designate what data is content and what is description; and thereby to provide a client, accessing a digital collection to distinguish content from description.
- Provide the means for a server to exploit existing descriptive aids.
- Provide the means for a server to model relationships among collections and objects.
- Provide semantics for navigational queries to be interpreted consistently across servers.
- Define data structures that provide semantics for navigation, to locate and retrieve objects of interest.

The profile models objects and collections: a collection is a group of related objects and/or collections. Thus it is a tree, where leaf nodes are objects and non-leaf nodes are subcollections. The profile defines structures that allow a client to navigate among superior, subordinate, related, and context collections.

As an example, suppose a client connects to the Library of Congress home page. It might provide access to the following *root* collections: "Exhibits", "Research", "Copyright", and "Government, Congress, and Laws". The latter might be split into three subcollections: "Government", "Congress", and "Laws". Subordinate to the "Congress" collection might be the two collections "Bills" and "Congressional Record". The latter might be partitioned into collections according to the congress number (for example the "102nd Congress: 1991-1992"). Furthermore, each congress might have subordinate collections for "House Resolutions", "House Bills", "Senate Resolutions", and "Senate bills". On the other hand, all "resolutions" (House and Senate) may be aggregated into a single collection and all Bills into another. And in fact, all of these different types of aggregations may co-exist, so that a client is not constrained to navigate along a strictly hierarchical path. Finally, at the leafs of the tree, are actual digital objects, which might be the electronic texts of the bills or resolutions.

As another example, Subordinate to the root collection "Exhibits" mentioned above, may be the "Federal Theater Project" collection. Consider a sub-collection of digitized theater costumes, in turn with various sub-collections. One of the objects is a digitization of an eighteenth century, yellow, Shakespearean costume. That single object might belong to a number of collections, including "18th century costumes", "yellow costumes" and "shakespearean costumes". A user interested in that object might navigate to the object via any of the three collections, based on the specific aspect of interest.

CIMI Profile

The Consortium for the Computer Interchange of Museum Information (CIMI) has supported the development of a Z39.50 Profile as part of its current Project CHIO (Cultural Heritage Information Online), for access to museum information.

Museum information includes a variety of physical and electronic objects, including physical artifacts and electronic derivatives, descriptive records designed for collection management, full-text documents, and online tools such as thesauri and authoritative lists of artists' names.

A digital collection of museum information needs to address not only the heterogeneous nature of the information objects but also the fact that such a collection will draw upon repositories of museum information distributed around the world.

CIMI initiated Project CHIO as a demonstration project to investigate a standards-based approach for searching and retrieving cultural heritage information from disparate and distributed information systems containing museum information. Project CHIO consists of two interrelated demonstration projects -- CHIO Structure and CHIO Access -- to show respectively the utility of SGML and Z39.50, to enhance electronic access to cultural heritage museum information in a distributed, networked environment.

Museum information includes physical and electronic objects -- physical artifacts and electronic derivatives of those artifacts, descriptive records designed for collection management, full-text documents, online tools such as thesauri and authoritative lists of artists' names, and more.

CIMI initiated Project CHIO as a demonstration project to investigate a standards-based approach for searching and retrieving cultural heritage information from distributed information systems containing museum information. Project CHIO consists of two interrelated demonstration projects -- CHIO Structure and CHIO Access -- to show respectively the utility of SGML and Z39.50, to enhance electronic access to cultural heritage museum information in a distributed, networked environment.

"CHIO Structure" uses SGML to mark up museum objects including (text) exhibition catalogues and wall text, and make them available for electronic access. "CHIO Access" demonstrates the utility of Z39.50 to access digitized museum objects.

Digital Library Objects

The *Z39.50 Profile for Access to Digital Library Objects* (DL Profile) addresses functional and user requirements for search and retrieval of information in digital library collections, specifically the Library of Congress digital library collections and similar collections.

The profile provides a general and flexible model for the structure of a digital object. In the model, a digital object may consist of constituent parts, any of which may in turn consist of constituent parts, and so on. Consider, for example, a single digital object consisting of several images (e.g. photos or text images). Although the set of images comprises a single digital object, each must be distinctly representable and the object must convey the fact that there are distinct images, how many, and their individual characteristics. Thus they are represented as separate elements of a Z39.50 record.

Next suppose that the digital object not only includes a number of images, but also additional constituent parts, further structured; for example, each such constituent part may consist of several images. This introduces an intermediate level of aggregation. The model of a digital object adopted by the DL profile assumes arbitrary levels of aggregation and is represented as a tree, where each non-leaf node has an arbitrary number of subtrees and/or leaves, and leaf nodes represent data.

Every node, whether a leaf or non-leaf node, may have metadata attached, including description, date of creation, terms and conditions, etc.

This model will support, for example, a digital object representing 10 boxes, each with 20 folders, each with 30 photos. Z39.50 string tags such as 'box', 'folder', and 'photo' could be used to convey the type of element. As a more complex example, a folder might include a variety of photos, maps, correspondences, etc. and perhaps the correspondences consist of several sequential digitized pages.

CIP Profile

CIP - the Catalogue Interoperability Protocol - addressed the ability to effectively exploit earth observation and associated data resources. That capability is impeded by the lack of homogeneity in services and interfaces offered by various data providers. CIP is being developed by the Protocol Task Team within the Committee on Earth Observation Satellites (CEOS).

CEOS provides coordination between international Earth observation missions and encompasses various national (civil) agencies involved in Earth Observation satellite programmes: the European Space Agency, NASA, DLR (Germany), NASDA (Japan), DDRS (Canada), BNSC (UK), and CEO (Centre for Earth Observation).

The objective of CIP is to enable users to logically search physically distributed data catalogues, without separately querying each and merging/correlating result sets, effectively allowing the various data archives to appear to be a single database. It includes a data dictionary to specify the common attributes that describe the primary objects within a catalogue system.

CIP models collections, permitting complex hierarchical groupings of data organized thematically over multiple databases, where both the collections and the individual collection members (objects and subcollections) have item descriptors, roughly analogous to the descriptive records defined by the Collections profile.

Cataloging Profile

A service named WORLD 1 will be offered by the National Library of Australia to replace the current Australian Bibliographic Network and Ozline services. The technical infrastructure to operate the WORLD 1 Service is being developed as a joint venture by the National Library of Australia and the National Library of New Zealand under the banner of the National Document and Information System (NDIS) Project.

The plan is to use union catalogues as tools for the identification of resources and their location, in a geographic area. The premise is that union catalogues with good coverage and authority control are still an attractive concept because of the limitations of multi-target searches, with performance degradation (for searches over several targets), where results are not well integrated, with duplicate records, and multiple versions of headings (e.g. author and subject).

Libraries contributing to a union catalog would require a cataloging system to update both their own local catalog and the union catalog in a single operation, and the project proposes to integrate the "cataloguing protocol" with Z39.50. To this end, they propose to use Z39.50 both for search and update, and they are profiling the Z39.50 Update Extended Service.

ZSTARTS

The *Z39.50 Profile for STARTS* (ZSTARTS) stems from the *Stanford Protocol for Internet Search and Retrieval* (STARTS), an initiative of the Stanford Digital Library Project. The STARTS project brought together a number of commercial companies to develop requirements for distributed searching and ranked retrieval. The ZSTARTS profile is a Z39.50 solution to these requirements.

The STARTS model assumes document databases; a client sends a query to multiple servers, where the query includes a filter and ranking expression. The filter is analogous to the Z39.50 type-1 query (i.e. a boolean query); while the ranking expression supplies guidance for the server to rank results -- the client may assign weights to individual terms. The STARTS model calls for the merging of the ranked results from the various servers.

Search results include document metadata: title, publication date, size, score (assigned to the document for the given search), occurrence information (pertaining to the terms in the query) and a pointer (url) to the document for subsequent retrieval.

Proposed New Z39.50 Query Types

Z39.50 defines several query types. The type-1, RPN query (named for its "reverse polish notation" structure) is the primary Z39.50 query, and its support is mandatory for conformance. A type-0, "private" query, is defined, for partners who have a private agreement about its usage. There are also type-2 and type-100, corresponding to ISO 8777 and Z39.58, respectively (both address common command language). There is also a type-101, which, in version 2, was an extension to type-1, adding proximity searching, as well as new feature called "result set restriction". In version 3, these type-101 extensions were adopted into type-1, making these two queries identical.

Two new query types are currently under development, type-102 and (tentatively) type-103.

Type-102 RLQ

The Type 102 Ranked List Query (RLQ) was originally intended to be developed as a natural language query, but it was deemed impossible to design a query that adequately supports all of the natural language search methodologies. Type 102 RLQ has instead been designed for the ranked searching technologies used by large-scale commercial information providers and information industry software vendors, several of whom have participated in the development of this query, including:

- Chemical Abstracts Service
- Clearinghouse for Networked Information Discovery and Retrieval
- Excalibur Technologies Corporation
- Knight-Ridder Information Services
- Lexis-Nexis
- National Library of Medicine
- West Publishing Company

RL Query features include:

- Results Ranking

  The client provides the server with information about the specific type of ranking it needs.
- User/client hints

  The client indicates which components of the query are most significant to the user, including:
    - weighting of terms and operators
    - ranking operators
    - query reformulation options
- Relevance Feedback

  The client 'seed's the server's search process by indicating records which are either precisely on-point or are totally off-the-mark.
- Restriction of search scope

  The client can restrict the set of records which are eligible for input to the ranked search process (analogous to the ZSTARTS ranking component discussed above).
- Iterative Query Reformulation

  The client may request the server to expand elements of the query based on the need for additional precision or recall (e.g. thesaural or morphological expansions). The client may then ask the server to return the reformulated query so that user can inspect and modify it before it is executed (or reformulated again).
- Precision vs. Recall control

Proposed SQL Query

The Distributed Database Unit at CRC for Distributed Systems Technology Centre at the Department of Computer Science, University of Queensland, is proposing changes to Z39.50 to support SQL databases. Their proposal includes:

- A new query type based on SQL-92.
- Addition to the Explain Syntax.
- Two supporting record syntaxes.

The proposal, named "Z39.50/SQL+", is seen to combine the advantages of Z39.50 and SQL: a stateful communication environment and the flexibility and query power of SQL including queries on multiple tables supporting cartesian products, unions, intersections, joins on matching columns, and projections on given columns, ability to use constructs for expressing conditions, performing aggregate and comparison operations, and partitioning tables into groups.

Attribute Architecture

Attributes define and qualify access points used in the construction of Z39.50 queries. Originally, there was a single attribute set in use, the "bib-1" attribute set, for searching bibliographic databases, and bib-1 was sufficient for most version 2 applications. Additional attribute sets have now been defined for non-bibliographic databases, for example, for scientific or geo-spatial information.

As a result of the growth of Z39.50 over the past several years, the proliferation of Z39.50 attribute sets has produced interoperability and maintenance problems.

One problem is the duplication of attributes. A severe limitation of version 2 (corrected in version 3) is that a query may reference only a single attribute set. Yet it is often desirable to construct a query combining access points defined in different sets (or to qualify a given access point with attributes from different sets). The result has been that newly defined attribute sets often copy existing attributes from other sets. In fact, some attribute sets seem to include most or all attributes from all other sets. The result, paradoxically, is the proliferation of a single "universal" attribute set, with many identifiers, each incorporating new attributes from other sets as they are added. This causes both administrative and interoperability problems. Unless a server maintains redundant tables to account for the different sets, it likely will receive queries with attribute sets it does not support, even though those queries are composed of well-known, familiar attributes.

Another problem is semantic ambiguity. Server behavior, for example, when attributes are repeated or omitted, is not well specified. Semantics of queries where attributes are intermixed from different sets is not well-understood. When an attribute from one set is imported into another set, the semantics of that attribute may change. (The semantic ambiguity problem is not limited to Z39.50 growth; there are also semantic ambiguities specific to the bib-1 attribute set.)

These are just a few examples of problems with Z39.50 attributes. To address these problems, an effort to develop a new "attribute architecture" has been initiated. Compatibility and interoperability with version 2 (specifically, its limitation of a single attribute set per query) will not confine the architecture; version 3 will be assumed. When the new architecture is in place, attribute sets will be developed by groups with content expertise, not by the ZIG (with some exceptions, such as special purpose attribute sets such as Explain and Extended Services). The ZIG will, of course, provide broad architectural guidance about how to structure, register, manage and maintain an attribute set.