Bolegweb Platform – Contribution to the Web Communities

T. Kliment[1], V. Cetl[1], M. Tuchyňa[2], M. Kliment[3], G. Bordogna[4]

[1] Faculty of Geodesy, University of Zagreb, Zagreb, Croatia

[2] Slovak Environment Agency, Data Center, Banská Bystrica, Slovakia

[3] Horticulture and Landscape Engineering Faculty, Slovak University of Agriculture in Nitra, Nitra, Slovakia

[4] Institute for the Dynamics of Environmental Processes, National Research Council of Italy, Milan, Italy

Abstract

The effective access to and reuse of geospatial information (GI) has come to be of critical value in modern knowledge based society. The standardized web services defined by Open Geospatial Consortium (OGC) are frequently used for the implementation of a spatial data infrastructure (SDI), to expose geospatial data, metadata and models on the Web. These GI are normally stored in an encoded geospatial layer, which is hidden from search engines. SDI uses a catalogue service for the web as a gateway to GI through the metadata defined by ISO standards, which are structurally diverse to OGC metadata. Therefore, a crosswalk needs to be implemented to bridge the OGC resources discovered on mainstream web with those documented by metadata in an SDI to enrich its information extent. We have to build mechanisms allowing entrepreneurs and developers access the information SDI is providing to build their apps. The paper reports a global wide and user friendly platform of OGC resources available on the web with the main goal to ensure and enhance the use of GI within a multidisciplinary context and to bridge the geospatial web from the end-user perspective, thus to open its borders to more web communities. The platform has been developed in the research project Borderless Geospatial Web (Bolegweb).

Keywords

Geospatial Web, Mainstream Web, Geospatial Information, SDI, OGC services, Discovery, Bolegweb.

Introduction

Rapid development of Spatial Data Infrastructure (SDI) in Europe triggered by INSPIRE (European Commission, 2007) and other similar initiatives make more and more geospatial information (GI) resources (data, metadata and models) available on the web. The main objective of an SDI is to enable data harmonization, sharing and reuse.. The fundamental component of an SDI which enables users to search and discover the GI resources is metadata, or data about data. Metadata are the first visible component of each SDI for users. In the frame of an SDI, the metadata are usually divided into metadata for geospatial data and metadata for geospatial services. Both are provided in a standardized way by discovery services. A very popular one is the Open Geospatial Consortium (OGC) catalogue service standard, which is implemented by many existing either commercial or open source software

solutions. Geospatial data users search for GI resources within an SDI using discovery clients of a Geoportal application (i.e. INSPIRE Geoportal) (Kliment et al., 2013). All aforementioned works effectively when both producers and users are aware of an SDI, corresponding services and how to use them. On the other hand, there are many potential users who are not aware of an SDI. They usually search for GI resources through web search engines such as Google, Yahoo, Bing etc. One such community is represented by the mainstream developers, who might already be aware of data coming from an SDI world; however they need an easy recipe how to integrate the geospatial data into their applications. In addition, there are still many GI data producers making their resources available on the Web without any or incomplete and unstandardized documentation. They would need to create and publish metadata describing their GI resources in a predefined structure in order

to make them discoverable and thus shareable in an SDI. This approach would allow for distributed searches or harvesting of the metadata from different SDI nodes.

State of the art

Discovery of Geospatial Resources: Methodologies, Technologies, and Emergent Applications book (Díaz et al., 2012) presents a collection of attempts to push forward the automated discovery of GI resources. Contributions in the book from different authors provide a wide spectrum of perspectives and possible methods. For the data producers an added value would be to facilitate the production of standardized metadata embedded into data production workflows, link data with metadata, ensure that all changes in data are automatically reflected in metadata. An ability to automatically generate standardized metadata from the content of a harvested data-publishing server would significantly facilitate maintenance and management of the description of large volumes of data as reported in previous research works (Florczyk et al., 2012; Kliment et al., 2015; Kalantari et al., 2016; Giuliani et al., 2016). In addition, many research activities (Abargues et al., 2009; López-Pellicer et al., 2011; Kliment et al., 2013; Hou et al., 2016) have reported that the mainstream web provides significant number of valuable GI resources of several types (e.g. OGC services, KML data, etc.). They can be discovered using the mainstream web search engines. The benefit in comparison with SDI engines (geospatial catalogues) is that the search engine robot automatically crawls the web in order to discover available information resources. GI resources discovered with web search engines may significantly extend the information richness of an SDI, and may be used to extend a specific domain oriented SDI portal with a combination of social media, voluntary geographic information (VGI) (Poorazizi et al., 2015) etc. The importance and relevance of GI discovery can also be seen in industry through the establishment of the Spatial Data on the Web working group by the World Wide Consortium (W3C) in close cooperation with Open Geospatial Consortium (OGC). W3C and OGC are industrial standardization bodies developing open standards for the mainstream and geospatial web. The main objective is to make it easier to publish and use geospatial data on the web. Improving discovery of geospatial data on the Web is recognized as a one of the use cases that demands a combination of geospatial and non-geospatial data sources and techniques. There are two important requirements identified by the W3C and OGC

cooperation: i) Crawlability - geospatial data need to be crawlable in the web, to be found and indexed by external agents; ii) Discoverability and accessibility- once geospatial data are published on the web, both humans and machines should be able to discover and access them.

This paper provides the results achieved in the research project Bolegweb, which aimed at the development of a geospatial meta-search crawler platform to collect GI resources accessible online and published on the Web using OGC services, harvest the geospatial metadata and deploy Graphic User and Application Programming Interfaces (GUI and API) facilitating access for wider web communities.

Research goal

The main objective of the project was to design, develop and implement a complex solution for the discovery of GI from OGC web services available on the Internet; generate and share geospatial metadata according to standards that might be facilitated by mainstream, SDI and Semantic web communities. The goal was composed of the following sub-objectives:

- Design and development of an OGC metasearch enhanced crawler for semi-automatic collection of OGC services end points (URLs).

- Publish the metadata describing both OGC services and the content published as geographic layers, tiles, features, coverages, observations through implementations of an SDI catalogue, and thus contribute to the information coverage provided by the current SDI's implementations at any level.

- Design and develop web graphic user and application programming interfaces providing search and access facilities.

Materials and methods

OGC Geospatial Services

The following seven OGC services and related GI resources have been collected within the Bolegweb project:

1. Web Map Service (WMS) – is the mostly used web service operating on both raster and vector geospatial data and generates map preview with a dynamic way of layer style symbolism each associated with a legend (Blower et al., 2013).

2. Web Feature Service (WFS) – is a web service operating on vector geospatial data and provides query and transactions mechanism to access and retrieve geometric and attribute data at the feature level in real time on the Web (Peng and Zhang, 2004).

3. Web Coverage Service (WCS) – is a web service for accessing and processing of raster data, more generally: coverages (Baumann, 2010) of multidimensional gridded data, extending WMS by formats used for complex modelling and analysis, usually used to encode large datasets for meteorology, oceanography and climatology.

4. Web Processing Service (WPS) – is a web service for creating and distributing web-based functions (Michaelis and Ames, 2009) operating on raster/vector data using as inputs and producing as outputs resulting from analysis executed by the service (e.g. buffering, zonal statistics, nearest feature, shortest path, clipping, etc.).

5. Sensor Observation Service (SOS) – is a web service sharing observational and measurement data collected on a spatial feature by providing the information about observed property (what

was measured), feature of interest (where was it measured), procedure (how was it measured), phenomenom time (when was it measured), the data quality and most important the result of the observation (Schleidt, 2013).

6. Catalogue Service for Web (CSW) – is a web service operating on collections of descriptive information (metadata) being created and maintained for geospatial data and services and provides access, filtering and transactions on metadata records. The service has become a major mechanism to catalogue resources and shares links to relevant GI encoded in the metadata (Li et al., 2011).

7. Web Map Tile Service (WMTS) – is a web service operating on map tiles caches of spatially referenced data using pregenerated tile images with predefined content, extent, resolution within a coordinate reference system.

System architecture

The system concept definition of the Bolegweb platform high-level architecture is represented in the architecture schema depicted in Figure 1.
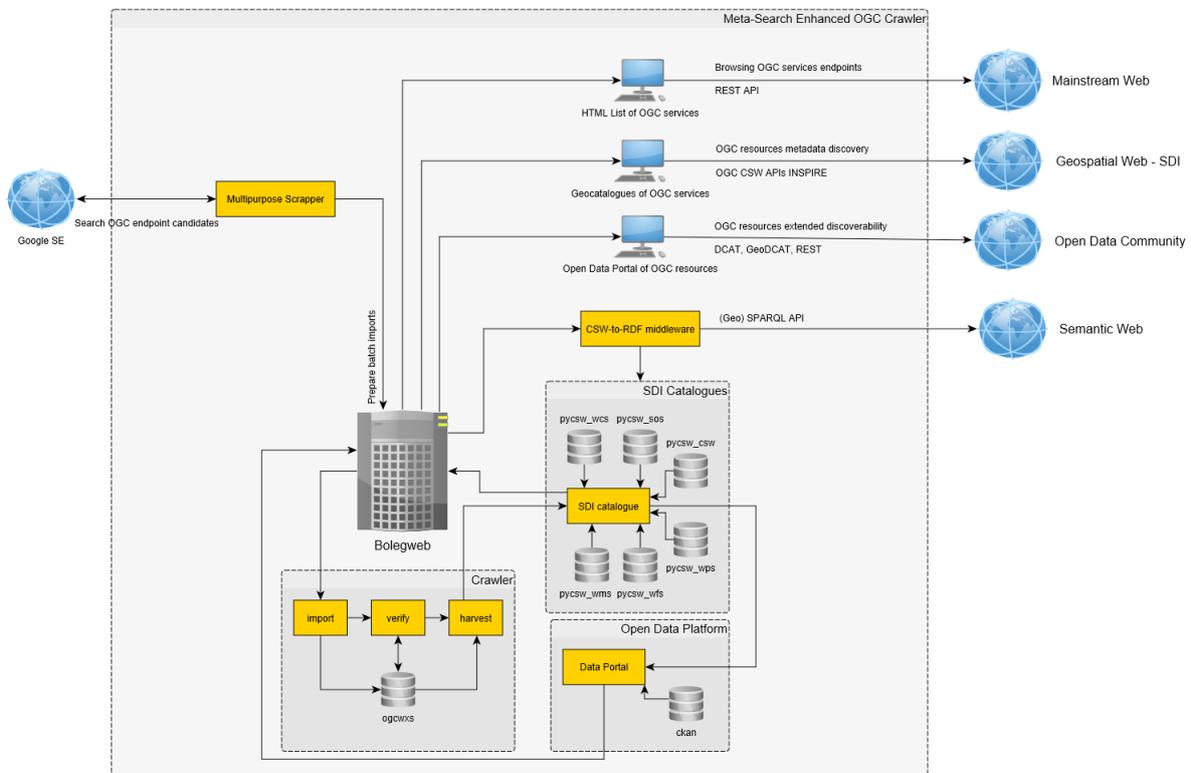
Source: Own processing

Figure 1: Bolegweb system architecture overview.

The overall infrastructure of the Bolegweb system is quite complex and consists of several components. The services collection flow starts with the Google search engine where potential candidates of OGC service endpoints are collected by the Multipurpose Scrapper component using advanced search parameters (e.g. parameter inurl contains strings as "service", "request", "WMS": "inurl:service inurl:WMS inurl:request"). In the next step the results are passed to the Crawler component to process these inputs as follows: (i) Gathered OGC services' endpoints candidates and related information are imported into the crawler database (import); (ii) the Crawler's verification script checks the availability of the collected services and extracts the service type, version, basic quality parameters, server location and other information (verify); (iii) the Harvesting script queries the identified services and related resources for metadata and stores this in an SDI catalogue component (harvest).

The SDI catalogue component is represented by set of Catalogue Service for Web (CSW) interfaces implemented using pycsw, a python based implementation of the OGC's CSW server. pycsw allows for the discovery and publishing of metadata for geospatial resources, and can be deployed as a standalone server or embedded in other applications (Sibolla et al., 2014). We use pycsw as a standalone metadata management system. Individual service types have their own virtual CSW endpoint (e.g. pycsw_wms), where the metadata are harvested for both services and related content.

The system provides an extension over the SDI catalogues, which serves as a gateway between SDI and Semantic web communities. CSW-to-RDF middleware is implemented using the TripleGeo-CSW software developed by the GeoKnow project team; this was developed as a proof of concept working with CSWs from public authorities across Europe, which involved datasets complying with the EU INSPIRE Directive. Experience gained testifies that TripleGeo-CSW can assist stakeholders to repurpose existing CSWs with minimal overhead and readily expose spatial metadata on the Semantic Web (Athanasiou et al., 2015). The metadata collected into SDI catalogues from available OGC services are exposed via a GeoSPARQL endpoint; SPARQL queries are parsed to identify filter criteria and generate a corresponding CSW GetRecords request, which is then send to the remote CSW service via HTTP/POST. The CSW service responds with ISO/XML metadata files that are transformed to the RDF/XML form and provide RDF triples in the result set; this RDF response is then returned to the original requester.

For the development of clients' platforms various GUIs and APIs were designed and are being developed. A simple tabular representation of the data about OGC services is provided as a HTML/JavaScript client, with filtering options as well as a simple REST service interface with the same capabilities. A light JavaScript map application allows users to search for the metadata stored in the CSW and display the results. An Open Data Platform based on the data portal implementation CKAN, the world's leading out-of-the-box open-source data portal platform, was deployed. It provides tools to streamline the process of publishing, sharing, finding and using data. It has native capability to harvest several types of open resources, including SDI catalogue CSWs. We use this functionality to collect the metadata about discovered OGC resources managed by SDI catalogues communicating to CSWs and storing this into the CKAN internal data model. Metadata can be exposed to communities using mainstream REST or DCAT as well as geospatial community specific GeoDCAT standards.

Results and discussion

The collection process of OGC services endpoints has been launched in October 2013. This process has been repeated on a monthly basis. The last metasearch process was performed on August 31st and the Crawler database contains a total of 22875 records. It is important to stress that these data also characterize those URLs that do not provide direct links to OGC services, linking instead to a web page, which may after further crawling of URLs provide a link to an access point. However, the current verification step in the crawler workflow does not further investigate those URLs which do not immediately provide an XML type response with a root attribute identifying the OGC Service version. (e.g. <sos:Capabilities version="2.0.0" ….. >). In addition, some of the discovered URLs might no longer be available, or even invalid. After filtering all the potential OGC service candidates, the current database provides interesting numbers of records representing functioning services (Figure 2).

The actual figures change frequently due to the dynamic nature of the Web environment and specifically the geospatial web. Some services available today might be offline or removed tomorrow while new ones may become publicly
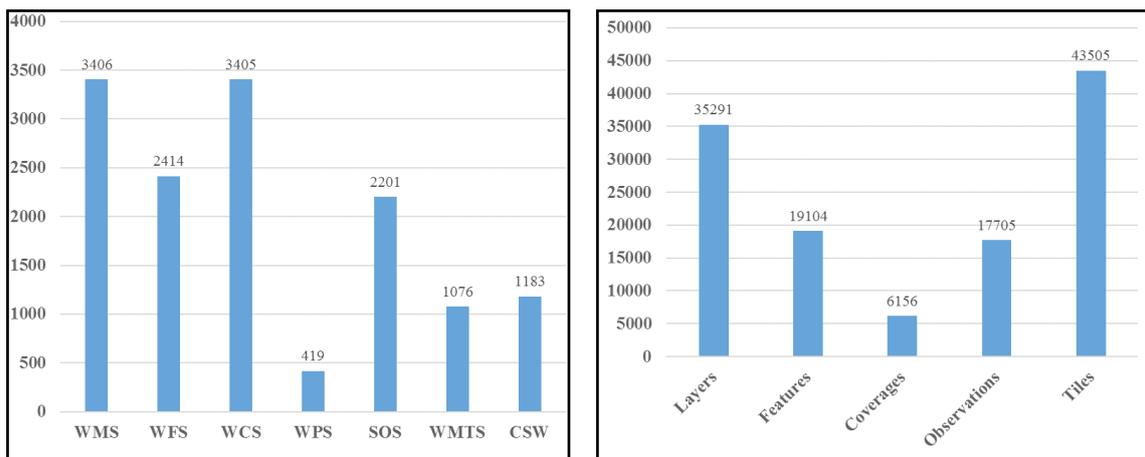
Source: Own processing

Figure 2: Number of functioning OGC service endpoint URLs available from the Bolegweb platform and harvested metadata for related GI resources (17 September 2016).

available. The number of available services does not necessarily provide a clear picture of GI availability on the web. Each service serves some type of GI resources as layers, features, etc. In order to understand the extent with this respect, the metadata from individual service endpoints were harvested for both service and related content (WMS Layers, WFS Features, WCS Coverages, etc.) and the results are represented by the chart in Figure 2, right part. However, the number of metadata records does not fully correspond to the number the OGC services due to various reasons as follows: i) not all versions of OGC services were harvested; ii) some of the supported services were not harvested due to exceptions raised during the harvesting requests. These facts decreased the final number of metadata records and will have to be resolved in future work by extending the functionality of pycsw and resolving the exception errors. Nevertheless, with regard to the main objective of the Bolegweb project, we were able to reach the goal of providing interfaces to the collected GI resources for wider web communities by deploying existing solutions and developing new ones where required, as described in the following section.

Mainstream web community

For the mainstream web community, represented by the entrepreneurs and developers having interest in discovering GI served by OGC services in an easy and straightforward manner, both GUI and API interfaces were developed. The GUI developed for the mainstream Web community and easy access to OGC services is entitled HTML List of OGC Services (Figure 3).

This simple HTML page provides a list of OGC Web Services represented in a simple tabular view based on the database implemented underneath the Bolegweb crawler component. Information such as the OGC Service type, version, location of the server, date when the service was discovered on Google SE, when its availability has been last checked and finally its status after the last check are made available. In addition, a link to metadata available as a query to the corresponding CSW is provided. A user can filter by applying free text criteria on each column. In the example represented by the Figure 3, the user filters the OGC service type SOS, restricting the version to 2.0, the country to be Italy and value 1 under status which stands for functioning service. To access the same information from the remote application code, a REST service is made available and provides the same filtering capabilities as the described GUI using the HTTP GET protocol. The same query would be represented by the following URL:

https://bolegweb.geof.unizg.hr/ogcwxs/rest/json.php?
type=SOS&
version=2.0.0&
location=Italy&
status=1

This would result in the JSON representation of the data from the database being returned, as illustrated in the following object fragment of the first result:

Currently displayed OGC services: 14

ID	Google Title	Type	Version	Metaseach URL discovered	Server location	Import Date	Status	Crawling Date	Metadata
		sos	2.0.0		italy		1		
3790	"StarterKit SOS StarterKit SOS OGC:SOS 1.0.0 2.0.0 http..."	sos	2.0.0	http://geosk.ve.ismar.cnr.it/observations/sos/kvpdemo?service...	Italy	2014-08-0...	●	2016-09-07T13:57...	14
3799	"OGS/NODC SOS OGS/National Oceanographic Data Center..."	sos	2.0.0	http://nodc.ogs.trieste.it/SOS/sos?service=SOS&REQUEST=Ge...	Italy	2014-08-0...	●	2016-09-07T13:57...	-1
7369	"Capabilities Link - Starter Kit - IREA CNR UOS Milano"	sos	2.0.0	http://skmi.irea.cnr.it/observations/sos/kvp?service=SOS&requ...	Italy	2014-09-07...	●	2016-09-14T10:00...	2
9932	"Capabilities Link - Starter Kit - Demo - Cnr"	sos	2.0.0	http://geosk.ve.ismar.cnr.it/observations/sos/kvp?service=SOS...	Italy	2014-10-16...	●	2016-09-06T21:5...	14
122...	"Capabilities Link - Cnr"	sos	2.0.0	http://sk.oristano.iamc.cnr.it/observations/sos/kvp?service=SO...	Italy	2015-01-07...	●	2016-09-06T21:3...	-1
12821	"Capabilities Link - Starter Kit - Demo - Cnr"	sos	2.0.0	http://vesk.ve.ismar.cnr.it/observations/sos/kvp?service=SOS&...	Italy	2015-02-0...	●	2016-09-07T04:0...	29
13539	"Capabilities Link - Cnr"	sos	2.0.0	http://mareesk.irea.cnr.it/observations/sos/kvp?service=SOS&r...	Italy	2015-03-26...	●	2016-09-06T22:0...	16
13664	"Capabilities - Starter Kit - CNR ISMAR Lesina"	sos	2.0.0	http://sk.fg.ismar.cnr.it/observations/sos/kvp?service=SOS&re...	Italy	2015-03-26...	●	2016-09-06T22:0...	-1
15147	"52N SOS 52North Sensor Observation Service - Data..."	sos	2.0.0	http://sk.ise.cnr.it/observations/sos/kvp?service=SOS&request...	Italy	2015-06-0...	●	2016-09-07T10:3...	23
15257	"Capabilities - Cnr"	sos	2.0.0	http://sk.ise.cnr.it/observations/sos/kvp?service=SOS&request...	Italy	2015-06-0...	●	2016-09-07T10:3...	23
15284	"Capabilities Link - ermes - get-it"	sos	2.0.0	http://get-it.ermes-fp7space.eu/observations/sos/kvp?service=S...	Italy	2015-06-0...	●	2016-09-07T10:3...	-1
15287	"Capabilities Link - Starter - Demo - Ogs"	sos	2.0.0	http://geonodenodc.ogs.trieste.it/observations/sos/kvp?service...	Italy	2015-06-0...	●	2016-09-07T10:3...	-1
16696	"Starter Kit ISMAR Venezia None ASD OGC:SOS 1.0.0 2.0.0..."	sos	2.0.0	http://vesk.ve.ismar.cnr.it/observations/sos/kvp?service=SOS&...	Italy	2015-08-17...	●	2016-09-07T05:2...	29
17885	"OGS - Starter Kit Node NODC-OGS (Trieste) RITMARE..."	sos	2.0.0	http://geonodenodc.ogs.trieste.it/observations/sos/kvp?service...	Italy	2015-10-09...	●	2016-09-07T11:19...	-1

Source: https://bolegweb.geof.unizg.hr/site/products#tabs-1

Figure 3: GUI of HTML List of OGC Services with a query applied searching for online SOS services, version 2.0 running on machines in Italy.

```json
{
  "id" : "3790",
  "title" : "\"StarterKit SOS StarterKit
    SOS OGC:SOS 1.0.0 2.0.0 http...\"",
  "url" :
"http:\/\/geosk.ve.ismar.cnr.it\/
observations\/sos\/kypdemo?service
=SOS&request=GetCapabilities",
    "importDate" : "2014-08-
    05T02:15:07+02:00",
  "location" : "Italy",
  "endpoint" :
"http:\/\/geosk.ve.ismar.cnr.it\/
observations\/sos\/kvpdemo",
  "type" : "sos",
  "version" : "2.0.0",
  "status" : "1",
  "statusDate" : "2016-09-
07T13:57:44+02:00",
  "harvested" : "1",
  "harvestingDate" : "2016-09-
  10T12:13:35+02:00",
  "metadata" : "14",
  "metadataDate" : "1970-
  0101T01:00:00+01:00"
}
```

Geospatial web community - SDI

For the geospatial community, while they may also use the interfaces described in the previous section, standard SDI techniques for search of GI resources might be of preference; for these we developed a map based discovery client and deployed CSW endpoints for individual OGC Services endpoints. Ideally the user searches, displays and uses the data from within the same web application, or within a desktop GIS e.g. using the QGIS plugin MetaSearch Catalogue Client. The same query

as above would be set up in the discovery map viewer as follows: i) selecting the Sensor Observation Service from the drop list; ii) selecting Italy from the place search box iii) activating the current spatial extent function and iv) typing the string 2.0.0 in a full text query box; v) triggering the search function (Figure 4).

The result list provides metadata records for SOS and related observations matching the query with links to metadata records and the service endpoint URL. The same scenario can be imitated from a third party discovery client, e.g. using the Metasearch Catalogue Client plugin in QGIS desktop and a CSW endpoint for OGC service type of interest. In our example we use SOS services CSW endpoint (Figure 5).

Semantic web community

The semantic web community can reuse the GI resources collected by the Bolegweb platform using the GeoSparql endpoint deployed using the middleware software TripleGeoCSW connected to the platform's CSWs. The sample query tested in the previous section would be represented by the following SPARQL encoding:

```
PREFIX dc: <http://purl.org/dc/
elements/1.1/>
PREFIX csw: <http://www.opengis.net/cat/
csw/2.0.2>
PREFIX geo: <http://www.opengis.net/ont/
geosparql#>
PREFIX geof: <http://www.opengis.net/def/
geosparql/function/>
SELECT *
WHERE {
?s dc:format ?type .
?s csw:AnyText ?anytext .
```

Source: https://bolegweb.geof.unizg.hr/site/products#tabs-3
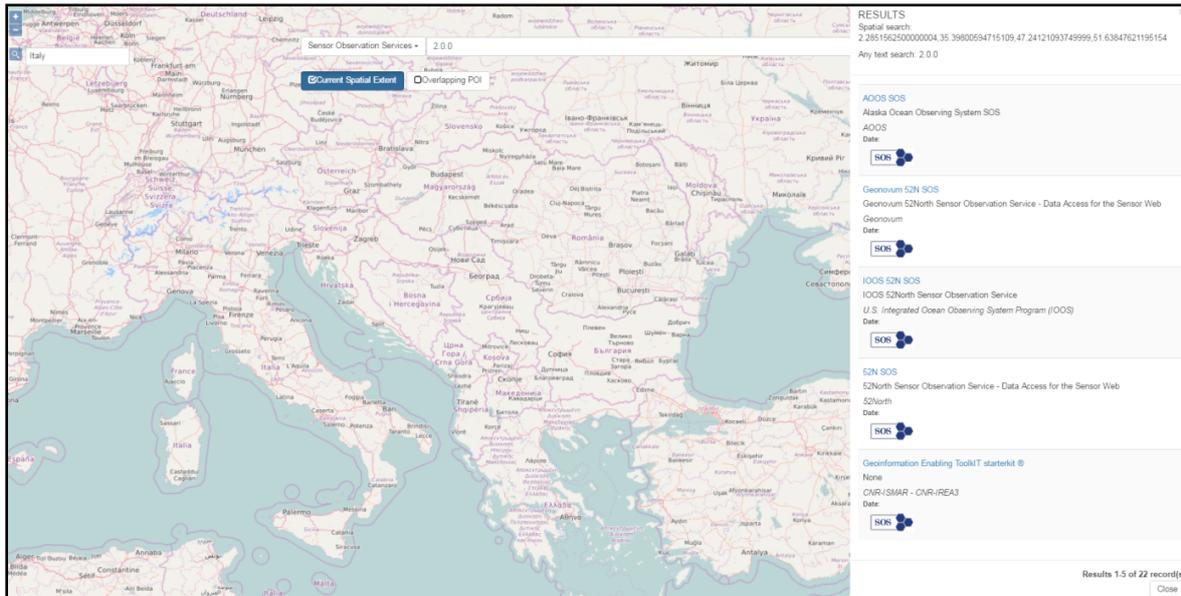
Figure 4: GUI of HTML List of OGC Services with applied query searching for online SOS services, version 2.0 running on machines in Italy.
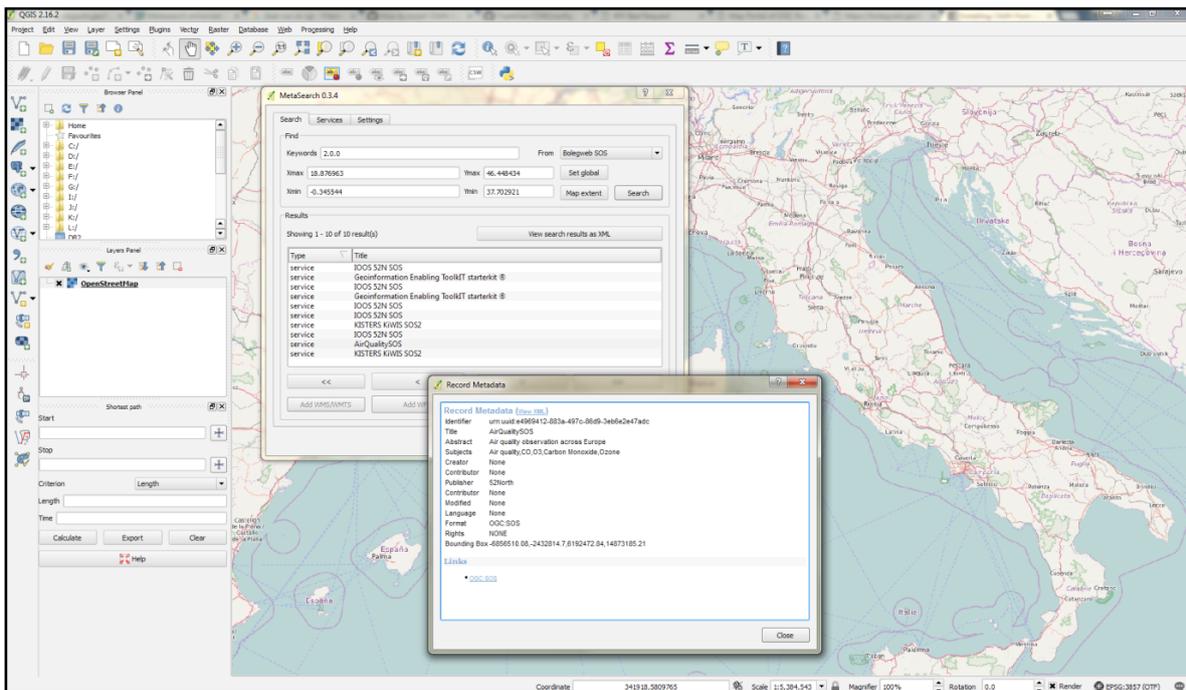
Source: Own processing with a CSW endpoint used from: https://bolegweb.geof.unizg.hr/site/products#tabs-2

Figure 5: GUI of Metasearch Catalogue Client plugin of QGIS desktop GIS software connected to the CSW endpoint for SOS services with applied a free text "2.0.0" and spatial query extracted from the map extent.

```
?s geo:hasGeometry ?fWKT .

FILTER ( REGEX(?anytext ,
"2.0.0") && (?type , "OGC:SOS") &&
geof:sfWithin(?fWKT, "BOX2D(-12.65625
33.23868752757414,32.2998046875
49.98655213050617)"^^geo:wktLiteral))

}
```

This query results in an RDF representation of metadata records matching the filter parameters as shown in the following fragment for a single record:

```
<rdf:Description
rdf:about="http://geosk.ve.ismar.cnr.it/
observations/sos/kvp">
```

```
    <foaf:primaryTopicOf rdf:resource="ur
n:uuid:urn:uuid:9a8744ff-74bc-457a-862c-
aef6cfbd1d63"/>
    <dct:language rdf:datatype="http://
purl.org/dc/terms/ISO639-2"/>
    <dct:title xml:lang="">Geoinformation
Enabling ToolkIT starterkit &#174;</
dct:title>
    <dct:description xml:lang="">National
Data Buoy Center SOS</dct:description>
    <rdf:type rdf:resource="http://www.
w3.org/ns/dcat#Dataset"/>
    <dcat:landingPage rdf:resource="http://
geosk.ve.ismar.cnr.it/observations/sos/
kvp"/>
    <dct:identifier rdf:datatype="http://
www.w3.org/2001/XMLSchema#string"/>
    <dct:subject/>
    <dcat:keyword xml:lang="">ASD1</
dcat:keyword>
    <dcat:keyword xml:lang="">ASD1</
dcat:keyword>
    <dct:spatial>
      <dct:Location>
        <locn:geometry
rdf:datatype="http://www.openlinksw.
com/schemas/virtrdf#Geometry">BOX2D(12.0
12.0,45.0 45.44)</locn:geometry>
      </dct:Location>
    </dct:spatial>
    <dct:provenance>
      <dct:ProvenanceStatement>
        <rdfs:label xml:lang=""/>
      </dct:ProvenanceStatement>
    </dct:provenance>
    <dcat:distribution>
      <dcat:Distribution>
        <dct:rights>
          <dct:RightsStatements>
            <rdfs:label/>
          </dct:RightsStatements>
        </dct:rights>
        <dct:accessRights>
          <dct:RightsStatement>
            <rdfs:label/>
          </dct:RightsStatement>
        </dct:accessRights>
      </dcat:Distribution>
    </dcat:distribution>
    <dct:rightsHolder>
      <foaf:Organisation>
        <foaf:name xml:lang="">National
Data Buoy Center</foaf:name>
        <foaf:mbox rdf:resource="mailto:"/>
      </foaf:Organisation>
    </dct:rightsHolder>
    <prov:qualifiedAttribution>
      <prov:Attribution>
        <prov:agent>
          <vcard:Kind>
            <vcard:organization-name
xml:lang="">National Data Buoy Center</
vcard:organization-name>
            <vcard:hasEmail
rdf:resource="mailto:"/>
          </vcard:Kind>
        </prov:agent>
        <dct:type rdf:resource="http://
inspire.ec.europa.eu/codelist/
ResponsiblePartyRole/resourceProvider"/>
      </prov:Attribution>
    </prov:qualifiedAttribution>
    <dcat:contactPoint>
      <vcard:Kind>
        <vcard:organization-name
xml:lang="">National Data Buoy Center</
vcard:organization-name>
        <vcard:hasEmail
rdf:resource="mailto:"/>
      </vcard:Kind>
    </dcat:contactPoint>
    <prov:qualifiedAttribution>
      <prov:Attribution>
        <prov:agent>
          <vcard:Kind>
            <vcard:organization-name
xml:lang="">National Data Buoy Center</
vcard:organization-name>
            <vcard:hasEmail
rdf:resource="mailto:"/>
          </vcard:Kind>
        </prov:agent>
        <dct:type rdf:resource="http://
inspire.ec.europa.eu/codelist/
ResponsiblePartyRole/pointOfContact"/>
      </prov:Attribution>
    </prov:qualifiedAttribution>
    <dct:conformsTo>
      <dct:Standard>
        <dct:title xml:lang="en"/>
        <dct:issued rdf:datatype="http://
www.w3.org/2001/XMLSchema#date"/>
      </dct:Standard>
    </dct:conformsTo>
  </rdf:Description>
```

Conclusion

The availability of GI resources on the Web is increasing on a daily basis. This progress can be clearly seen on many geoportals around the world (e.g. on INSPIRE geoportal). This is mainly driven through overall SDI developments. On the other hand, also the number of interested

users is increasing at a similar pace. It brings us new ideas and new added value applications. However, for a non GIS expert it is not easy to understand how to search, find and efficiently use GI resources provided by OGC services. The most convenient way for non GIS users to search for something on the web is through either mainstream or semantic search engines. The outcome will be that most of the GI resources published in standardized SDI environments will not be discovered. This fact triggered many researchers to try and find better solutions of how to identify and utilize available spatial services. This paper described the idea and products of the Bolegweb platform aimed at addressing this issue. It provides various products to web communities, foremost a tabular view of the services available online deployed using OGC standards like Web Feature / Map / Tile / Observation Service, etc. Results reported in the paper already provide a solid basis of information about available GI resources on the global scale. On the Bolegweb project web page it is possible to get information about currently available OGC Web Services that have been discovered from the Google Search Engine together with their geographic distribution and basic statistics. It is possible to use other components of the platform as described in the paper and filter results by different attributes (e.g. title, type of service, version, server location etc.). In order to provide an intuitive and user friendly GUI, a pilot version of a discovery map portal of OGC resources metadata is provided. Simple full text and spatial queries are supported by defining the location of the point of interest and the map extent. This pilot is operational but still in the development phase, and is foreseen to provide more advanced functionalities in the near future. The contribution to other web communities is also supported by REST, SPARQL and CKAN interfaces, which have been deployed on the top of the Bolegweb platform GI resources. A likely outcome of the OGC/W3C Spatial Data on the Web working group is the definition of how to wrap those services in software that automatically creates human and machine readable Web pages for each of the information items behind the service. This would mean that not just the service itself, but also the data provided by the service would be discoverable, thus enabling tools such as the Metasearch Enhanced OGC Crawler to provide a much richer information coverage and user experience. The Bolegweb project ended in October 2016; and same system components are still in the development and pilot phase; however, they are functional and publicly available. The major future plans are to further develop the geospatial map portal of OGC resources with visualization, download and processing functionalities, with the main goal to reuse available OGC Services capabilities as much as possible.

Acknowledgements

Corresponding author:
Ing. Tomáš Kliment, PhD.
Faculty of Geodesy, University of Zagreb, Kačićeva 26, Zagreb, Croatia
Phone: + 385 (1) 4639 367, E-mail: tomas.kliment@geof.hr

References

[1]   Abargues, C., Granell, C., Díaz, L., Huerta, J. and Beltran, A. (2009) "*Discovery of User-Generated Geographic Data Using Web Search Engines*", Advances in Geoscience and Remote Sensing, Gary Jedlovec (Ed.), pp. 207-228. ISBN 978-9533070056.

[2]   Athanasiou, S., Georgomanolis, N., Patroumpas, K., Alexakis, M. and Stratiotis, T. (2015) "TripleGeo-CSW: A Middleware for Exposing Geospatial Catalogue Services on the Semantic Web", *EDBT/ICDT Workshops*, Vol. 1330, pp. 229-236. ISSN 1613-0073.

[3]   Baumann, P. (2010) "Beyond rasters: introducing the new OGC web coverage service 2.0", *GIS '10 Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 320-329. ISBN 978-1-4503-0428-3. DOI 10.1145/1869790.1869835.

[4]   Blower, J. D., Gemmell, A. L., Griffiths, G. H., Haines, K., Santokhee, A. and Yang, X. (2013) "A Web Map Service implementation for the visualization of multidimensional gridded environmental data", *Environmental Modelling & Software*, Vol. 47, pp. 218-224. ISSN 1364-8152. DOI 10.1016/j.envsoft.2013.04.002.

[5]   Díaz, L., Granell, C. and Huerta, J. (2012) "*Discovery of Geospatial Resources: Methodologies, Technologies and Emergent Applications*", IGI Global, 335 p. ISBN 978-1-4666-0945-7.

[6]   European Commission (2007) "Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)". Published in the official Journal on the 25th April, 2007.

[7]   Florczyk, A. J., López-Pellicer, F. J., Nogueras-Iso, J. and Zarazaga-Soria, F. J. (2012) "Automatic Generation of Geospatial Metadata for Web Resources", *International Journal of Spatial Data Infrastructures Research*, Vol.7, pp. 151-172. ISSN 1725-0463.

[8]   Giuliani, G., Guigoz, Y., Lacroix, P., Ray, N. and Lehmann, A. (2016) "Facilitating the production of ISO-compliant metadata of geospatial datasets", *International Journal of Applied Earth Observation and Geoinformation*, Vol. 44, pp. 239-243. ISSN 0303-2434.

[9]   Hou, D., Chen, J. and Wu, H. (2016) "Discovering Land Cover Web Map Services from the Deep Web with JavaScript Invocation Rules", *ISPRS International Journal of Geo-Information*, Vol. 5, No. 7. ISSN 2220-9964.

[10]   Kalantari, M., Rajabifard, A., Olfat, H., Pettit, C. and Keshtiarast, A. (2016) "Automatic spatial metadata systems–the case of Australian urban research infrastructure network", *Cartography and Geographic Information Science*, pp. 1-11. ISSN 1523-0406, E- ISSN 1545-0465.

[11]   Kliment, T., Bordogna, G. amd Frigerio, L. (2015). "A Geospatial Data Infrastructure Deploying an Information and Knowledge Platform for the Agriculture Sector", *Kartografija i geoinformacije (Cartography and Geoinformation)*, Vol. 14, No. 24, pp. 4-24. ISSN 1333-896X, E-ISSN 1848-0713.

[12]   Kliment, T., Cetl, V., Kliment, M. and Tuchyňa, M. (2015) "Making more OGC services available on the web discoverable for the SDI community", *15th International Multidisciplinary Scientific GeoConference SGEM 2015*, Book2, Vol. 2, pp. 955-962. ISBN 978-619-7105-35-3. ISSN 1314-2704.

[13]   Kliment, T., Granell, C., Cetl, V. and Kliment, M. (2013) "Publishing OGC resources discovered on the mainstream web in an SDI catalogue", *Proceedings of the 16th AGILE International Conference on Geographic Information Science, Leuven, Belgium*. ISBN 978-3-319-00615-4.

[14]   Li, Z., Yang, C. P., Wu, H., Li, W. and Miao, L. (2011). "An optimized framework for seamlessly integrating OGC Web Services to support geospatial sciences", *International Journal of Geographical Information Science*, Vol 25, No 4, pp. 595-613. ISSN 1365-8816, E-ISSN 1362-3087. DOI 10.1080/13658816.2010.484811.

[15]   López-Pellicer, F. J., Florczyk, A. J., Béjar, R., Muro-Medrano, P. R. and Zarazaga-Soria, F. J. (2011) "Discovering geographic web services in search engines", *Online information Review*, Vol. 35, No. 6, pp. 909-927. ISSN 1468-4527.

[16]   Michaelis, C. D. and Ames, D. P. (2009) "Evaluation and implementation of the OGC web processing service for use in client-side GIS", *Geoinformatica*, Vol 13, No 1, pp. 109-120. ISSN 1384-6175, E- ISSN 1573-7624. DOI 10.1007/s10707-008-0048-1.

[17]   Peng, Z. R. and Zhang, C. (2004). "The roles of geography markup language (GML), scalable vector graphics (SVG), and Web feature service (WFS) specifications in the development of Internet geographic information systems (GIS)", *Journal of Geographical Systems*, Vol. 6, No. 2, pp. 95-116. ISSN 1435-5930, E-ISSN 1435-5949. DOI 10.1007/s10109-004-0129-0.

[18]   Poorazizi, M. E., Hunter, A. J. S. and Steiniger, S. (2015) "A Volunteered Geographic Information Framework to Enable Bottom-Up Disaster Management Platforms", *ISPRS International Journal of Geo-Information*, Vol 4, pp. 1389-1422. E-ISSN 2220-9964.

[19]   Schleidt, K. (2013). "INSPIREd air quality reporting", *Environmental Software Systems. Fostering Information Sharing*, Vol 413, pp. 439-450. ISSN 1868-4238. DOI 10.1007/978-3-642-41151-9_41.

[20]   Sibolla, B., van Zyl, T., McFerren, G. and Hohls, D. (2014) "Adding temporal data enhancements to the advanced spatial data infrastructure platform", *Tenth International Conference of the African Association of Remote Sensing of the Environment (AARSE 2014)*, University of Johannesburg, South Africa.