CATALOGUE SERVICES LITE - INSPIRE PROFILE

Jan RŮŽIČKA

Institute of geoinformatics, VSB-TU of Ostrava, 17. listopadu 15,
708 33, Ostrava – Poruba, Czech Republic
jan.ruzicka@vsb.cz

Abstract

A very important part of the architecture for GeoWeb services orchestration is a catalogue. It does not matter what kind of a catalogue is used and what type of technology is used for metadata storage. The most important part of the catalogue are two items. The first is a used standard for the description of GeoWeb services and the second one is a standard used for building a catalogue's interface. Interface CSW (Catalogue Services of Web), has shown itself as too complicated for implementation. We have decided to use CSW as a base for building a less-complicated standard for the catalogue used in the architecture for orchestration. In the last year we have published CSW Lite standard. CSW Lite is easy to implement and have all the necessary parts for the architecture of orchestration. CSW Lite is compatible with complete CSW specification. CSW Lite is based on Dublin Core standard for metadata. This seems to be not applicable for several implementations mainly based on INSPIRE directive. We have prepared profile of CSW Lite, that is based on INSPIRE metadata profile. The paper should describe this standard in a brief, but complex way.

Keywords: catalogue, CSW Lite, CSW-L, GeoWeb, INSPIRE, metadata, monitoring

INTRODUCTION

The company that would like to have a CSW compatible catalogue has three possible ways of how to do this. In the commercial sphere, there are several systems that implement CSW specification. Most of them are part of general software solutions that are very expensive. There are only two known system in the area of open source software (GeoNetwork Open Source and Deegree). The third way is based on one's own CSW implementation. All of the options bring challenges. The reason why there are not enough implementations in this area (or they are costly) and why the last way is very difficult for a typical organization is the complexity of the CSW specification. CSW Lite (CSW-L) is a reaction to a quite unnoticed situation in the area of catalogue services for geodata and geoweb services. CSW-L should be a simple specification that could be implemented by a single programmer with basic knowledge of the standards and technologies. CSW-L is compatible with the full CSW 2.0.2 specification and it uses only a minimal set of capabilities of CSW 2.0.2. This simplification should make the CSW more understandable to a general developer. CSW-L is available at [OGCLite 2010]. What is missing in the CSW-L specification is a possibility to encode metadata according to INSPIRE metadata implementation rules. That's why I have created CSW-L INSPIRE Profile implementation specification. The specification is not fully available in this paper, but referenced to external documents at the web [OGCLite 2010].

DESCRIPTION OF GENERAL CSW-L INSPIRE PROFILE CAPABILITIES

Metadata records are only based on ISO 19115 and ISO 19119 standards.

Communication is based only on HTTP protocol. To make this specification simple enough, other protocols are not supported.

Supported operations are basic methods for catalogue discovery (GetCapabilities, GetRecords, DescribeRecord and GetRecordById). A distributed search is not supported.

GETCAPABILITIES OPERATION

The GetCapabilities operation [OGC 2009a, page 131] is used as the initial handshake between a client and a service. The service responds with a document that contains a description of the service capabilities (supported operations, formats, etc.).

Request

The GetCapabilities operation is requested via the HTTP GET method with three encoded parameters using KVP (Key-Value Pairs). Values of the parameters are fixed as is shown in the following example.

```
http://gis.vsb.cz/catalogue/cswl?
REQUEST=GetCapabilites&SERVICE=CSW&VERSION=2.0.2
```

Response

The response is an XML formatted document as is described in the following example. Values marked by -- should be modified, the others should stay untouched. This document specifies which part of the CSW capabilities must be implemented.

```
<?xml version="1.0" encoding="UTF-8"?>
<csw:Capabilities version="2.0.2" …
  Because of space limitations for a paper. The normative example is available
at http://www.ogclite.org/cswlite/inspire/capabilities.xml
</csw:Capabilities>
```

DESCRIBERECORD OPERATION

The DescribeRecord operation [OGC 2009a, page 134] returns the description of a metadata record returned in a response to GetRecords or GetRecordById operations.

Request

The request is XML encoded and transferred via the HTTP POST method. The request is encoded as is described in the following example.

```
<?xml version="1.0" encoding="UTF-8"?>
<DescribeRecord service="CSW"  version="2.0.2" outputFormat="application/xml"
   schemaLanguage="http://www.w3.org/2001/XMLSchema"
   xmlns="http://www.opengis.net/cat/csw/2.0.2">
   <TypeName>gmd:MD_Metadata</TypeName>
</DescribeRecord>
```

Response

The response for this operation is an XML schema that describes (specifies) the structure of the metadata record. The CSW-L compatible service must return document in the following structure.

```
<?xml version="1.0" encoding="UTF-8"?>
<DescribeRecordResponse
   xmlns="http://www.opengis.net/cat/csw/2.0.2">
 <SchemaComponent>content of http://www.isotc211.org/2005/gmd/identification.xsd
file</SchemaComponent>
 <SchemaComponent>content of
http://schemas.opengis.net/iso/19139/20060504/srv/serviceMetadata.xsd
file</SchemaComponent>
</DescribeRecordResponse>
```

There are two elements SchemaComponent, that must contatin content of files specified in the example.

GETRECORDS OPERATION

The GetRecords operation [OGC 2009a, page 142] returns metadata records according to specified conditions in a request.

Request

The request is XML encoded and delivered via the HTTP POST method. The request consists of four parts:

- the GetRecords element

- the Query element

- the ElementSetName element

- the Constraint element

The structure of a request is described in the following example.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<GetRecords
   service="CSW"
   version="2.0.2"
   maxRecords="5"
   startPosition="1"
   resultType="results"
   xmlns="http://www.opengis.net/cat/csw/2.0.2"
   xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
   xmlns:ogc="http://www.opengis.net/ogc"
   xmlns:gml="http://www.opengis.net/gml">
   <Query typeNames="gmd:MD_Metadata">
      <ElementSetName typeNames="gmd:MD_Metadata">full</ElementSetName>
      <Constraint version="1.1.0">
         <ogc:Filter>
            <ogc:And>
              <ogc:PropertyIsLike escapeChar="\" singleChar="?" wildCard="*">
                 <ogc:PropertyName>gmd:identificationInfo/gmd:MD_DataIdentifica
tion/gmd:citation/gmd:CI_Citation/gmd:title/gco:CharacterString</ogc:PropertyNam
e>
                 <ogc:Literal>*Elevation*</ogc:Literal>
              </ogc:PropertyIsLike>
              <ogc:PropertyIsEqualTo>
                 <ogc:PropertyName>gmd:hierarchyLevel/gmd:MD_ScopeCode</ogc:Pro
pertyName>
                 <ogc:Literal>service</ogc:Literal>
              </ogc:PropertyIsEqualTo>
              <ogc:PropertyIsGreaterThanOrEqualTo>
                 <ogc:PropertyName>gmd:identificationInfo/gmd:MD_DataIdentifica
tion/gmd:citation/gmd:CI_Citation/gmd:date/gmd:CI_Date/gmd:date/gco:Date[../../g
md:dateType/gmd:CI_DateTypeCode = 'publication']</ogc:PropertyName>
                 <ogc:Literal>2004-03-15</ogc:Literal>
              </ogc:PropertyIsGreaterThanOrEqualTo>
              <ogc:BBOX>
                 <ogc:PropertyName>gmd:identificationInfo/gmd:MD_DataIdentifica
tion/gmd:extent/gmd:EX_Extent/gmd:geographicElement/gmd:EX_GeographicBoundingBox
</ogc:PropertyName>
                 <gml:Envelope>
                    <gml:lowerCorner>14.05 46.46</gml:lowerCorner>
                    <gml:upperCorner>17.24 48.42</gml:upperCorner>
                 </gml:Envelope>
              </ogc:BBOX>
            </ogc:And>
         </ogc:Filter>
      </Constraint>
   </Query>
</GetRecords>
```

The GetRecords Element

The CSW-L compatible service must implement five attributes that are described in the following table.

Table 1. Attributes for the GetRecords element

Attribute	Description
service	Fixed value "CSW"
version	Fixed value "2.0.2"
maxRecords	Maximal count of returned records. Default is 10. 0 means resultType = hits.
startPosition	Index of the first returned record. Starts at 1.
resultType	Hits or results. Option hits must return an empty result set with an identified number of records.

The Query Element

The Query element is fixed as is shown in the previous example.

The ElementSetName Element

The element ElementSetName contains the fixed attribute typeNames as is shown in the previous example. There are three options of the possible value; they are described in the following table.

Table 2. Values for the ElementSetName element

Value	Description
full	Returns a full set of elements for each record in the response.
summary	Returns a selected set of elements described in the Response subchapter for each record in the response.
brief	Returns a selected set of elements described in the Response subchapter for each record in the response.

The Constraint Element

The Constraint element is based on the OGC Filter Implementation Specification [OGC 2009a, page 152]. It has one fixed attribute and one fixed element ogc:Filter. Specified conditions for searching in a catalogue are found inside ogc:Filter . The structure is described in the previous example. The CSW-L compatible service must support operators described in the following table.

Table 3. Mandatory constraint operators

	Description
And	Conditions can be in a logical And relationship.
Or	Conditions can be in a logical Or relationship.
Not	Conditions can be negated.
PropertyIsEqualTo	Simple equal (==) comparison.
PropertyIsNotEqualTo	Simple not-equal (!=) comparison.
PropertyIsLessThan	Simple lest than (<) comparison.
PropertyIsGreaterThan	Simple greater than (>) comparison.
PropertyIsLessThanOrEqualTo	Simple less than or equal to (<=) comparison.
PropertyIsGreaterThanOrEqualTo	Simple greater than or equal to (>=) comparison.
PropertyIsLike	Like operator including information for: • escaping special characters such as quotes (attribute

	escapeChar),
	● wild-card for one letter (attribute `singleChar`),
	● wild-card for more than one letter (attribute `wildCard`).
BBOX	Spatial operator that selects all overlapping envelopes.

Properties can be specified from all metadata elements listed in the Response subchapter.

Response

The response structure [OGC 2009a, page 120] depends on two parameters in the request resultType and ElementSetName. When the requested resultType is hit, only then a count of records is returned. When the requested resultType is results, then metadata records are returned.

Structure of each record depends on the ElementSetName parameter. The richest record is returned when the value is full.

Response on hits

Response on resultType=hits returns basic information about the search result, but without any records included. The structure of the response is described in the following example.

```
<?xml version="1.0" encoding="UTF-8"?>
<csw:GetRecordsResponse xmlns="http://www.opengis.net/cat/csw/2.0.2"
   xmlns:csw="http://www.opengis.net/cat/csw/2.0.2">
   <csw:SearchStatus timestamp="2009-10-04T09:05:25"/>
   <csw:SearchResults
       nextRecord="0"
       numberOfRecordsMatched="5"
       numberOfRecordsReturned="5"
   />
</csw:GetRecordsResponse>
```

All elements and attributes included in the example are mandatory. The following table describes attributes.

Table 4. Attributes for response on hits

Value	Description
timestamp	Date and time when the response was created.
nextRecord	Index of the first record from the search result that was not included in the response because of some limits (data size) of the communication. Zero means that all records have been returned in the response.
numberOfRecordsMatched	Number of records that have been matched by specified search criteria.
numberOfRecordsReturned	Number of records that have been returned in the response.

Response on full

Response on resultType=results and ElementSetName=full is the same as response on resultType=hits described above, but it includes records (elements Record) in the SearchResults element. The structure of the Record element is described in the following example. All elements are mandatory. If the information for the element is not available, it must be filled with NULL.

```
<gmd:MD_Metadata ... >
```

```
   Because of space limitations for a paper. The normative example is available
at http://www.ogclite.org/cswlite/inspire/dataset.xml
</gmd:MD_Metadata>
```

Response on summary

Response on resultType=results and ElementSetName=summary is the same as response on ElementSetName=full described above, but the record element is named SummaryRecord and it includes only selected metadata items (gmd:fileIdentifier, gmd:language, gmd:title, gmd:abstract, gmd:spatialResolution, gmd:distributionInfo and gmd:EX_GeographicBoundingBox).

Response on brief

Response on resultType=results and ElementSetName=brief is the same as response on ElementSetName=full described above, but the record element is named BriefRecord and it includes only selected metadata items (gmd:fileIdentifier, gmd:title and gmd:EX_GeographicBoundingBox).

GETRECORDBYID OPERATION

The GetRecordById operation [OGC 2009a, page 160] is KVP encoded and requested via HTTP GET method. It returns a record (records) according to specified id (ids). The same possibility exists to request a full, summary, or brief record from the catalogue.

Request
```
http://gis.vsb.cz/catalogue/cswl?
REQUEST=GetRecordById&SERVICE=CSW&VERSION=2.0.2&ElementSetName=full&id=
http://gis.vsb.cz/datasets/0015600,http://gis.vsb.cz/datasets/0115600
```

Parameters REQUEST, SERVICE, and VERSION are mandatory and fixed. The parameter ElementSetName is optional (default is summary). Parameter Id can contain a comma to separate ids and is mandatory.

Response

The response is the same as in the GetRecords operation.

DESCRIPTION OF SERVICES

The CSW-L specification comes from the GeoWeb services orchestration project (GA 205/07/0797) [Růžička 2009]. To make CSW-L specification more clear in a context of GeoWeb services orchestration, it should include information on how to describe specific information about services using ISO 19119/19139.

Service type

One of the most important issues that is connected with service description is an identification of the service type. When searching in the catalogue, a user (client) must have the possibility to find similar services based not only on service interface (WMS), but on data used as well.

For purposes of an orchestration, it is necessary to be able find similar services. This can be managed through several ways. Some of them are described in [Růžička 2008]. The most suitable seems to be the concept based on ontology, where similar services are described as related terms in one branch of the ontology. The service type consists of interface definition, used data, and used algorithms.

The service type is defined by the identifier in the thesaurus. The identifier is constructed from parts of the service description to make the id more human readable. Example of identifier follows.
```
ogc:WMS:1.1.1/datasets:ZABAGED,DMÚ200/algorithms:simplify
gis.vsb.cz:erosion:0.0.1/datasets:FreeGEOCZ/algorithms:RUSLE
```

This type identifier is encoded in the serviceTypeVersion element (see 7.3.), because serviceType according to INSPIRE element can contain only selected values: discovery, view, download, transformation, invoke or other.

```
<srv:serviceType>
  <gco:LocalName>view</gco:LocalName>
</srv:serviceType>
<srv:serviceTypeVersion>
  <gco:CharacterString>Service/gis.vsb.cz:erosion:0.0.1/datasets:FreeGEOCZ/algor
ithms:RUSLE</gco:CharacterString>
</srv:serviceTypeVersion>
```

What is still not defined is the dependency of the similarity on a context of a provided process (running orchestra). This should be defined later and included in the CSW-L specification and CSW-L INSPIRE Profile.

Monitoring

For purposes of an orchestration, there are very important measures in service monitoring [Kaszper 2008]. This information can be directly written into the *gmd:MD_Metadata*, but it could make *gmd:MD_Metadata* too complex and it will not be compatible with CSW-L. That is why this kind of information is only referenced via the gmd:onLine element. One of the *gmd:onLine* elements in *gmd:MD_Metadata/gmd:distributionInfo* should (necessary for dynamic orchestration) include a reference to an XML document (static or dynamic). The reference must be done as at the following example where the mandatory and fixed is the name of the onLine resource.

```
<gmd:onLine>
  <gmd:CI_OnlineResource>
    <gmd:linkage>
      <gmd:URL>http://gis.vsb.cz/catalogue/monitor?id=0015647</gmd:URL>
    </gmd:linkage>
    <gmd:protocol>
      <gco:CharacterString>HTTP</gco:CharacterString>
    </gmd:protocol>
    <gmd:name>
      <gco:CharacterString>monitoring</gco:CharacterString>
    </gmd:name>
  </gmd:CI_OnlineResource>
</gmd:onLine>
```

The monitoring document should contain results from monitoring. We have specified three monitoring parameters that should be included, but others can be added to the results. An example of the XML file that describes results from monitoring follows.

```
<?xml version="1.0" encoding="UTF-8"?>
<Service xmlns="http://gis.vsb.cz/monitoring/0.0.1"
id="http://gis.vsb.cz/services/0015647">
  <stability>0.95</stability>
  <lastState>ssOK</lastState>
  <speed>1000</speed>
</Service>
```

The stability element contains a float value that describes the long time stability of the service measured in percentage of valid responses according to all valid requests.

The lastState element describes the state of the service from the last monitoring. The possible values are: ssOK, ssWarning, ssCritical, and ssUnknown.

The speed element contains the average time spent on one request-response operation. It is measured in milliseconds.

Full example

The following example describes how a service description should look using ISO 19119/19139.

```
<gmd:MD_Metadata ... >
  Because of space limitations for a paper. The normative example is available
at http://www.ogclite.org/cswlite/inspire/service.xml
</gmd:MD_Metadata>
```

CONCLUSION

The prepared CSW-L INSPIRE Profile specification should be simple enough to be easily implemented by any developer. There are still some open problems with the specification, but not significant. If the specification is not used, it, at least, can help with understanding how to use CSW for INSPIRE.

REFERENCES

GeoNode (2010). CSW_GetRecordById. http://projects.opengeo.org/CAPRA/wiki/CSW_GetRecordById

ISO (2005). ISO 19139. http://www.isotc211.org/2005/

Drafting Team Metadata and European Commission Joint Research Centre (2010) *INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119*. European Commission Joint Research Centre. Available at:
http://inspire.jrc.ec.europa.eu/documents/Metadata/INSPIRE_MD_IR_and_ISO_v1_2_20100616.pdf

Kaszper, R. (2008) *Sledování stavu služeb pro GeoWeb*. Diploma thesis. VŠB-TUO, Ostrava. 2008.

OGC (2009a). http://portal.opengeospatial.org/files/?artifact_id=20555, CSW 2.0.2.

OGC (2009b). http://schemas.opengis.net/csw/2.0.2/.

OGC (2007). http://schemas.opengis.net/iso/19139/20060504/srv/.

OGCLite (2010). http://www.ogclite.org/.

Růžička, J. (2008) ISO 19115 for GeoWeb services orchestration. *P*roceedings from Geoinformatics on CTU 2008. CTU, Prague. Available at:
http://geoinformatics.fsv.cvut.cz/wiki/index.php/ISO_19115_for_GeoWeb_services_orchestration

Růžička, J., Šeliga, M. (2010) Catalogue Services Lite. Proceedings from GIS Ostrava 2010. VSB-TUO. Ostrava. Available at: http://gis.vsb.cz/GIS_Ostrava/GIS_Ova_2010/sbornik/Lists/Papers/EN_1_11.pdf

Růžička et al. (2009) Geoweb services orchestration. http://kokos.vsb.cz/wiki/.