

SRW and CQL

Overview:

- ◆ Motivations
- ◆ Administration
- ◆ Explain Operation
- ◆ ZeeRex and Thin Clients
- ◆ SearchRetrieve Operation
- ◆ State, Records, Identifiers
- ◆ Scan Operation
- ◆ Diagnostics
- ◆ Extensibility

SRW/U: Introduction

SRW/U is an Information Retrieval Web Service.

Motivations:

- ◆ Create an easy to implement protocol with the power of Z39.50
- ◆ Need for a generic IR web service
- ◆ Use existing off the shelf solutions where possible
- ◆ Re-evaluate Z39.50 “a good idea at the time” baggage
- ◆ Avoid library-centric preconception

Information:

<http://www.loc.gov/srw/>

<http://www.loc.gov/cql/>

<http://explain.z3950.org/>

SRW/CQL: Administration

Standards Body: NISO (Or ISO as profile of Z39.50)

Maintenance Agency: Library of Congress

Editorial Board:

Managing Editor: Ray Denenberg

Senior Editor: Rob Sanderson

Technical Editor: Matthew Dovey

Plus about eight others from Australia, US and Europe.

All additions for future versions will first require implementation as extensions to prevent creeping bloat.

SRW/U: Protocol Overview

SRW is the SOAP binding. (POST of XML request and response)
SRU is a REST-like binding. (GET parameters for XML response)

Current Version: 1.1

Related Specifications:

- ◆ CQL (Common Query Language)
- ◆ ZeeRex (Z39.50 Explain Explained and Re-Engineered in XML)

Three Operations Available:

- ◆ Explain
- ◆ SearchRetrieve
- ◆ Scan

SRW/U: Explain Operation

Purpose: Retrieve Service Description

Request Parameters:

- ◆ Version '1.1'
- ◆ RecordPacking 'string' or 'xml'
- ◆ Stylesheet URI
- ◆ ExtraRequestData Namespaced XML

Response Parameters:

- ◆ Version '1.1'
- ◆ Record Record (ZeeRex schema)
- ◆ Diagnostics Sequence of Diagnostic
- ◆ ExtraResponseData Namespaced XML
- ◆ EchoedExplainRequest Request

ZeeRex is semantic, WSDL is syntactic.

SRW/U: New Parameters in 1.1

Version (all messages):

Request: Version of the request and indication that the client wants the response's version to be less than, or preferably equal to, that version.

Response: version of the response.

Stylesheet (all requests):

URL to a stylesheet to include in the response.

Extra{Request|Response}Data (all messages):

Space for extensions. Extensions must be requested, and requests may be silently ignored. Carried as embedded, namespaced XML in SRW, or on the URL with an 'x-' prefix for SRU requests.

Echoed*Request (all responses):

Thin clients may not even know what they just asked for.

SRW/U: Explain Record

ZeeRex Service Description Schema

Six sections to a record:

- ◆ ServerInfo: Host, port, database path
- ◆ DatabaseInfo: Title, description, access rights, etc.
- ◆ MetaInfo: Updates, aggregation
- ◆ IndexInfo: CQL contextSets and indexes
- ◆ SchemaInfo: Available record schemas
- ◆ ConfigInfo: Settings, supports and defaults

Can be (and is already) used to automatically configure a client, in the way that Z39.50's Explain service was supposed to work. Even simple XSLT transformations can be used to generate a usable HTML interface (a 'dumb client').

Website: <http://explain.z3950.org/>

SRW/U: XSLT Thin Client

The image shows a Mozilla browser window displaying the 'Legend of the Five Rings Database' website. The browser's address bar shows the URL 'http://srw.cheshire3.org:8080/15r'. The website content includes a search interface with a table of search criteria and a 'Browse' section. An 'XSLT Thin Client' window is overlaid on the right, showing the source code of the page, which is an XML document with XSLT instructions for displaying the database information.

Legend of the Five Rings Database

A database containing complete descriptions of cards from the L5R CCG.

Search

| Index | Relation | Term | Boolean |
|--------------------|----------|------|---------|
| Anywhere | = | | and |
| Card Name | = | | and |
| Card Type | = | | and |
| Card Text | = | | and |
| Keyword | = | | and |
| Card Flavor | = | | and |
| Artist | = | | and |
| Set | = | | and |
| Legality | = | | and |
| Rarity | = | | and |
| Location of Card | = | | and |
| Number at Location | = | | and |
| Force | = | | and |
| Chi | = | | and |
| Experience Level | = | | and |
| Focus Value | = | | and |
| Province Strength | = | | and |
| Personal Honor | = | | and |
| Gold Cost | = | | and |
| Honor Requirement | = | | and |

Browse

Browse: Any

Response Pos: 1

Maximum T: 1

Record Schema: Dublin Core

Number of Records: 1

Record Position: 1

Done

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="/xsl/srw.xsl"?>
<srw:explainResponse xmlns:srw="http://www.loc.gov/zing/srw/" xmlns:diag="http://www.loc.gov/zing/srw/"
<srw:version>1.1</srw:version>
<srw:record>
<srw:recordSchema>http://explain.z3950.org/dtd/2.0</srw:recordSchema>
<srw:recordPacking>xml</srw:recordPacking>
<srw:recordData>
<explain id="org.o-r-g.srw-card" authoritative="true" xmlns="http://explain.z3950.org/dtd/2.0/">
  <serverInfo protocol="srw/u" version="1.1" transport="http">
    <host>srw.cheshire3.org</host>
    <port>8080</port>
    <database numRecs="3492" lastUpdate="2002-11-26 23:30:00">15r</database>
  </serverInfo>
  <databaseInfo>
    <title lang="en" primary="true">Legend of the Five Rings Database</title>
    <description lang="en" primary="true">
      A database containing complete descriptions of cards from the L5R CCG.
    </description>
  </databaseInfo>
  <metaInfo>
    <dateModified>2002-11-29 12:00:00</dateModified>
  </metaInfo>
  <indexInfo>
    <set identifier="http://srw.o-r-g.org/indexSets/ccg/1.0/" name="ccg"/>
    <set identifier="http://srw.o-r-g.org/indexSets/ccg/15r/1.0/" name="ccg_15r"/>
    <set identifier="info:srw/cql-context-set/1/dc-v1.1" name="dc"/>
    <index id="DISTRIB-15r-1">
      <title>Anywhere</title>
      <map><name set="cql">anywhere</name></map>
      <map><name set="cql">serverChoice</name></map>
    </index>
    <index id="DISTRIB-15r-5">
      <title>Card Name</title>
      <map><name set="dc">title</name></map>
    </index>
    <index id="DISTRIB-15r-9">
      <title>Card Type</title>
      <map><name set="dc">type</name></map>
    </index>
    <index id="DISTRIB-15r-8">
      <title>Card Text</title>
      <map><name set="ccg">text</name></map>
    </index>
    <index id="DISTRIB-15r-20">
      <title>Keyword</title>
      <map><name set="ccg">keyword</name></map>
    </index>
  </indexInfo>
</explain>
</srw:recordData>
</srw:record>
</srw:explainResponse>
```


SRW/U: SearchRetrieve Request

Purpose: Find, Sort and Display Records

Request Parameters:

- ◆ Version '1.1'
- ◆ Query CQL Query (string)
- ◆ RecordSchema URI Schema Identifier or alias
- ◆ RecordPacking 'string' or 'xml'
- ◆ RecordXPath XPath Expression (string)
- ◆ StartRecord Positive Integer
- ◆ MaximumRecords Non-negative Integer
- ◆ ResultSetTTL Non-negative Integer
- ◆ SortKeys Sort Expression (string)
- ◆ Stylesheet URI
- ◆ ExtraRequestData Namespaced XML data

SRW/U: New Parameters in 1.1

RecordPacking:

Either embed the record or return as an encoded string. Raw XML is convenient for thin clients, but can be problematic in SOAP, particularly if it is not well-formed.

RecordXPath:

An XPath expression to select nodes from the record to return. Useful for paging through long records (EAD, X3D, SVG), or selecting a known section (title/author display). Returned in an XML encoded form.

ResultSetTTL:

Servers need to know if the client wants a result set to be maintained or not. This parameter allows the client to request the set be maintained for a certain time.

SRW/U: Result Sets and State

State:

Z39.50 has session state due to a persistent connection.
HTTP does not have a persistent connection between requests.

State is optional in SRW as it is only necessary for persistent result sets. These result sets might persist between connections in Z39.50 as well. Not every SRW service will require persistent result sets, e.g. an SRW interface to Google.

Result Sets:

A result set is an ordered list of records. Each result set has a unique identifier and a time to live. It is named by the server and cannot change once created. For example, if a result set is sorted, it will receive a new identifier.

SRW/U: SearchRetrieve Response

Purpose: Find and Display Records

Response Parameters:

- | | |
|-------------------------------|------------------------|
| ◆ Version | '1.1' |
| ◆ NumberOfRecords | Non-negative Integer |
| ◆ ResultSetId | String |
| ◆ ResultSetIdleTime | Non-negative Integer |
| ◆ Records | Sequence of Record |
| ◆ NextRecordPosition | Non-negative Integer |
| ◆ Diagnostics | Sequence of Diagnostic |
| ◆ ExtraResponseData | Namespaced XML |
| ◆ EchoedSearchRetrieveRequest | Request + XCQL |

SRW/U: Records

Records are transferred in SRW in a simple structure:

- ◆ RecordSchema URI identifier for the schema
- ◆ RecordPacking 'string' or 'xml'
- ◆ RecordData The actual record
- ◆ RecordPosition Record's position in the resultSet
- ◆ ExtraRecordData Space for further metadata

Like Z39.50, SRW does not assume a record structure in the database, but all records must be transferred in an XML format, either raw or string encoded. A list of registered schema identifiers is available, but any schema may be used.

SRW/U: URIs as Identifiers

We need to know which XML schema to use for the response. Can't use namespace (multiple namespaces per schema) or schema location (non unique). The solution is to use URIs as Identifiers.

'info' URI scheme was developed in the OAI/OpenURL world.

SRW info URIs have the form:

`info:srw/<type>/<owner>/<identifier>`

e.g. `info:srw/cql-context-set/1/cql-v1.1`

Identifiers are used for record schemas, context sets, diagnostics, extensions and profiles. They can be info, but many are also regular http URIs.

e.g. `http://srw.cheshire3.org/contextSets/rec/1.1/`

SRW/U: Scan Operation

Purpose: Retrieve Ordered List of Searchable Terms

Request Parameters:

- ◆ Version '1.1'
- ◆ ScanClause CQL Search Clause (string)
- ◆ ResponsePosition Non-negative Integer
- ◆ MaximumTerms Non-negative Integer
- ◆ Stylesheet URI
- ◆ ExtraRequestData Namespaced XML

Response Parameters:

- ◆ Version '1.1'
- ◆ Terms Sequence of Term
- ◆ Diagnostics Sequence of Diagnostic
- ◆ ExtraResponseData Namespaced XML
- ◆ EchoedScanRequest Request

SRW/U: Terms

Scan Term Structure:

- ◆ Value String
- ◆ NumberOfRecords Non-negative Integer
- ◆ DisplayTerm String
- ◆ WhereInList 'first', 'last', 'only' or 'inner'
- ◆ ExtraTermData Namespaced XML

The Scan operation is practically identical to the Z39.50 operation.

Value is the indexed term to be used in a search, as opposed to DisplayTerm is an alternative form to be used to display to the end user. WhereInList carries the location of the term within the server's list of terms, so that the beginning and end of a list can be flagged.

SRW/U: Diagnostics

Sometimes things go wrong.

Diagnostic Structure:

- ◆ Uri URI diagnostic code and diagnostic set
- ◆ Details Machine readable or debugging string
- ◆ Message Human readable string

Both fatal and non-fatal, surrogate and non-surrogate.
URI is new in 1.1, and allows user communities to define their own profiled diagnostics.

Mappings available both to and from Z39.50's diagnostics.

SRW/CQL: Extensibility

Operations:

- ◆ ExtraRequestData/ExtraResponseData
- ◆ SOAP/HTTP Headers
- ◆ New Operations

Individual Entries:

- ◆ ExtraRecordData/ExtraTermData
- ◆ Diagnostics
- ◆ Record Schemas

CQL Context Sets:

- ◆ Indexes
- ◆ Relations
- ◆ Relation/Boolean Modifiers