CSW 3.0

Part 4: OpenSearch Query Interface

74th OGC Technical Committee
Toulouse, France
Pedro Gonçalves, Terradue
September 20, 2010
Agenda

• OpenSearch Revisited
  – Geo and Time extension
  – Usage
  – Adoption

• CSW 3.0 Part 4
  – Status
  – Core elements mappings
  – Planning for document release
OpenSearch

- OpenSearch started in an effort built around Amazon's A9.com and now maintained in a community process at opensearch.org
- Allows syndication of search results
- Search engines have a description document used by client applications
- OASIS Search Web Services group is publishing searchRetrieve Operation with Bindings for SRU 1.2 and OpenSearch
  http://docs.oasis-open.org/search-ws/
OpenSearch

- OpenSearch is flexible, results can be returned as HTML interfaces or in Atom, XML/RDF, KML, WKT, JSON,...
- Directly integrated in Web Browsers
- Includes result’s paging
OpenSearch Description

- Provides metadata about the contents along with a set of URL Templates which illustrate the parameters accepted by the service and the variety of output formats in which results can be obtained.
- The OpenSearch request interface is simple, consisting of a description of a HTTP GET request with a series of optional key-value parameters that can be used to constrain the search.

```xml
<Url type="text/html" template="http://example.com/?q={searchTerms}+'&amp;pw={startPage ?}"/>
```

http://www.google.com/?q=question
http://www.google.com/?q={searchTerms}
<?xml version="1.0" encoding="UTF-8"?>
<OpenSearchDescription xmlns="http://a9.com/-/spec/OpenSearch/1.1/"
    <$ShortName>Web Search</$ShortName>

 <$Description>Use Acme.com to search the Web.</$Description>
 <$Tags>example web</$Tags>
 <$Contact>admin@acme.com</$Contact>

 <Url type="application/rss+xml"
     template="http://acme.com/?q={searchTerms}&amp;pw={startPage?}&amp;format=rss"/>
</OpenSearchDescription>
<?xml version="1.0" encoding="UTF-8"?>
<OpenSearchDescription xmlns="http://a9.com/-/spec/OpenSearch/1.1/">
  <ShortName>Web Search</ShortName>
  <Description>Use Acme.com to search the Web.</Description>
  <Tags>example web</Tags>
  <Contact>admin@acme.com</Contact>
  <Url type="application/atom+xml" template="http://acme.com/?q={searchTerms}&amp;pw={startPage?}&amp;format=atom"/>
  <Url type="application/rss+xml" template="http://acme.com/?q={searchTerms}&amp;pw={startPage?}&amp;format=rss"/>
  <Url type="text/html" template="http://acme.com/?q={searchTerms}&amp;pw={startPage?}"/>
  <LongName>Example.com Web Search</LongName>
  <Image height="64" width="64" type="image/png">http://acme.com/websearch.png</Image>
  <Image height="16" width="16" type="image/vnd.microsoft.icon">http://acme.com/websearch.ico</Image>
  <Query role="example" searchTerms="cat"/>
  <Developer>Acme.com Development Team</Developer>
  <Attribution>Search data Copyright 2005, Acme.com, Inc., All Rights Reserved</Attribution>
  <SyndicationRight>open</SyndicationRight>
  <AdultContent>false</AdultContent>
  <Language>en-us</Language>
  <OutputEncoding>UTF-8</OutputEncoding>
  <InputEncoding>UTF-8</InputEncoding>
</OpenSearchDescription>
GeoSpatial Extension

• Specify a series of parameters that can be used to spatially constrain search results.
• Provision is made to filter results by:
  – A bounding box
  – A geometry using Well Known Text
  – Within a certain radius from a given latitude-longitude point
  – Having a certain containment relation (within, overlaps, disjoint) with a geographic constraint
  – Matching a geographic name
• All geographic information is to be expressed using the EPSG 4326 (WGS84)
  – Following the GeoRSS “precedent”
GeoSpatial Extension Elements

- **geo:box** -> Geographic bounding box The box is defined by "west, south, east, north" coordinates of longitude, latitude, in a EPSG:4326 decimal degrees.
- **geo:geometry** -> Geographic area (geometry): The polygon is defined using the Well Known Text standard for geographic shapes, using EPSG:4326 (POINT, LINESTRING, POLYGON, MULTIPOINT, MULTILINESTRING, MULTIPOLYGON)
- **geo:lat, geo:lon, geo:radius** -> The latitude, longitude of a given point with a given radius
- **geo:relation** -> Spatial relation to result set, Character String; One of "overlaps", "contains", "disjoint" (default is "overlaps")
- **geo:name** -> A string describing the location to perform the search

The namespace of the OpenSearch GeoSpatial Extension is: http://a9.com/-/opensearch/extensions/geo/1.0/
Time Extension Elements

- **time:start** - Start of the temporal interval to search (RFC-3339)
- **time:end** - End of the temporal interval to search (RFC-3339)

The namespace of the OpenSearch Time Extension is:
http://a9.com/-/opensearch/extensions/time/1.0/
Geo and Time Extension

- The geospatial extension allows to formulate geospatial requests e.g. point-plus-radius, a bounding box, or a polygon

  \[ \text{bbox=} \{\text{geo:box}?) \]

- Together with the Time extension, OpenSearch can specify time start, finish, and slices for searching data.

  \[ \text{start=} \{\text{time:start}?) \text{& stop=} \{\text{time:end}?) \]

<Url type="text/html"
    template="http://example.com/xml/?q={searchTerms?}&start_date={time:start?}&stop_date={time:end?}&bbox={geo:box}"/>
Integration

- REST approach allowing the query and registry of datasets
- Allows the easy definition of custom result types to ease the integration and creation of mashup applications
- Responses as HTML interfaces or in Atom, XML/RDF, KML, WKT, metalink, EOP/HMA
Integration using feeds
Experience from GENESI-DR

• OpenSearch with geo-temporal extensions provided the necessary flexibility to federate different Earth Science communities
• Search results can be in mass-market formats (e.g. KML, Atom) but also in community specific like the HMA/EOP
• Implementation of gateways demonstrated a future path to access community wide infrastructures (e.g. CNES’s SIPAD)
• Follow-up in new GENESI-DEC project focusing on Digital Earth Communities enlarging the GENESI-DR infrastructure in terms of both resources availability and geographical extent
Adoption

• GENESI-DR
  – ESA, NILU, JRC, ENEA, DLR, CNES, INFOTERRA-UK, K-SAT

• ESA Grid - G-POD Catalogue
  (ESA biggest online data repository)

• OGC Web Services Phase 7 (OWS-7)

• Software
  – Geocommons
  – Gi-CAT (ESSI Lab)
  – GeoNetwork (development trunk)
  – OpenLayers (development trunk)
Adoption

- ngEO - ESA next generation of user services for Sentinel ground segment (GMES) seems to adopt it ...
CSW 3.0 OpenSearch Vote Nov 2009

• OpenSearch with Geo and Time extensions as the "baseline" query operation for all profiles:
  – the complete CSW 2.0.2 implementation without any additional profiles or extension packages.
  – Supported regardless of profile (or even without a profile).
  – It defines core queryables, query and response message syntax,
  – Allows a client to query any catalog by any profile without having to know the extended profile-specific details

• Requests using HTTP GET and responding with Atom (w/geoRSS) results at a minimum,

• Based on 09-084 “OpenSearch Geospatial Extensions” discussion paper and 08-169 Change Request
CSW 3.0 Part 4 (OGC 10-032) Status

• Final version submitted in pending documents (aligned with the OpenSearch.org)

• Codifies geospatial and temporal extensions in a form compatible with OGC

• Includes:
  – Conformance Classes
  – Mappings for queirables
  – Mappings for response

• OpenSearch is included in annex for reference
# Response Mappings

<table>
<thead>
<tr>
<th>Name</th>
<th>OGC core returnable</th>
<th>Atom Element</th>
<th>Data type and values</th>
<th>Multiplicity and use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Title</td>
<td>atom:title</td>
<td>Character String</td>
<td>1 (mandatory)</td>
</tr>
<tr>
<td>Creator</td>
<td></td>
<td>atom:author</td>
<td>Character String</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Subject</td>
<td>Subject</td>
<td>atom:category</td>
<td>Character String</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Description</td>
<td>Abstract</td>
<td>atom:summary</td>
<td>Character String</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Publisher</td>
<td></td>
<td>atom:feed/atom:author</td>
<td>Character String</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Contributor</td>
<td></td>
<td>atom:contributor</td>
<td>Character String</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Date</td>
<td></td>
<td>dc:date</td>
<td>ISO-8601 date</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Modified</td>
<td>Modified</td>
<td>atom:updated</td>
<td>ISO-8601 date</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Identifier</td>
<td>Identifier</td>
<td>atom:id</td>
<td>Character String</td>
<td>1 (mandatory)</td>
</tr>
<tr>
<td>Source</td>
<td>Source</td>
<td>atom:source</td>
<td>Character String</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Language</td>
<td></td>
<td>@xml:lang</td>
<td>Character String</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Type</td>
<td></td>
<td>atom:content/@type,</td>
<td>MIME-type</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Format</td>
<td>Format</td>
<td>atom:link/@type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relation</td>
<td>Association</td>
<td>atom:link [@rel=&quot;alternate&quot;]</td>
<td>URL (to the full metadata for the resource)</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Data</td>
<td></td>
<td>atom:link [@rel=&quot;enclosure&quot;]</td>
<td>URL (to the location of the data resource)</td>
<td>1 (optional)</td>
</tr>
<tr>
<td>Coverage</td>
<td>Bounding Box</td>
<td>georss:where</td>
<td>GeoRSS (geometry of the resource)</td>
<td>1 (mandatory)</td>
</tr>
<tr>
<td>Rights</td>
<td>Rights</td>
<td>atom:rights</td>
<td>Character String</td>
<td>1 (optional)</td>
</tr>
</tbody>
</table>
Motion

The Catalog Service 3.0 SWG recommends that the OGC Technical Committee approve using the Fast Track process for OGC Document Number 10-032 “OpenSearch Query Interface” as an OGC Candidate Standard.

– Motion: Pedro Goncalves
– Second: Arnulf Christl
– There was no objection to unanimous consent
Document Motion template

• Sample Document Motion:

• The <DWG or SWGname> recommends that the OGC Technical Committee approve release of <OGC Document Number> “<OGC Document Name>” as an OGC <White Paper/Discussion Paper/ Engineering Report>
  – Pending and final edits and review by OGC staff
  – <Any Discussion Points that the TC needs to be aware of>
  – Motion: <Name of person making the motion>
  – Second: <name of the person seconding the motion>
  – If there is a hand vote, the results of the vote. Otherwise, the phrase <There was no objection to unanimous consent> should be used
Motion template to Request an electronic vote

- The <DWG or SWGname> recommends that the OGC Technical Committee approve an electronic vote to approve release of <OGC Document Number> “<OGC Document Name>” as an OGC <Best Practice/Adopted Standard>
  - Pending and final edits and review by OGC staff
  - <Any Discussion Points that the TC needs to be aware of>
  - Motion: <Name of person making the motion>
  - Second: <name of the person seconding the motion>
  - If there is a hand vote, the results of the vote. Otherwise, the phrase <There was no objection to unanimous consent> should be used