Open Geospatial Consortium

Publication Date: 2015-10-01

Approval Date: 2015-08-03

Posted Date: 2015-06-23

Reference number of this document: OGC 15-066r1

Reference URL for this document: http://www.opengis.net/doc/PER/t11-linked-data-nhd

Category: Public Engineering Report

Editor(s): Gobe Hobona;Roger Brackin

Testbed-11 Use of Semantic Linked Data with RDF for National Map NHD and Gazetteer Data Engineering Report

Copyright © 2015 Open Geospatial Consortium.
To obtain additional rights of use, visit http://www.opengeospatial.org/legal/.

Warning

This document is not an OGC Standard. This document is an OGC Public Engineering Report created as a deliverable in an OGC Interoperability Initiative and is not an official position of the OGC membership. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an OGC Standard. Further, any OGC Engineering Report should not be referenced as required or mandatory technology in procurements.

Document type: OGC® Engineering Report
Document subtype: NA
Document stage: Approved for public release
Document language: English

License Agreement

Permission is hereby granted by the Open Geospatial Consortium, ("Licensor"), free of charge and subject to the terms set forth below, to any person obtaining a copy of this Intellectual Property and any associated documentation, to deal in the Intellectual Property without restriction (except as set forth below), including without limitation the rights to implement, use, copy, modify, merge, publish, distribute, and/or sublicense copies of the Intellectual Property, and to permit persons to whom the Intellectual Property is furnished to do so, provided that all copyright notices on the intellectual property are retained intact and that each person to whom the Intellectual Property is furnished agrees to the terms of this Agreement.

If you modify the Intellectual Property, all copies of the modified Intellectual Property must include, in addition to the above copyright notice, a notice that the Intellectual Property includes modifications that have not been approved or adopted by LICENSOR.

THIS LICENSE IS A COPYRIGHT LICENSE ONLY, AND DOES NOT CONVEY ANY RIGHTS UNDER ANY PATENTS THAT MAY BE IN FORCE ANYWHERE IN THE WORLD.

THE INTELLECTUAL PROPERTY IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, AND NONINFRINGEMENT OF THIRD PARTY RIGHTS. THE COPYRIGHT HOLDER OR HOLDERS INCLUDED IN THIS NOTICE DO NOT WARRANT THAT THE FUNCTIONS CONTAINED IN THE INTELLECTUAL PROPERTY WILL MEET YOUR REQUIREMENTS OR THAT THE OPERATION OF THE INTELLECTUAL PROPERTY WILL BE UNINTERRUPTED OR ERROR FREE. ANY USE OF THE INTELLECTUAL PROPERTY SHALL BE MADE ENTIRELY AT THE USER’S OWN RISK. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR ANY CONTRIBUTOR OF INTELLECTUAL PROPERTY RIGHTS TO THE INTELLECTUAL PROPERTY BE LIABLE FOR ANY CLAIM, OR ANY DIRECT, SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES, OR ANY DAMAGES WHATSOEVER RESULTING FROM ANY ALLEGED INFRINGEMENT OR ANY LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR UNDER ANY OTHER LEGAL THEORY, ARISING OUT OF OR IN CONNECTION WITH THE IMPLEMENTATION, USE, COMMERCIALIZATION OR PERFORMANCE OF THIS INTELLECTUAL PROPERTY.

This license is effective until terminated. You may terminate it at any time by destroying the Intellectual Property together with all copies in any form. The license will also terminate if you fail to comply with any term or condition of this Agreement. Except as provided in the following sentence, no such termination of this license shall require the termination of any third party end-user sublicense to the Intellectual Property which is in force as of the date of notice of such termination. In addition, should the Intellectual Property, or the operation of the Intellectual Property, infringe, or in LICENSOR’s sole opinion be likely to infringe, any patent, copyright, trademark or other right of a third party, you agree that LICENSOR, in its sole discretion, may terminate this license without any compensation or liability to you, your licensees or any other party. You agree upon termination of any kind to destroy or cause to be destroyed the Intellectual Property together with all copies in any form, whether held by you or by any third party.

Except as contained in this notice, the name of LICENSOR or of any other holder of a copyright in all or part of the Intellectual Property shall not be used in advertising or otherwise to promote the sale, use or other dealings in this Intellectual Property without prior written authorization of LICENSOR or such copyright holder. LICENSOR is and shall at all times be the sole entity that may authorize you or any third party to use certification marks, trademarks or other special designations to indicate compliance with any LICENSOR standards or specifications.

This Agreement is governed by the laws of the Commonwealth of Massachusetts. The application to this Agreement of the United Nations Convention on Contracts for the International Sale of Goods is hereby expressly excluded. In the event any provision of this Agreement shall be deemed unenforceable, void or invalid, such provision shall be modified so as to make it valid and enforceable, and as so modified the entire Agreement shall remain in full force and effect. No decision, action or inaction by LICENSOR shall be construed to be a waiver of any rights or remedies available to it.

None of the Intellectual Property or underlying information or technology may be downloaded or otherwise exported or reexported in violation of U.S. export laws and regulations. In addition, you are responsible for complying with any local laws in your jurisdiction which may impact your right to import, export or use the Intellectual Property, and you represent that you have complied with any regulations or registration procedures required by applicable law to make this license enforceable

 Copyright © 2015 Open Geospatial Consortium.

Contents

Page

Copyright © 2015 Open Geospatial Consortium.

Figures

Page

Tables

Page

 Copyright © 2015 Open Geospatial Consortium.

Abstract

Over the past few years there has been an increase in the number, size and complexity of databases across government sectors. This has undoubtedly created challenges relating to the discovery and access of information and services on multiple databases across static and deployed networks. Linked Data has been suggested as a method able to tackle those challenges. The aim of the Hydrographic Linked Data activity in the OGC Testbed 11 was to advance the use of Linked Data for hydrographic data by building on the achievements of the previous testbeds and to improve the understanding of how to better build relations between hydro features and non-hydro features (e.g., stream gauge measurement/location vs bridge or other built features upstream or downstream). This aspect of the testbed focused on the National Hydrography Dataset (NHD) which is published by the United States Geological Survey (USGS). This OGC Engineering Report provides guidelines on the publication of hydrographic and hydrological data serialized as Resource Description Framework (RDF) using Linked Data principles and technologies based on OGC standards. The document also presents the experimentation conducted by Testbed 11 in order to identify those guidelines.

Business Value

This OGC Engineering Report describes approaches that could improve semantic interoperability by enhancing the ability of data consumers to discover data that is associated with other data, thereby providing insight beyond that offered by any single dataset. The content of the engineering report is important to achieving interoperability in location-based technologies because it offers an approach that could add value to existing and future geospatial data products.

Keywords

ogcdocs, ogc documents, testbed-11, hydrography, hydrology, semantic web, linked data, rdf

Copyright © 2015 Open Geospatial Consortium.

Testbed-11 Use of Semantic Linked Data with RDF for National Map NHD and Gazetteer Data Engineering Report

1 Introduction

1.1 Scope

Over the past few years there has been an increase in the number, size and complexity of databases across government sectors. This has undoubtedly created challenges relating to the discovery and access of information and services on multiple databases across static and deployed networks. Linked Data has been suggested as a method able to tackle those challenges.

Linked Data presents a method of publishing structured data so that it can be interlinked and become more practical through use of semantic queries [7]. The World Wide Web Consortium (W3C) also defines Linked Data as “a way to create a network of standards-based machine interpretable data across different documents and Web sites” [6]. Linked Data can therefore be considered a specialization of the Semantic Web. Linked Data applies standard Semantic Web technologies such as the Resource Description Framework (RDF) and the Web Ontology Language (OWL). However, instead of using them to serve web pages for human readers, it uses them to share information in a way that can be processed by machines. This method has established a capability within the World Wide Web in which not just hyperlinked documents are accessible to the world, but also primary data can be connected and queried.

The aim of the Hydrographic Linked Data aspect of OGC Testbed-11 was to advance the use of Linked Data for Hydrographic Data by building on the achievements of the previous testbeds and improving the understanding of how to better build relations between hydro features and non-hydro features (e.g., stream gauge measurement/location vs bridge or other built features upstream or downstream). Such advancement could help with addressing several challenges relating to the enduring need to associate data with other data in order to derive useful information and infer new insight. This OGC Engineering Report provides guidelines on the publication of hydrographic and hydrological data using Linked Data principles applied to technologies based on OGC standards. The document also presents the experimentation conducted by the testbed.

This Engineering Report is applicable to initiatives involving hydrographic data production, hydrological science and Linked Data (e.g. data.gov and data.gov.uk).

Copyright © 2015 Open Geospatial Consortium.

1.2 Background

Since its conception by Tim Berners-Lee in 2006, the vision of Linked Data has established a sizable community behind it. Figure 1 represents the different initiatives providing documents into the global Linked Data space, which shall hereinafter be referred to in this report as the Linked Data Cloud. A more legible figure can be found at http://lod-cloud.net/

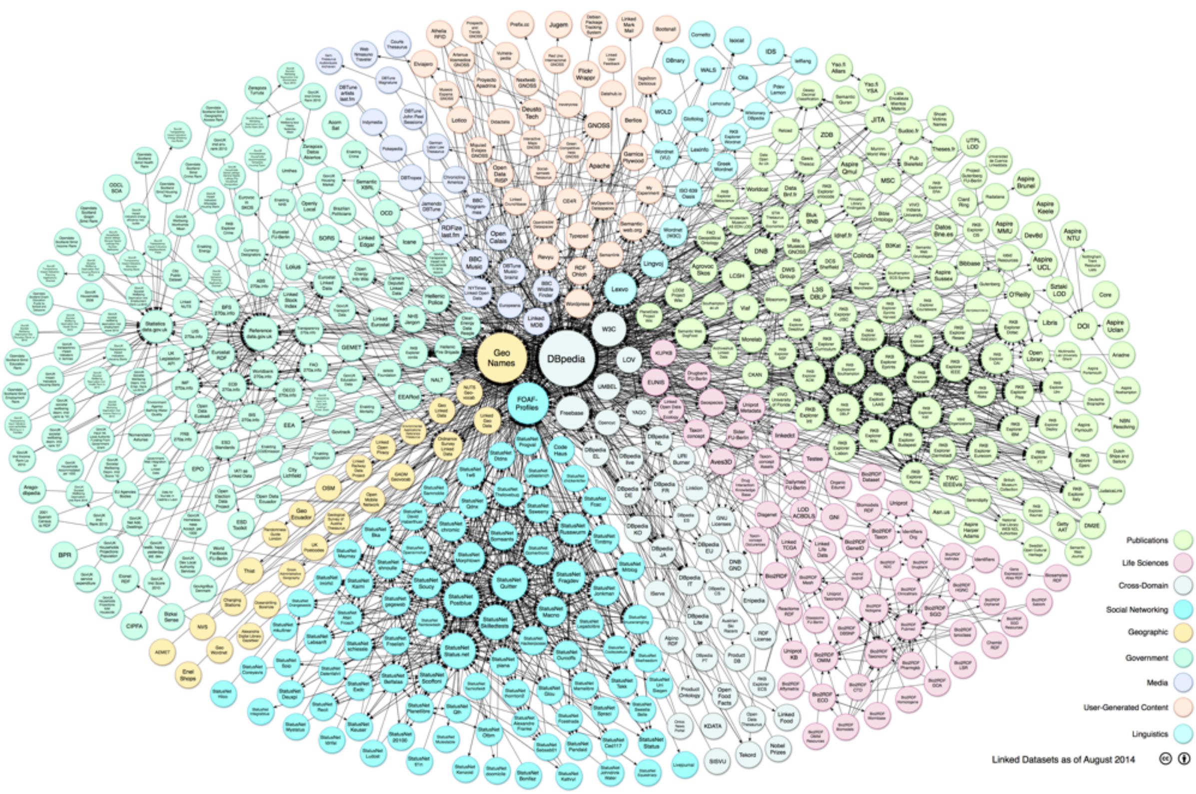

Figure 1. The Linked Data Cloud (source: http://lod-cloud.net/)

Several of the organizations that have published Linked Data are government departments within the United States (US) and in Europe. For example, the data.gov.uk program has published over 5400 datasets from government departments across the United Kingdom, and the data.gov site has made 6.4 billion triples of US government data available for exploitation [8].

1.2.1 Previous Work

Linked Data has been seen as a potential enabler of cross community interoperability for several years and previous OGC testbeds have provided some insight into how such enablement could be achieved. This section presents summaries of the work conducted in the most previous testbed.

 Copyright © 2015 Open Geospatial Consortium.

The OGC Web Services Testbed 10 (OWS-10) included multiple research activities related to Linked Data and the Semantic Web [4]. The first activity involved the investigation of the potential for a Virtual Global Gazetteer (VGG) that integrated two gazetteers: the USGS Geographic Names Information System (GNIS) gazetteer and the GEOnet Names Server (GNS) gazetteer of the National Geospatial Intelligence Agency (NGA). The VGG provided the capability to link types of places offered by one gazetteer with place types offered by another. Semantic mappings used by the VGG were served by a SPARQL Server, whereas the instances of places were provided by web services based on the Gazetteer profile of the Web Feature Service (WFS-G) specification. The semantic mappings allowed the VGG to offer semantic mediation between the USGS and NGA WFS-G services.

A second activity was concerned with ontology engineering involving an examination of ontologies in the context of OGC data modeling, handling, and organization [2]. The testbed defined a set of ontologies implementing solid theoretical foundations and semantics. Another aspect of the ontology engineering involved an examination of the definition of core ontologies for representing incident information within the emergency and disaster management domain.

A third testbed activity addressed ontology mapping between hydrology feature models [5]. The activity had a goal of advancing interoperability of approaches for sharing geospatial data within hydro communities and to also advance semantic mediation approaches for data discovery, access, and use of heterogeneous hydro data models (and heterogeneous hydro metadata models).

1.2.2 Related Work

Several initiatives in the hydrology community have also been exploring approaches for publishing hydro data as Linked Data. An example of such an initiative is by the Centre of Excellence in Geographic Information Science (CEGIS) at the USGS which undertook a pilot study to make USGS data available to the Semantic Web and the Linked Open Data Community[1]. CEGIS converted USGS data to RDF and Geography Markup Language (GML) for a group of test areas. The conversion processes adopted by CEGIS involved extracting all data for eight layers of The National Map for the test areas, and converting the point vector data to RDF whilst maintaining the coordinates in GML. The triples (formed of a subject, predicate and object) were constructed from the entities defined in the National Map. An example of an NHD entity is a flowline, which can be used to represent a stream reach that provides connections within a hydrographic network. In the case of an NHD flowline the subject is the feature identifier derived from the reach code, the predicate is the particular characteristic of the flowline (e.g. its length), and the objects can be literal values or references to other features. The CEGIS approach modelled the geometry objects as containing GML coordinates of the flowline.

The approach adopted by CEGIS has the following advantages:

---

[1] http://www.semantic-web-journal.net/sites/default/files/swj180.pdf

Copyright © 2015 Open Geospatial Consortium.

- Allows most (if not all) of the information offered by NHD data to be preserved by the generated RDF.
- Results in a graph structure that can be traversed between different NHD feature types, through use of the same property to represent unique identifiers across different feature types.
- Allows locations to be expressed using the complete expressivity of GML.
- Allows for spatial queries using GeoSPARQL.

OGC Testbed 11 offered an opportunity to partially align the approach adopted by CEGIS with other information models adopted in the hydro community; for example, the OGC Hy_Features model[2]. In addition to such alignment, the testbed provided an opportunity to examine the potential for publishing NHD data as Linked Data consistent with guidelines adopted in the wider Linked Data community.

1.3 Document contributor contact points

All questions regarding this document should be directed to the editor or the contributors:

Name	Organization
Gobe Hobona PhD.	Envitia Ltd.
Roger Brackin MSc.	Envitia Ltd.
Stefano Cavazzi PhD.	Envitia Ltd.
Barbara Klis	Envitia Ltd.
Stephane Fellah	Image Matters
Josh Lieberman PhD.	OGC/Tumbling Walls
Dean Hintz	Safe Software
Buck Shou	Feng Chia University (GIS FCU)
Chen-Yu (How) Hao	Feng Chia University (GIS FCU)
Eugene Yu PhD.	George Mason University (GMU)
Lingjun Kang MS.	George Mason University (GMU)
David Blodgett	US Geological Survey
David Wesloh	National Geospatial Intelligence Agency

[2] http://www.opengis.net/doc/DP/hy-features

 Copyright © 2015 Open Geospatial Consortium.

1.4 Future work

Improvements in this document are not planned.

1.5 Forward

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium (OGC) shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

2 References

The following documents are referenced in this document. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the latest edition of the normative document referred to applies.

1. OGC 06-121r3, *OpenGIS® Web Services Common Standard*
2. OGC 14-049, *OGC® Testbed 10 Cross Community Interoperability (CCI) Ontology Engineering Report*
3. OGC 12-103r3, *OWS-9 Semantic Mediation Engineering Report*
4. OGC 14-029r2, *OWS-10 Gazetteer Engineering Report*
5. OGC 14-048. *Testbed 10 Cross Community Interoperability (CCI) Hydro Model Interoperability Engineering Report*
6. W3C. JSON-LD 1.0, last visited 20-04-2015, available from http://www.w3.org/TR/json-ld/
7. Berners-Lee T., (2009) Linked Data, last visited 20-04-2015, available from *http://www.w3.org/DesignIssues/LinkedData.html*
8. Hart, G., Dolbear, C. (2013) Linked Data : A Geographic Perspective, CRC Press
9. Howard M., Payne S., Sunderland R., (2010) Technical Guidance for the INSPIRE Schema Transformation Network Service available from http://inspire.jrc.ec.europa.eu/documents/Network_Services/JRC_INSPIRE-TransformService_TG_v3-0.pdf
10. US EPA. NHDPlus Version 2: User Guide (Data Model Version 2.1), 2015

Copyright © 2015 Open Geospatial Consortium.

3 Terms and definitions

For the purposes of this report, the definitions specified in Clause 4 of the OWS Common Implementation Specification [OGC 06-121r3] and in OpenGIS® Abstract Specification shall apply. In addition, the following terms and definitions apply.

3.1
feature
representation of some real world object or phenomenon

3.2
interoperability
capability to communicate, execute programs or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units

3.3
metadata
data about data

3.4
model
abstraction of some aspects of a universe of discourse

3.5
ontology
a formal specification of concrete or abstract things, and the relationships among them, in a prescribed domain of knowledge [ISO/IEC 19763]

3.6
semantic interoperability
the aspect of interoperability that assures that the content is understood in the same way in both systems, including by those humans interacting with the systems in a given context

3.7
syntactic interoperability
the aspect of interoperability that assures that there is a technical connection, i.e. that the data can be transferred between systems

4 Conventions

4.1 Abbreviated terms

CCI	Cross Community Interoperability
CSW	Catalogue Service for the Web

Copyright © 2015 Open Geospatial Consortium.

ER	Engineering Report
GML	Geography Markup Language
HTML	HyperText Markup Language
JSON-LD	JavaScript Object Notation for Linked Data
NGA	National Geospatial Intelligence Agency
NHD	National Hydrographic Dataset
NSG	National System for Geospatial Intelligence
OGC	Open Geospatial Consortium
OWL	Web Ontology Language
RDF	Resource Description Framework
SDI	Spatial Data Infrastructure
TNM	The National Map
URI	Unique Resource Identifier
URL	Uniform Resource Locator
URL	Uniform Resource Name
USGS	US Geological Survey
WFS	Web Feature Service
WKT	Well Known Text
WMS	Web Map Service

5 Methodology

To achieve the research aims outlined above, the experimentation conducted in the testbed followed the following process:

1. Identify data and services to support the use case
2. Design ontology, reusing where possible standard or commonly used predicates

Copyright © 2015 Open Geospatial Consortium.

3. Configure individual components to support, convert or generate the Linked Data
4. Integrate the individual components into the testbed demonstrator
5. Assess interoperability and identify lessons to learn

6 National Hydrography Dataset (NHD)

The NHD is a digital spatial data product that offers information about naturally occurring and manmade water-bodies, water flowpaths and related features. The NHD offers a variety of information about such features, including classification, delineation, geographic name, position, and other characteristics. Also provided is a "reach code" through which other information can be related to the NHD.

Several different applications exploit NHD data, for example:

- Map making.
- Geocoding and linking through the reach code.
- Modeling the flow of water.
- Providing a reference of unique identifiers for features found in the NHD.

The NHD adopts an object oriented model within which entities are modelled as features. Features, within the NHD, include naturally occurring and manmade water-bodies, water flowpaths and related features. Feature types such as “stream/river”, “canal/ditch” and “lake/pond” are defined through grouping and classification of features that share specific characteristics. The features are represented geometrically as points, lines and polygons following a process of delineation according to the following rules:

1. The delineated feature must be contiguous with related features.
2. The delineated feature must have consistent dimensionality; that is, it must be consistently one point, one or more lines, or one or more areas.
3. The delineated feature can have only one feature type and must have the same set of characteristics and choices of values throughout its extent.

Unique identifiers are used extensively within the NHD to reference instances of feature types. Whereas some software applications have been implemented with built-in logic for associating NHD feature types, a platform-independent approach offers the potential for interoperability between different applications.

The following sections describe the unique identifiers that the NHD offers for cross referencing.

 Copyright © 2015 Open Geospatial Consortium.

6.1.1 Common identifier

Each feature is uniquely identified through a 10-digit integer known as the “common identifier”. The common identifier is held in an attribute named “COMID” in the datasets used in the testbed. A common identifier number is permanently assigned to a feature and is thus retired when that feature is deleted.

6.1.2 Reachcode

In hydrology, a ‘reach’ is a continuous, unbroken stretch or expanse of surface water. The NHD expands this definition to define a reach “as a significant segment of surface water that has similar hydrologic characteristics, such as a stretch of stream/river between two confluences, or a lake/pond”[3]. Reaches are assigned unique 14-digit numeric labels known as a “reach code”. The first 8 digits of a reach code indicate the hydrologic unit for the sub-basin in which the reach exists, whereas the last 6 digits are assigned arbitrarily and sequentially among the reaches. A reach code is permanently assigned to a reach and is retired when that reach is deleted. Reach code can be effective in associating observation data to other information. For example, observations can be referenced to an entire reach by tagging the observations with the reach code, or to sections of a transport, coastline, or shoreline reach. Reach codes are held in attributes named "REACHCODE" in datasets used in the testbed.

7 Principles of Linked Data

Whereas the primary units of the World Wide Web can be considered to be hypertext documents encoded in HyperText Markup Language (HTML) and linked by untyped hyperlinks, the primary units of the Linked Data Cloud are documents holding data in RDF format and linked through typed statements. The resultant structure has in some cases been referred to as the Web of Data.

The RDF data model is designed to support the representation of information that is integrated from multiple sources, represented using different schemas, and is heterogeneously structured. Encoding data in RDF allows Linked Data applications to reason about the meaning of the data and make inferences based on the assertions specified in the data (and other data linked to it). This ability is not supported by traditional web documents encoded in HTML, which only allow applications to interpret the formatting specified by the documents. Such reasoning is made possible by the standardized encoding of concepts and the relationships between them in such a way that makes it possible for different applications to adopt a consistent understanding of the meaning of the concepts. Such formal specifications of concepts and the relationships between them are commonly referred to as ontologies, and are typically encoded in OWL.

---

[3] http://nhd.usgs.gov/chapter1/chp1_data_users_guide.pdf

Copyright © 2015 Open Geospatial Consortium.

OWL is considered the standard language for defining and instantiating ontologies on the World Wide Web. An ontology specified in OWL may include descriptions of classes, properties, instances and the relationships between them. The cross referencing of some instances based on a common vocabulary of class and properties results in a graph that allows applications to traverse through from one dataset to another, provided the datasets are linked. The linking of classes, properties and their instances in the Linked Data Cloud (and the ontologies that support the Cloud) relies on unique resource identifiers (URI).

A URI is a string of characters assigned to a resource and no other resource such that the referencing of that string of characters unambiguously refers to that resource. There are generally considered to be two types of URI, namely uniform resource names (URN) and Uniform Resource Locators (URL). A URN is generally used to identify a resource (such as a coordinate system), whereas a URL provides an address through which a resource can be accessed on the World Wide Web. An example of a URN is "urn:ogc:def:crs:EPSG:6.9:4326" and an example of a URL is http://spatialreference.org/ref/epsg/wgs-84/ . URLs are therefore often referred to as HTTP URI to distinguish them from URNs. The act of retrieving data using an HTTP URI is known as dereferencing that URI.

Neither OWL nor RDF mandates the use of HTTP URIs over URN. Moreover, neither of these standards requires that it be possible to deference a URI. As the ability to dereference an HTTP URI is fundamental to being able to traverse the Linked Data Cloud, the following guidelines have been proposed by Tim Berners-Lee (2006):

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names
3. When someone looks up a URI, provide useful information, using standards like RDF and SPARQL
4. Include links to other related URIs, so that they can be discovered, thus expanding the concept of the Web of Data

These four points are considered to be the "Linked Data Principles" and provide fundamental instructions for publishing and linking data using the infrastructure of the Web while maintaining its architecture and standards at the same time.

8 Use Cases Adopted

8.1 Find a Placename, Return Related Flowlines and/or Gauges

A user would like to find all flowlines and gauges that pass through a specific place. In a flood scenario, this use case would be applied in the strengthening of flood defenses along streams that are likely to affect towns at risk.

 Copyright © 2015 Open Geospatial Consortium.

Actors:

- Client.
- WFS-G Semantic Mediator.
- USGS NHD WFS.
- NationalMap Geonames WFS.
- CSW 3.0.
- GeoSPARQL Server.

Basic Steps

1. The client retrieves metadata from the CSW 3.0.
2. The client sends a request to a WFS-G semantic mediator.
3. WFS-G semantic mediator retrieves mappings from the GeoSPARQL Server.
4. WFS-G semantic mediator applies mappings to the request from the Client.
5. WFS-G semantic mediator retrieves places from the NationalMap Geonames WFS.
6. WFS-G semantic mediator returns the places to the client.
7. Client retrieves flowlines from the GeoSPARQL Server.
8. Client retrieves stream gauge locations from the GeoSPARQL Server.
9. Client links received features and presents output.

An illustration of the use case using a sequence diagram to describe the interactions between the components is presented in Figure 2.

Copyright © 2015 Open Geospatial Consortium.

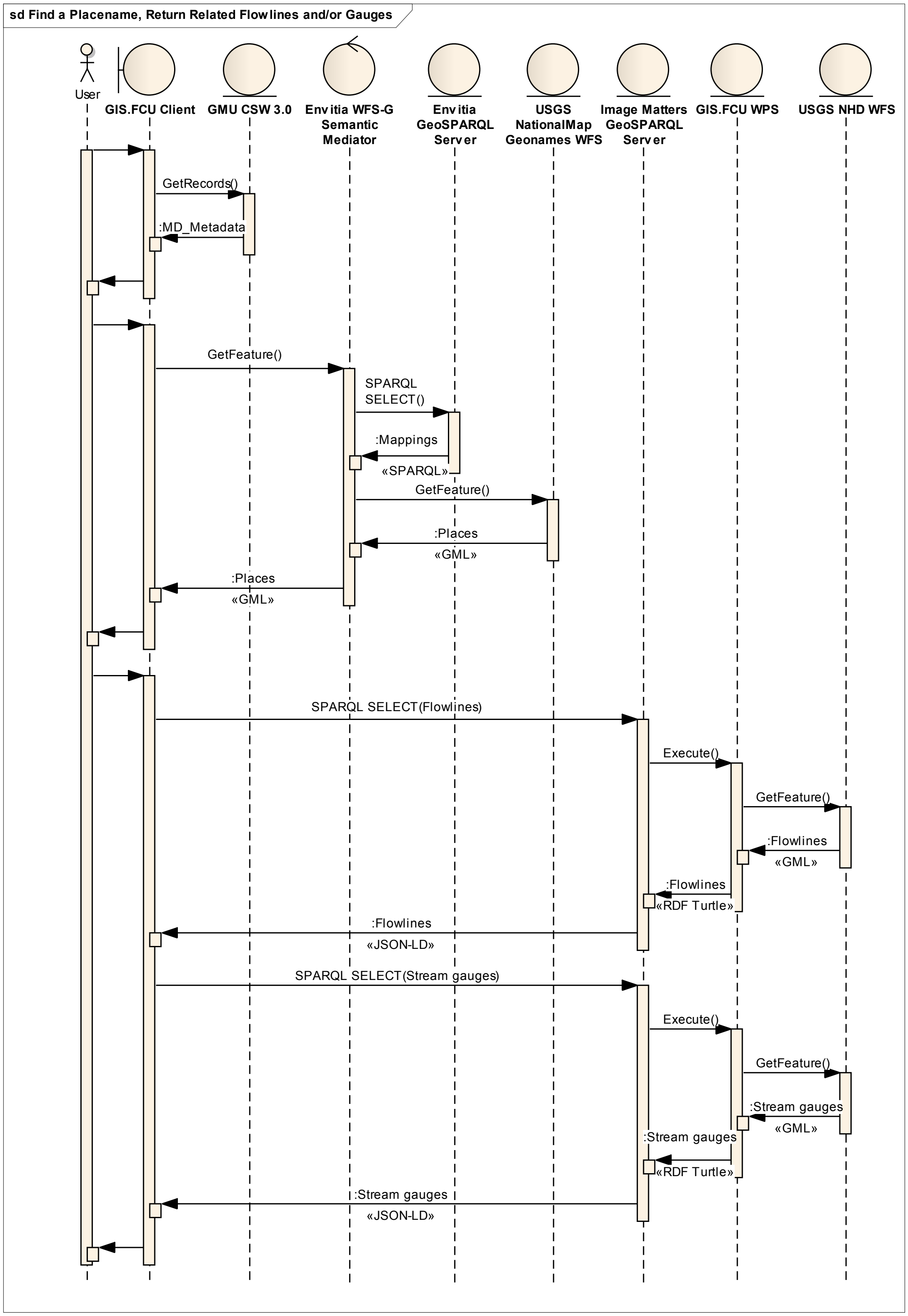

Figure 2. Sequence diagram for the first use case

 Copyright © 2015 Open Geospatial Consortium.

8.2 Find a Flow Line as a Place, Return Related Other Place Names

A user would like to find all places along a specific flowline. In a flood scenario, this use case could be applied in initial response efforts such as identifying the towns that are likely to be affected by a stream that is about to breach its banks.

Actors:

- Client.
- WFS-G Semantic Mediator.
- USGS NHD WFS.
- NationalMap Geonames WFS.
- CSW 3.0.
- GeoSPARQL Server.

Basic Steps

1. The client retrieves metadata from the CSW3.0.
2. The client sends a request for flowlines to the USGS NHD WFS.
3. The client retrieves flowlines from NHD WFS.
4. The client selects a stream of interest.
5. The client sends a request with the bounds of the selected stream to the WFS-G semantic mediator.
6. WFS-G semantic mediator retrieves mappings from the GeoSPARQL Server.
7. WFS-G semantic mediator applies mappings to the request from the Client.
8. WFS-G semantic mediator retrieves places from the NationalMap WFS.
9. WFS-G semantic mediator returns the places to the client.
10. The client retrieves places from the GeoSPARQL Server.
11. The client presents the output.

Copyright © 2015 Open Geospatial Consortium.

An illustration of the use case using a sequence diagram to describe the interactions between the components is presented in Figure 3.

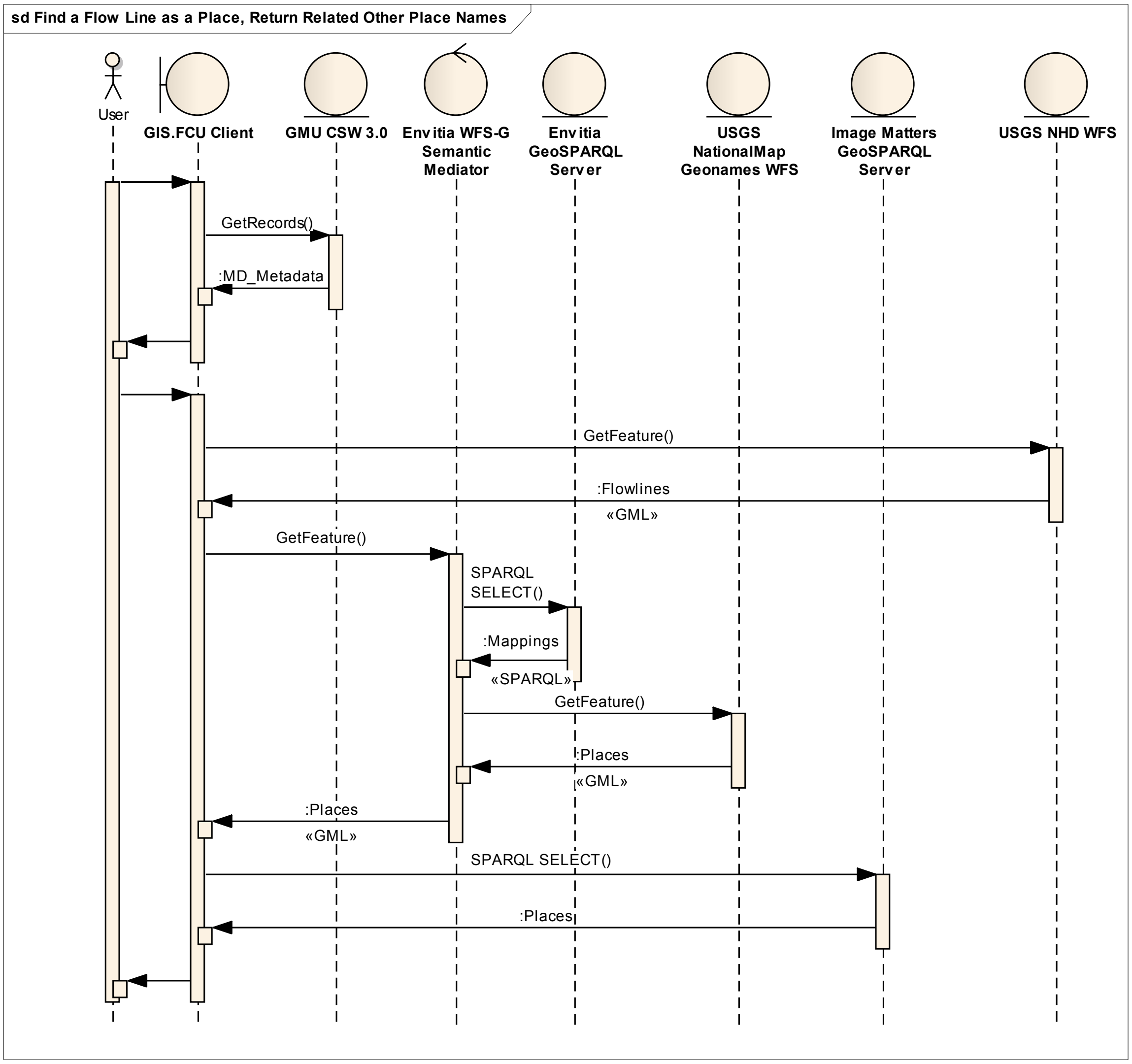

Figure 3. Sequence diagram for the second use case

9 Testbed Architecture

The architecture adopted by the Cross Community Interoperability (CCI) thread for Linked Data is shown in Figure 4. The GIS-FCU client application retrieves metadata from the GMU CSW. The metadata is structured according to the ISO 19115 and 19119 international standards. The client application obtains places from the Envitia WFS-G[4] Semantic Mediator, which in turn retrieves places from external WFS including the

[4] Web Feature Service-Gazetteer

 Copyright © 2015 Open Geospatial Consortium.

USGS WFS. The WFS-G makes use of semantic mappings obtained from the Envitia GeoSPARQL Server. The WFS-G returns GML-encoded places that reference their RDF-encoded equivalent places from the Image Matters GeoSPARQL Server (derived from the USGS gazetteer). The places returned by the WFS-G also reference hydrographic features from the Envitia GeoSPARQL Server. Within this architecture the GIS-FCU WPS and Safe Software Workbench are used for generating RDF-encoded data.

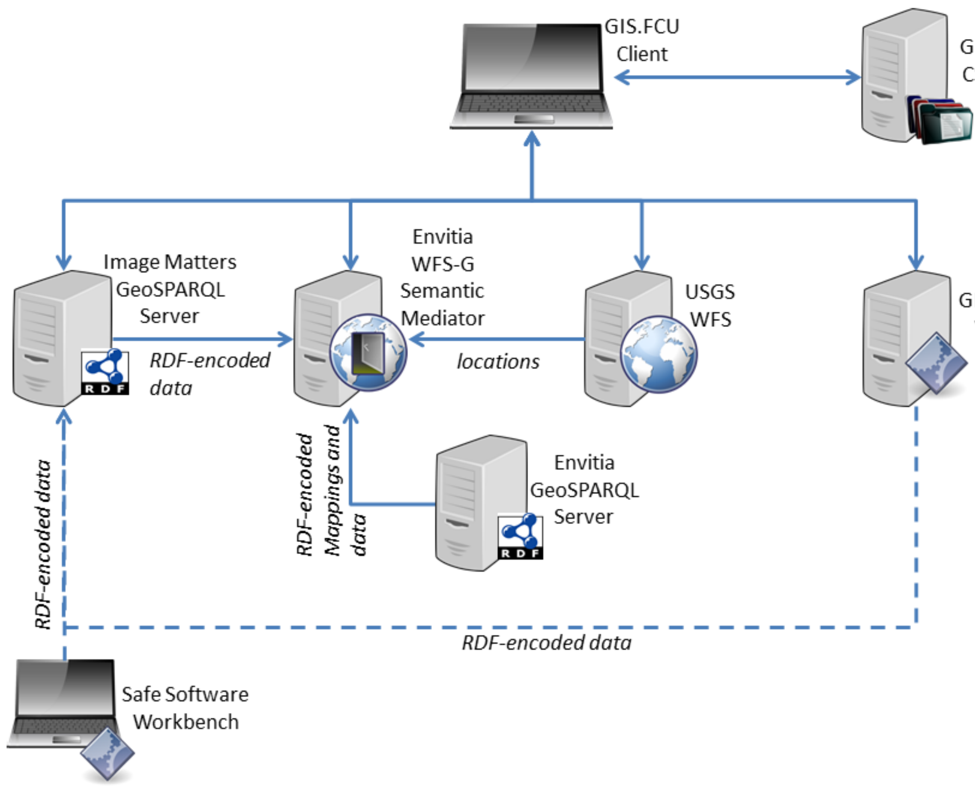

Figure 4. Architecture for CCI Linked Data components

10 Implementation

This section describes the implementation of the architecture, the issues encountered and lessons identified.

10.1 Encoding the Linked Data

The RDF model encodes data in the form of subject, predicate, object triples. The subject can be a URI or a blank node. The predicate is always a URI. The object can be a URI or a blank node. RDF triples can be grouped into two principal types, namely literal triples and link triples. Literal triples have, as a target object, a literal value such as a string, number, date or a Well Known Text (WKT) geometry. Link triples describe the

Copyright © 2015 Open Geospatial Consortium.

relationship between two resources. Link triples can be further categorized according to whether they associate resources within the same namespace (e.g. within the USGS namespace) or whether they associate resources in different namespaces (e.g. between namespaces of the USGS and NGA).

The predicate describes how the subject and object are related and is also represented by a URI. Predicate URIs come from vocabularies associated with particular purposes or communities of interest. Collections of predicate URIs can therefore be used to structure information relating to a particular domain. The links represented as predicate URIs can indicate relationship, identity and vocabulary links. Relationship links associate related entities such as a 'dam' on a 'river' that passes through a 'place'. Identity links associate objects with their names or labels, for example, some rivers might have separate local and official names. Vocabulary links offer descriptions of data to make data self-descriptive.

The Web of data, formed of Linked Data, can be seen as an additional layer that is interlaced with the classic document Web and has many of the same characteristics:

- Linked Data is universal and can include any type of data.
- Anyone can publish Linked Data.
- There are no constraints on data publishers on choice of vocabularies for describing data.
- Entities are connected by RDF links, constructing a global data graph that spreads data sources.
- Linkage between entities enables the discovery of new data resources.

There are various languages for encoding RDF triples by. The following sub-sections discuss the three languages used in the testbed for encoding RDF.

10.1.1 Turtle Language[5]

The Turtle language (TTL) allows RDF graphs to be written down in a compact natural text form within documents. The approach uses abbreviations for common usage patterns and datatypes, thereby making the documentation of RDF graphs more efficient than the XML-based alternative.

10.1.2 RDF/XML

The RDF/XML syntax is an application schema for writing RDF graphs in XML, using XML constructs — element names, attribute names, element contents and attribute

[5] http://www.w3.org/TeamSubmission/turtle/

 Copyright © 2015 Open Geospatial Consortium.

values. RDF/XML uses namespaces in XML to abbreviate URIs, however documents written in RDF/XML are often much larger than those written in other languages such as Turtle.

10.1.3 JSON-LD

JSON-LD is a syntax for serializing Linked Data into JavaScript Object Notation (JSON). It is mainly designed for use within web-based environments, particularly where JavaScript is supported by default. Although JSON-LD is more compact than RDF/XML, its efficiency is comparable to that offered by TTL. In contrast to TTL however, JSON-LD offers the benefit of built-in support by JSON parsers that are already available on popular web browsers.

10.2 Identification of Data Sources

A set of data sources in the form of feature collections served through WFS were provided by the USGS. After a review of the feature types available, the testbed selected feature types for stream gauges, flowlines and catchments for implementing the use cases described above.

Table 1 presents attributes of the stream gage feature type as exported from the WFS DescribeFeatureType response[6] and defined in the NHDPlus User Guide[7].

Table 1. NHDPlus stream gage attributes offered by the USGS WFS

Attributes	Type	Definition	Comment
the_geom	geometry	Location of the NHD feature	Used to generate GeoSPARQL as WKT
comid	Integer	Common identifier of the NHD feature	Value across all features was Nil so was not used as an identifier
eventdate	dateTime	Date event was created	Generated as a literal
reachcode	string	Reachcode on which Stream Gage is located. Converted to a resource URI	Used to generate a URI to reference the related Reach
reachsmdat	dateTime	Reach Version Date - Not populated	Generated as a literal

6 http://cida-test.er.usgs.gov/nhdplus/geoserver/ows?request=DescribeFeatureType&service=WFS&TypeName=nhdPlus:gage

7 http://www.fws.gov/r5gomp/gom/nhd-gom/NHDPLUS_UserGuide.pdf

Copyright © 2015 Open Geospatial Consortium.

reachresol	string	Reach Resolution, always “Medium”	Generated as a literal
featurecom	integer	Reserved for future use	Generated as a literal
featurecla	integer	Reserved for future use	Generated as a literal
source_ori	string	Originator of Event	Generated as a literal
source_dat	string	Description of point entity	Generated as a literal
source_fea	string	Gage Id/USGS Site Number	Generated as a literal
featuredet	string	URL where detailed gage data can be found (NWISWEB)	Generated as a literal
measure	double	Measure along reach where Stream Gage is located in percent from downstream end	Generated as a literal
offset	double	Always zero	Generated as a literal
eventtype	string	Fixed value of “StreamGage”	Generated as a literal
flcomid	integer	Internal identifier	Generated as a literal
agency_cd	string	Gov. Agency responsible for the streamgage	Generated as a literal
station_nm	string	Station name	Generated as a literal
state_cd	string	2 digit state FIPS code of the WSC maintaining the gage. Puerto Rico is listed as a state	Generated as a literal
state	string	2 char. state postal abrev. of the WSC maintaining the gage. Puerto Rico is listed as a state.	Generated as a literal
dasqmile	double	Reported drainage area in square miles. Stations with DA of -999999 means there is no reported DA in NWIS.	Generated as a literal

Copyright © 2015 Open Geospatial Consortium.

lonsite	double	Longitude of the streamgage (site) location - gage house in decimal degrees, NAD83	Generated as a literal
latsite	double	Latitude of the streamgage (site) location - gage house in decimal degrees, NAD83	Generated as a literal

Table 2 presents attributes of the flowline feature type as exported from the WFS DescribeFeatureType response[8] and defined in the NHDPlus User Guide.

Table 2. NHDPlus flowline attributes offered by the USGS WFS

Attribute	Type	Definition	Comment
the_geom	geometry	Location of the flowline	Used to generate GeoSPARQL as WKT
comid	integer	Common identifier of the NHD feature	Used to create a URI to reference the Flowline
fdate	dateTime	Feature Currency Date	Generated as a literal
resolution	string	Always "Medium"	Generated as a literal
gnis_id	string	Geographic Names Information System ID for the value in GNIS_Name	Generated as a literal
gnis_name	string	Feature Name from the Geographic Names Information System	Generated as a literal
lengthkm	double	Feature length in kilometers	Generated as a literal
reachcode	string	Reach Code assigned to feature	Used to generate a URI to reference the related Reach
flowdir	string	Flow direction is "WithDigitized" or "Uninitialized"	Generated as a literal
wbareacomi	integer	ComID of an NHD polygonal water feature through which an	Generated as a literal

[8] http://cida-test.er.usgs.gov/nhdplus/geoserver/ows?request=DescribeFeatureType&service=WFS&TypeName=nhdPlus:nhdflowline_nonnetwork

Copyright © 2015 Open Geospatial Consortium.

		NHD "Artificial Path" flowline flows	
ftype	string	NHD Feature Type	Generated as a literal
fcode	integer	Numeric codes for various feature attributes in the NHDFCode lookup table	Generated as a literal
shape_leng	double	Feature length	Generated as a literal
enabled	string	Always "True"	Generated as a literal

Table 2 presents attributes of the catchment feature type exported from the WFS DescribeFeatureType response[9] and defined in the NHDPlus User Guide.

Table 3. NHDPlus catchment attributes offered by the USGS WFS

Attribute	Type	Definition	Comment
the_geom	geometry	Location of the catchment	Used to generate GeoSPARQL as WKT
gridcode	integer	a unique identification number for each catchment	Generated as a literal
featureid	integer	Common Identifier of the Feature	Used to generate a URI for the feature
sourcefc	string	Source Feature Class	Generated as a literal
areasqkm	double	Feature area in square kilometers	Generated as a literal
shape_leng	double	Feature length	Generated as a literal
shape_area	double	Feature area	Generated as a literal

[9] http://cida-test.er.usgs.gov/nhdplus/geoserver/ows?request=DescribeFeatureType&service=WFS&TypeName=nhdPlus:catchment

Copyright © 2015 Open Geospatial Consortium.

10.3 Configure Individual Components

10.3.1 CSW 3.0

The testbed participants deployed a component based on version 3.0 of the Catalogue Service for the Web (CSW) standard. The CSW 3.0 component provided the ability to publish and search collections of metadata records for geospatial data, services and other resources. Metadata describe resource characteristics in way that enables CSW to query and present the characteristics for evaluation and discovery by both humans and applications.

The CSW 3.0 instance deployed in the testbed offered a variety search interfaces. An example request using a typical key value pair (KVP) request is:

http://www.exampleserver.com/cat3/csw?service=CSW&version=3.0.0&request=GetRecords&resultType=results&ElementSetName=full&outputSchema=http://www.opengis.net/cat/csw/2.0.2&typenames=csw:Record&outputFormat=application/xml&startPosition=1&maxRecords=10&constraintlanguage=CQL_TEXT&constraint=csw%3AAnyText%20Like%20%27%25water%25%27

The response to this query from the CSW 3.0 is shown in the following listing:


```xml
<?xml version="1.0" encoding="UTF-8"?>
<csw:GetRecordsResponse xmlns:csw="http://www.opengis.net/cat/csw/3.0"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:rim="urn:oasis:names:tc:ebxml-regrep:xsd:rim:3.0"
xmlns:atom="http://www.w3.org/2005/Atom"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:dct="http://purl.org/dc/terms/"
xmlns:ows="http://www.opengis.net/ows"
xmlns:apiso="http://www.opengis.net/cat/csw/apiso/1.0"
xmlns:gml="http://www.opengis.net/gml"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:gco="http://www.isotc211.org/2005/gco"
xmlns:gmd="http://www.isotc211.org/2005/gmd"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:srv="http://www.isotc211.org/2005/srv"
xmlns:ogc="http://www.opengis.net/ogc"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:os="http://a9.com/-/spec/opensearch/1.1/"
xmlns:sitemap="http://www.sitemaps.org/schemas/sitemap/0.9"
xmlns:wrs="http://www.opengis.net/cat/wrs/1.0" version="2.0.2"
xsi:schemaLocation="http://www.opengis.net/cat/csw/3.0
../../../csw/3.0/CSW30-discovery.xsd">
  <csw:SearchStatus timestamp="2015-05-09T04:10:02Z"/>
  <csw:SearchResults numberOfRecordsReturned="1" nextRecord="0"
numberOfRecordsMatched="1"
recordSchema="http://www.isotc211.org/2005/gmd"
elementSet="brief">
    <gmd:MD_Metadata
xsi:schemaLocation="http://www.isotc211.org/2005/gmd
http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd">
```

```
      <gmd:fileIdentifier>
        <gco:CharacterString>22610244-2374-4447-9705-
465268442787</gco:CharacterString>
      </gmd:fileIdentifier>
      <gmd:hierarchyLevel>
        <gmd:MD_ScopeCode
codeList="http://.../gmxCodelists.xml#MD_ScopeCode"
codeSpace="ISOTC211/19115"
codeListValue="dataset">dataset</gmd:MD_ScopeCode>
      </gmd:hierarchyLevel>
      <gmd:identificationInfo>
        <gmd:MD_DataIdentification id="22610244-2374-4447-9705-
465268442787">
          <gmd:citation>
            <gmd:CI_Citation>
              <gmd:title>
<gco:CharacterString>NHDPlus Stream Gages
</gco:CharacterString>
              </gmd:title>
              <gmd:date>
                <gmd:CI_Date>
                  <gmd:date>
                    <gco:Date>2015-04-12</gco:Date>
                  </gmd:date>
                  <gmd:dateType>
                    <gmd:CI_DateTypeCode
codeList="http://../gmxCodelists.xml#CI_DateTypeCode"
codeSpace="ISOTC211/19115"
codeListValue="creation">creation</gmd:CI_DateTypeCode>
                  </gmd:dateType>
                </gmd:CI_Date>
              </gmd:date>
            </gmd:CI_Citation>
          </gmd:citation>
          <gmd:extent>
            <gmd:EX_Extent>
              <gmd:geographicElement>
                <gmd:EX_GeographicBoundingBox>
                  <gmd:westBoundLongitude>
                    <gco:Decimal>158.22</gco:Decimal>
                  </gmd:westBoundLongitude>
                  <gmd:eastBoundLongitude>
                    <gco:Decimal>158.22</gco:Decimal>
                  </gmd:eastBoundLongitude>
                  <gmd:southBoundLatitude>
                    <gco:Decimal>6.96</gco:Decimal>
                  </gmd:southBoundLatitude>
                  <gmd:northBoundLatitude>
                    <gco:Decimal>6.96</gco:Decimal>
                  </gmd:northBoundLatitude>
                </gmd:EX_GeographicBoundingBox>
              </gmd:geographicElement>
            </gmd:EX_Extent>
          </gmd:extent>
        </gmd:MD_DataIdentification>
```


Copyright © 2015 Open Geospatial Consortium.

```
      </gmd:identificationInfo>
    </gmd:MD_Metadata>
  </csw:SearchResults>
</csw:GetRecordsResponse>
```

The example response in the listing shows the returned ISO 19115 metadata encoded in XML based on ISO 19139 — arguably the most widely implemented geospatial metadata specification. The increasing popularity of JSON, however, has raised the need to explore the possibility of encoding metadata based on ISO 19115 but serialized in JSON. The testbed configured the deployed CSW 3.0 to also offer JSON encoded metadata through an OpenSearch interface.. An example query and response are shown below:

http://www.exampleserver.com/cat3/opensearch?service=CSW&version=3.0.0&maxRecords=10&q=ortho&startPosition=1&bbox=-180,-90,180,90&time=2000-01-01T00:00:00Z/2014-12-31T23:59:59Z&outputFormat=application/json


```
{
    "attributes": {
        "version": "2.0.2",
        "xsi:schemaLocation": "http://www.opengis.net/cat/csw/2.0.2
http://schemas.opengis.net/csw/2.0.2/CSW-discovery.xsd"
    },
    "tag": "csw:GetRecordsResponse",
    "children": [
        {
            "attributes": {
                "timestamp": "2015-05-09T04:23:45Z"
            },
            "tag": "csw:SearchStatus"
        },
        {
            "attributes": {
                "numberOfRecordsMatched": "2",
                "nextRecord": "0",
                "numberOfRecordsReturned": "2",
                "elementSet": "full",
                "recordSchema": "http://www.opengis.net/cat/csw/2.0.2"
            },
            "tag": "csw:SearchResults",
            "children": [
                {
                    "tag": "csw:Record",
                    "children": [
                        {
                            "text": "5f37e0f8-4fb1-4637-b959-b415058bdb68",
                            "tag": "dc:identifier"
                        },
                        {
                            "text": "Ortho",
```


Copyright © 2015 Open Geospatial Consortium.

```
                    "tag": "dc:title"
                },
                {
                    "text": "dataset",
                    "tag": "dc:type"
                },
                {
                    "text": "Orthoimagery",
                    "tag": "dc:subject"
                },
                {
                    "text": "http://www.ypaat.gr",
                    "tag": "dct:references",
                    "attributes": {
                        "scheme": "None"
                    }
                },
                {
                    "text": "2009-10-07",
                    "tag": "dct:modified"
                },
                {
                    "text": "Ortho",
                    "tag": "dct:abstract"
                },
                {
                    "text": "2009-10-07",
                    "tag": "dc:date"
                },
                {
                    "text": "otherRestrictions",
                    "tag": "dc:rights"
                },
                {
                    "attributes": {
                        "crs": "urn:x-ogc:def:crs:EPSG:6.11:4326",
                        "dimensions": "2"
                    },
                    "tag": "ows:BoundingBox",
                    "children": [
                        {
                            "text": "39.71 21.53",
                            "tag": "ows:LowerCorner"
                        },
                        {
                            "text": "39.74 21.58",
                            "tag": "ows:UpperCorner"
                        }
                    ]
                }
            ]
        },
        {
```


 Copyright © 2015 Open Geospatial Consortium.

```
                    "tag": "csw:Record",
                    "children": [
                        {
                            "text": "NS06agg",
                            "tag": "dc:identifier"
                        },
                        {
                            "text": "PacIOOS Nearshore Sensor 06: Micronesia",
                            "tag": "dc:title"
                        },
                        {
                            "text": "dataset",
                            "tag": "dc:type"
                        },
                        {
                            "text": "Oceans > Ocean Chemistry > Chlorophyll",
                            "tag": "dc:subject"
                        },
                        {
                            "text": "Oceans > Ocean Optics > Turbidity",
                            "tag": "dc:subject"
                        },
                        {
                            "text": "Oceans > Ocean Temperature > Water
Temperature",
                            "tag": "dc:subject"
                        },
                        {
                            "text": "Oceans > Salinity/Density >
Conductivity",
                            "tag": "dc:subject"
                        },
                        {
                            "text": "Oceans > Salinity/Density > Salinity",
                            "tag": "dc:subject"
                        },
                        {
                            "text": "Oceans > Water Quality",
                            "tag": "dc:subject"
                        },
                        {
                            "text": "sea_water_salinity",
                            "tag": "dc:subject"
                        },
                        {
                            "text": "depth",
                            "tag": "dc:subject"
                        },
                        {
                            "text": "latitude",
                            "tag": "dc:subject"
                        },
                        {
                            "text": "longitude",
```


Copyright © 2015 Open Geospatial Consortium.

```
                            "tag": "dc:subject"
                        },
                        {
                            "text": "time",
                            "tag": "dc:subject"
                        },
                        {
                            "text": "http://
oos.soest.hawaii.edu/thredds/nss.html",
                            "tag": "dct:references",
                            "attributes": {
                                "scheme": "None"
                            }
                        },
                        {
                            "tag": "dc:relation"
                        },
                        {
                            "text": "2014-04-16",
                            "tag": "dct:modified"
                        },
                        {
                            "text": "The nearshore sensors are part of the
PacIOOS.",
                            "tag": "dct:abstract"
                        },
                        {
                            "text": "2014-04-16",
                            "tag": "dc:date"
                        },
                        {
                            "attributes": {
                                "crs": "urn:x-ogc:def:crs:EPSG:6.11:4326",
                                "dimensions": "2"
                            },
                            "tag": "ows:BoundingBox",
                            "children": [
                                {
                                    "text": "6.96 158.22",
                                    "tag": "ows:LowerCorner"
                                },
                                {
                                    "text": "6.96 158.22",
                                    "tag": "ows:UpperCorner"
                                }
                            ]
                        }
                    ]
                }
            ]
        }
    ]
```


 Copyright © 2015 Open Geospatial Consortium.

```
}
```

10.3.2 RDF-Generating WPS

The testbed explored different approaches for generating Linked Data. One approach used a Web Processing Service (WPS) to convert NHD data into RDF. The WPS approach has the benefit of web access through its provision of a service endpoint for client applications to remotely bind to. Another approach, discussed in Section 10.3.3, used an Extraction Transform Load (ETL) tool to convert NHD data into RDF. The OGC WPS Interface Standard provides a standard interface for publishing simple or complex computational processes via web services. WPS is designed to be location aware through its support for GML-encoded inputs. WPS is also designed to be self-descriptive through its offering of process descriptions and capabilities metadata. In order to produce RDF-encoded data from both the TNM and NHD datasets, a WPS was provided with the following process description:

```
<wps:ProcessDescriptions xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:wps="http://www.opengis.net/wps/1.0.0"
xmlns:ows="http://www.opengis.net/ows/1.1"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xml:lang="en"
service="WPS" version="1.0.0"
xsi:schemaLocation="http://www.opengis.net/wps/1.0.0
http://schemas.opengis.net/wps/1.0.0/wpsAll.xsd">
   <ProcessDescription wps:processVersion="1.0.0"
statusSupported="true" storeSupported="true">
      <ows:Identifier>gs:Tnm2Rdf</ows:Identifier>
      <ows:Title>TNM to RDF</ows:Title>
      <ows:Abstract>Get TNM with RDF format</ows:Abstract>
      <DataInputs>
         <Input maxOccurs="1" minOccurs="1">
            <ows:Identifier>bbox</ows:Identifier>
            <ows:Title>bbox</ows:Title>
            <ows:Abstract>bounding box</ows:Abstract>
            <LiteralData>
               <ows:AnyValue/>
            </LiteralData>
         </Input>
         <Input maxOccurs="1" minOccurs="1">
            <ows:Identifier>dataName</ows:Identifier>
            <ows:Title>dataName</ows:Title>
            <ows:Abstract>data name of tnm</ows:Abstract>
            <LiteralData>
               <ows:AllowedValues>
                  <ows:Value>Flowline</ows:Value>
               </ows:AllowedValues>
            </LiteralData>
         </Input>
      </DataInputs>
```

Copyright © 2015 Open Geospatial Consortium.

```
      <ProcessOutputs>
         <Output>
            <ows:Identifier>GetByBBox</ows:Identifier>
            <ows:Title>GetByBBox</ows:Title>
            <LiteralOutput/>
         </Output>
      </ProcessOutputs>
   </ProcessDescription>
</wps:ProcessDescriptions>
```

An example WPS request to trigger the generation of RDF from TNM data is shown in the following listing.

```
<?xml version="1.0" encoding="UTF-8"?>
<wps:Execute version="1.0.0" service="WPS"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="http://www.opengis.net/wps/1.0.0"
xmlns:wfs="http://www.opengis.net/wfs"
xmlns:wps="http://www.opengis.net/wps/1.0.0"
xmlns:ows="http://www.opengis.net/ows/1.1"
xmlns:gml="http://www.opengis.net/gml"
xmlns:ogc="http://www.opengis.net/ogc"
xmlns:wcs="http://www.opengis.net/wcs/1.1.1"
xmlns:xlink="http://www.w3.org/1999/xlink"
xsi:schemaLocation="http://www.opengis.net/wps/1.0.0
http://schemas.opengis.net/wps/1.0.0/wpsAll.xsd">
  <ows:Identifier>gs:Tnm2Rdf</ows:Identifier>
  <wps:DataInputs>
    <wps:Input>
      <ows:Identifier>BBOX</ows:Identifier>
       <wps:Data>
          <wps:ComplexData mimeType="text/xml">
             <gml:Box>
          <gml:coordinates>-12909854.299551187,-12909832.258292047
4380430.9646445736,4380281.8091793861</gml:coordinates>
             </gml:Box>
          </wps:ComplexData>
       </wps:Data>
    </wps:Input>
    <wps:Input>
      <ows:Identifier>DataName</ows:Identifier>
      <wps:Data>
        <wps:LiteralData>Flowline</wps:LiteralData>
      </wps:Data>
    </wps:Input>
  </wps:DataInputs>
  <wps:ResponseForm>
    <wps:RawDataOutput>
      <ows:Identifier>GetByBBox</ows:Identifier>
    </wps:RawDataOutput>
  </wps:ResponseForm>
```

 Copyright © 2015 Open Geospatial Consortium.

```
</wps:Execute>
```

10.3.3 Generating RDF with ETL tools

An alternative approach for generating RDF encoded data was developed using Safe Software's FME software, an ETL tool. Use of the ETL tool to generate RDF has the benefit of flexibility as users can modify the workbench to customize the generated RDF.

An example of the RDF Turtle data generated is shown below:

```
@prefix : <http://ows.usersmarts.com/nhd/flowline#> .
@prefix geosparql: <http://www.opengis.net/ont/geosparql#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix feature: <http://www.opengis.net/ont/feature#> .
@prefix nhd: <http://www.opengis.net/ont/testbed11/hydro/nhd#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .

:20245062 a nhd:Flowline , nhd:HydroFeature ;
 dct:created "1999-11-24"^^xsd:date ;
 feature:featureType
<http://www.opengis.net/taxonomies/testbed11/hydro/nhd#StreamRiver> ;
 nhd:comID "20245062"^^xsd:int ;
 nhd:geometry [ a geosparql:Geometry ;
 geosparql:asWKT "MULTILINESTRING ((-115.97119433042951
36.53619487662189, -115.95031919712858 36.5737368765636, -
115.9506327971281 36.575156276561415, -115.95094619712762
36.576323809892926, -115.95097499712756 36.576804476558834, -
115.95123159712716 36.5778804765572))"^^geosparql:wktLiteral
 ] ;
 nhd:hasReach <http://ows.usersmarts.com/nhd/reach/16060014044574> ;
 nhd:lengthInKM "5.415"^^xsd:double ;
 nhd:resolution nhd:Medium .

<http://ows.usersmarts.com/nhd/reach/16060014044574>
 a nhd:Reach ;
 nhd:reachCode "16060014044574" ;
 nhd:reachOf :20245062 .

<http://www.opengis.net/taxonomies/testbed11/hydro/nhd#StreamRiver>
 a feature:FeatureType ;
 rdfs:label "StreamRiver" ;
 skos:notation "46006"^^xsd:int .


nhd:Medium a nhd:Resolution .


:24085230 a nhd:Flowline , nhd:HydroFeature ;
 dct:created "2001-02-27"^^xsd:date ;
 feature:featureType
```

```
<http://www.opengis.net/taxonomies/testbed11/hydro/nhd#StreamRiver> ;
 nhd:comID "24085230"^^xsd:int ;
 nhd:geometry [ a geosparql:Geometry ;
 geosparql:asWKT "MULTILINESTRING ((-121.15726798904637
42.88877720009475, -121.17686218901594 42.88773180009639, -
121.17831932234702 42.888627666761636, -121.17885338901289
42.8892846667606, -121.17894538901271 42.89022266675914, -
121.17866238901314 42.89052366675867))"^^geosparql:wktLiteral
 ] ;
 nhd:hasReach <http://ows.usersmarts.com/nhd/reach/17120005008721> ;
 nhd:lengthInKM "2.787"^^xsd:double ;
 nhd:resolution nhd:Medium .


<http://ows.usersmarts.com/nhd/reach/17120005008721>
 a nhd:Reach ;
 nhd:reachCode "17120005008721" ;
 nhd:reachOf :24085230 .
```

10.3.4 WFS-G Semantic Mediator

A semantic mediator with support for WFS-G was deployed in the testbed to provide mediation capabilities between WFS provided by USGS and other services. The WFS-G semantic mediator is designed to enable heterogeneous gazetteers offered through WFS to be accessed from a single point of entry and using a common language (based on the ISO 19112 standard for spatial referencing by identifiers). The WFS-G semantic mediator was connected to the USGS Geonames WFS. As the latter service (USGS Geonames WGS) was not based on the WFS-G specification, the former service (the mediator) was configured to translate the properties specified in filter constraints from ISO 19112 to the schema supported by the latter service.

The WFS-G semantic mediator was configured to retrieve semantic mappings from a GeoSPARQL Server and use the semantic mappings to translate place types from one vocabulary to another (e.g. NGA to USGS gazetteer place types). An example of a WFS-G request is shown below:

```
<?xml version="1.0" encoding="UTF-8"?>
<GetFeature
    xmlns="http://www.opengis.net/wfs"
    xmlns:fes="http://www.opengis.net/fes/2.0"
xmlns:iso19112="http://www.isotc211.org/19112"
    xmlns:ogc="http://www.opengis.net/ogc"
    xmlns:xlink="http://www.w3.org/1999/xlink"
    xmlns:xsi="http://www/w3/org/2001/XMLSchema-instance"
    xmlns:gml="http://www.opengis.net/gml"
    service="WFS"
    version="1.1.0"
    outputFormat="text/xml; subtype=gml/3.1.1"
    maxFeatures="10"
    handle="">
```

 Copyright © 2015 Open Geospatial Consortium.

```
    <Query typeName="iso19112:SI_LocationInstance"
           srsName="urn:ogc:def:crs:EPSG::4326">
       <ogc:Filter>
       <ogc:And>
<ogc:PropertyIsSemanticallyRelatedTo>
<ogc:PropertyName>iso19112:locationType/@xlink:title</ogc:PropertyName>

<ogc:Literal>water tank</ogc:Literal>
</ogc:PropertyIsSemanticallyRelatedTo>
             <ogc:BBOX>
                <ogc:PropertyName>iso19112:position</ogc:PropertyName>
                <gml:Envelope srsName="urn:ogc:def:crs:EPSG::4326">
         <gml:lowerCorner>43 -91</gml:lowerCorner>
         <gml:upperCorner>47 -87</gml:upperCorner>
                </gml:Envelope>
             </ogc:BBOX>
       </ogc:And>
       </ogc:Filter>
    </Query>
</GetFeature>
```

The following is part of the response to the previous request:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<ns5:FeatureCollection xmlns="http://www.opengis.net/gml"
xmlns:ns2="http://www.isotc211.org/19112"
xmlns:ns3="http://www.isotc211.org/2005/gmdsf1"
xmlns:ns4="http://www.w3.org/1999/xlink"
xmlns:ns5="http://www.opengis.net/wfs"
timeStamp="2015-04-15T15:39:55.409+01:00" numberOfFeatures="3">
    <boundedBy/>
    <featureMember>
        <ns2:SI_LocationInstance>
            <ns2:guid>ENV.1429108795409.0</ns2:guid>

<ns2:geographicIdentifier>1958165</ns2:geographicIdentifier>
            <ns2:alternativeGeographicIdentifiers>
                <ns2:alternativeGeographicIdentifier>
                    <ns2:name>Lake Labelle (historical)</ns2:name>
                </ns2:alternativeGeographicIdentifier>
            </ns2:alternativeGeographicIdentifiers>
            <ns2:position>
                <Point srsName="urn:ogc:def:crs:EPSG::4326">
                    <pos>43.20554490000006 -90.2354016999999</pos>
                </Point>
            </ns2:position>
            <ns2:geographicExtent>
                <ns3:EX_GeographicExtent/>
            </ns2:geographicExtent>
            <ns2:spatialObject>http://1-dot-
env072015.appspot.com/query?
query=SELECT%09%3Fsubject+%3Fobject+where+%7B%3Fsubject+%3Chttp%3A......
geosparql%23wktLiteral%3E%2C%3Fobject%29%7D&amp;output=json
</ns2:spatialObject>
```

Copyright © 2015 Open Geospatial Consortium.

```
            <ns2:locationType ns4:href="http://someURL/Reservoir"
ns4:title="Reservoir"/>
        </ns2:SI_LocationInstance>
    </featureMember>
    <featureMember>
        <ns2:SI_LocationInstance>
            <ns2:guid>ENV.1429108795409.1</ns2:guid>

<ns2:geographicIdentifier>1569493</ns2:geographicIdentifier>
            <ns2:alternativeGeographicIdentifiers>
                <ns2:alternativeGeographicIdentifier>
                    <ns2:name>Mill Pond (historical)</ns2:name>
                </ns2:alternativeGeographicIdentifier>
            </ns2:alternativeGeographicIdentifiers>
            <ns2:position>
                <Point srsName="urn:ogc:def:crs:EPSG::4326">
                    <pos>43.68736164300009 -89.04159427199994</pos>
                </Point>
            </ns2:position>
            <ns2:geographicExtent>
                <ns3:EX_GeographicExtent/>
            </ns2:geographicExtent>
            <ns2:spatialObject>http://1-dot-
env072015.appspot.com/query?
query=SELECT%09%3Fsubject+%3Fobject+where+%7B%3Fsubject+%3Chttp%3A%2F%2
F
89.04159427199994+43.68736164300009+0.8%29%22%5E%5E%3Chttp%3A%2F%2Fwww.
object%29%7D&amp;output=json</ns2:spatialObject>
            <ns2:locationType ns4:href="http://someURL/Reservoir"
ns4:title="Reservoir"/>
        </ns2:SI_LocationInstance>
    </featureMember>
</ns5:FeatureCollection>
```

10.3.5 GeoSPARQL Servers

The testbed participants deployed GeoSPARQL Servers to provide the ability to query the RDF-encoded data within the testbed. An example GeoSPARQL query is shown below. The query shows a FILTER constraint inside a WHERE statement that uses a GeoSPARQL operator to limit query results to only those within the spatial extent of the specified polygon. The server allows a client to specify whether the response should be encoded in any of a number of languages, including RDF/XML, JSON and JSON-LD:

```
SELECT ?subject ?label ?object WHERE {
?subject <http://www.w3.org/2000/01/rdf-schema#label> ?label .
?subject <http://www.opengis.net/ont/geosparql#asWKT> ?object
FILTER <http://www.opengis.net/def/function/geosparql/sfWithin>
("POLYGON((-92.775 46.546, -92.775 47.546,-91.723 47.025, -91.723
46.025,-92.775
46.546))"^^<http://www.opengis.net/ont/geosparql#wktLiteral>,?object)
}
```


Copyright © 2015 Open Geospatial Consortium.

Part of an example response from a SPARQL SELECT query is shown below:


```
{
  "head": {
    "vars": [ "subject" , "label" , "object" ]
  } ,
  "results": {
    "bindings": [
      {
        "subject": { "type": "uri" , "value": "http://1-dot-
env072015.appspot.com/resource/nhd/streamgages/gageloc_11316000" } ,
        "label": { "type": "literal" , "value": "BEAR R NR SALT SPRINGS DAM
CA" } ,
        "object": { "datatype":
"http://www.opengis.net/ont/geosparql#wktLiteral" , "type": "typed-literal" ,
"value": "POINT(-120.28919999999999391 38.49340000000000117)" }
      } ,
      {
        "subject": { "type": "uri" , "value": "http://1-dot-
env072015.appspot.com/resource/nhd/streamgages/gageloc_11425410" } ,
        "label": { "type": "literal" , "value": "ROCK CREEK LK NR AUBURN CA" }
,
        "object": { "datatype":
"http://www.opengis.net/ont/geosparql#wktLiteral" , "type": "typed-literal" ,
"value": "POINT(-121.08490000000000464 38.91199999999999903)" }
      } ,
      {
        "subject": { "type": "uri" , "value": "http://1-dot-
env072015.appspot.com/resource/nhd/streamgages/gageloc_11299995" } ,
        "label": { "type": "literal" , "value": "TULLOCH RES NR KNIGHTS FERRY
CA" } ,
        "object": { "datatype":
"http://www.opengis.net/ont/geosparql#wktLiteral" , "type": "typed-literal" ,
"value": "POINT(-120.60460000000000491 37.87630000000000052)" }
      }
    ]
  }
}
```


Part of the response to the above query is shown below, encoded in JSON-LD:


```
{
  "@id" : "nhd-gage:gageloc_11298000",
  "@type" : [ "nhd:StreamGage", "nhd:HydroFeature" ],
  "asWKT" : "POINT(-120.16880000000001871 38.09259999999999735)",
  "active" : "1",
  "agency_cd" : "USGS",
  "comid" : "0",
  "dasqkm" : "66.9",
  "eventdate" : "2014-12-30T00:00:00",
  "eventtype" : "StreamGage",
```

```
  "featurecla" : "0",
  "featurecom" : "0",
  "featuredet" : "http://waterdata.usgs.gov/nwis/nwisman/?site_no=11298000",
  "flcomid" : "343847",
  "gagesii" : "Non-ref",
  "hasReach" : "reach/18040010000044",
  "latsite" : "38.09242097",
  "lonsite" : "-120.1687993",
  "measure" : "88.95987",
  "offset" : "0.0",
  "reachcode" : "18040010000044",
  "reachresol" : "Medium",
  "reachsmdat" : "NULL",
  "source_dat" : " ",
  "source_fea" : "11298000",
  "source_ori" : "USGS, Water Resources Division",
  "state" : "CA",
  "state_cd" : "6",
  "station_nm" : "SF STANISLAUS R NR LONG BARN CA",
  "label" : "SF STANISLAUS R NR LONG BARN CA",
  "seeAlso" : [ "http://live.dbpedia.org/resource/Stream_gauge",
"http://cegis.usgs.gov/NHDOntology/GagingStation" ],
  "@context" : {
    "agency_cd" : "http://www.opengis.net/ont/testbed11/hydro/nhd#agency_cd",
    "featurecla" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#featurecla",
    "hasReach" : {
      "@id" : "http://www.opengis.net/ont/testbed11/hydro/nhd#hasReach",
      "@type" : "@id"
    },
    "reachcode" : "http://www.opengis.net/ont/testbed11/hydro/nhd#reachcode",
    "lonsite" : "http://www.opengis.net/ont/testbed11/hydro/nhd#lonsite",
    "flcomid" : "http://www.opengis.net/ont/testbed11/hydro/nhd#flcomid",
    "active" : "http://www.opengis.net/ont/testbed11/hydro/nhd#active",
    "comid" : "http://www.opengis.net/ont/testbed11/hydro/nhd#comid",
    "asWKT" : {
      "@id" : "http://www.opengis.net/ont/geosparql#asWKT",
      "@type" : "http://www.opengis.net/ont/geosparql#wktLiteral"
    },
    "featurecom" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#featurecom",
    "offset" : "http://www.opengis.net/ont/testbed11/hydro/nhd#offset",
    "state" : "http://www.opengis.net/ont/testbed11/hydro/nhd#state",
    "dasqkm" : "http://www.opengis.net/ont/testbed11/hydro/nhd#dasqkm",
    "featuredet" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#featuredet",
    "eventdate" : "http://www.opengis.net/ont/testbed11/hydro/nhd#eventdate",
    "seeAlso" : {
      "@id" : "http://www.w3.org/2000/01/rdf-schema#seeAlso",
      "@type" : "@id"
    },
    "latsite" : "http://www.opengis.net/ont/testbed11/hydro/nhd#latsite",
    "state_cd" : "http://www.opengis.net/ont/testbed11/hydro/nhd#state_cd",
```


 Copyright © 2015 Open Geospatial Consortium.

```
    "gagesii" : "http://www.opengis.net/ont/testbed11/hydro/nhd#gagesii",
    "station_nm" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#station_nm",
    "source_dat" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#source_dat",
    "measure" : "http://www.opengis.net/ont/testbed11/hydro/nhd#measure",
    "reachsmdat" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#reachsmdat",
    "label" : "http://www.w3.org/2000/01/rdf-schema#label",
    "source_fea" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#source_fea",
    "reachresol" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#reachresol",
    "eventtype" : "http://www.opengis.net/ont/testbed11/hydro/nhd#eventtype",
    "source_ori" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#source_ori",
    "@base" : "http://ows.usersmarts.com/nhd/flowline#",
    "" : "http://ows.usersmarts.com/nhd/flowline#",
    "nhd-gage" : "http://1-dot-
env072015.appspot.com/resource/nhd/streamgages/",
    "geo" : "http://www.opengis.net/ont/geosparql#",
    "foaf" : "http://xmlns.com/foaf/0.1/",
    "symbol" : "http://www.opengis.net/ont/portrayal/symbol#",
    "community" : "http://www.opengis.net/ont/community#",
    "j.1" : "http://www.opengis.net/ont/portrayal/symbol#",
    "j.0" : "http://purl.org/dc/terms/",
    "cegis" : "http://cegis.usgs.gov/surfacewater/GIS-NHD/",
    "rdfs" : "http://www.w3.org/2000/01/rdf-schema#",
    "geosparql" : "http://www.opengis.net/ont/geosparql#",
    "nhd-catch" : "http://1-dot-
env072015.appspot.com/resource/nhd/catchments/",
    "dct" : "http://purl.org/dc/terms/",
    "owl" : "http://www.w3.org/2002/07/owl#",
    "xsd" : "http://www.w3.org/2001/XMLSchema#",
    "rdf" : "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
    "feature" : "http://www.opengis.net/ont/feature#",
    "nhd" : "http://www.opengis.net/ont/testbed11/hydro/nhd#",
    "skos" : "http://www.w3.org/2004/02/skos/core#",
    "dbpedia-ont" : "http://live.dbpedia.org/ontology/"
  }
}
```


The following is an example of a catchment exported in JSON-LD. The approach adopted allowed for StreamGage instances to be referenced from a hasStreamGage predicate, the value of which was dereferencable:


```
{
  "@id" : "http://1-dot-
env072015.appspot.com/resource/nhd/catchments/catchment_2853539",
  "@type" : [ "nhd:Catchment", "nhd:HydroFeature" ],
  "asWKT" : "MULTIPOLYGON(((-122.11754157499997575 38.78749377707858059,-
122.11763618499996653 38.78775138507855047,-122.11797172699994007
```

```
38.78767874907858726,-122.11687049100000024 38.78763904607855295,-
122.11754157499997575 38.78749377707858059)))",
  "areasqkm" : "3.4479",
  "featureid" : "2853539",
  "gml_id" : "catchment.2538261",
  "gridcode" : "1410574",
  "hasGridCode" : "http://ows.usersmarts.com/nhd/grid/",
  "shape_area" : "0.000357519227295",
  "shape_length" : "0.109974897791",
  "sourcefc" : "NHDFlowline",
  "label" : "1410574",
  "@context" : {
    "gml_id" : "http://www.opengis.net/ont/testbed11/hydro/nhd#gml_id",
    "shape_length" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#shape_length",
    "shape_area" :
"http://www.opengis.net/ont/testbed11/hydro/nhd#shape_area",
    "label" : "http://www.w3.org/2000/01/rdf-schema#label",
    "asWKT" : {
      "@id" : "http://www.opengis.net/ont/geosparql#asWKT",
      "@type" : "http://www.opengis.net/ont/geosparql#wktLiteral"
    },
    "sourcefc" : "http://www.opengis.net/ont/testbed11/hydro/nhd#sourcefc",
    "gridcode" : "http://www.opengis.net/ont/testbed11/hydro/nhd#gridcode",
    "featureid" : "http://www.opengis.net/ont/testbed11/hydro/nhd#featureid",
    "areasqkm" : "http://www.opengis.net/ont/testbed11/hydro/nhd#areasqkm",
    "hasGridCode" : {
      "@id" : "http://www.opengis.net/ont/testbed11/hydro/nhd#hasGridCode",
      "@type" : "@id"
    },
    "rdfs" : "http://www.w3.org/2000/01/rdf-schema#",
    "geosparql" : "http://www.opengis.net/ont/geosparql#",
    "geo" : "http://www.opengis.net/ont/geosparql#",
    "foaf" : "http://xmlns.com/foaf/0.1/",
    "symbol" : "http://www.opengis.net/ont/portrayal/symbol#",
    "dct" : "http://purl.org/dc/terms/",
    "owl" : "http://www.w3.org/2002/07/owl#",
    "xsd" : "http://www.w3.org/2001/XMLSchema#",
    "community" : "http://www.opengis.net/ont/community#",
    "j.1" : "http://www.opengis.net/ont/portrayal/symbol#",
    "rdf" : "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
    "j.0" : "http://purl.org/dc/terms/",
    "nhd" : "http://www.opengis.net/ont/testbed11/hydro/nhd#",
    "skos" : "http://www.w3.org/2004/02/skos/core#"
  }
}
```


 Copyright © 2015 Open Geospatial Consortium.

10.3.6 Client Component

The various web services were connected to a client component that had been implemented as a web application. A screenshot of the client component is presented in Figure 5.

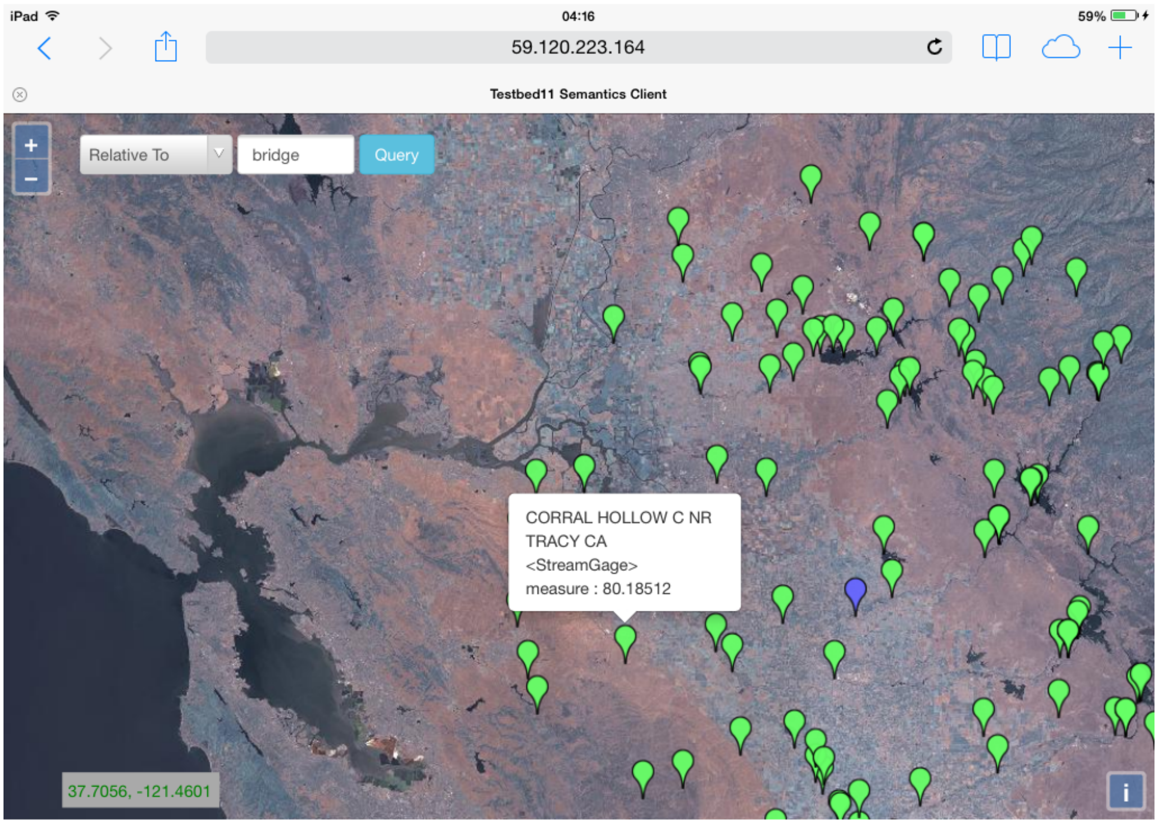

Figure 5. Screenshot of the client component

11 Discussion

A number of observations were made during the implementation of the testbed architecture.

The CSW was deployed in the testbed to provide a resource discovery component, as described in Section 10.3.1. Such a component is particularly relevant to the USGS because of the vast number and variety of web services that the agency publishes on the World Wide Web. Publishing metadata describing the location, content and lineage of these services is not only necessary for facilitating access by the general public, but is also necessary for enabling USGS personnel to manage the data that the separate groups within the organization publish. This engineering report recommends that the USGS

Copyright © 2015 Open Geospatial Consortium.

establishes a registry of all USGS services offering NHD data to support the management of unique identifiers within a future Linked Data framework.

The WPS takes as inputs the name of the feature type (e.g. Flowline) and the bounding rectangle (box) around the area of interest, as described in Section 10.3.2. Considering that various groups within the USGS have published datasets derived from NHD data, a centralized WPS configured to process all of the published datasets would be appropriate for establishing a Linked Data service for the USGS. Such a WPS would also need to maintain cross references between the various NHD features as was demonstrated in the testbed through referencing of different feature instances to NHD reaches through a reachcode.

The testbed observed that there are several instances of WFS offering NHD data. In some cases those WFS offer data for overlapping areas. Link Data offers the use of owl:sameAs to represent URI aliases that refer to resources that are similar, such as representations of the same stream gauge from different feature collections. URI aliases can be dereferenced to descriptions of the same resource thereby allowing for different expressions of views about the same resource to be represented. URI aliases also reduce the possibility of a single point of failure. In the case of the USGS, an approach could be for all Linked Data representations of the same NHD feature instances to cross reference one another.

Recommendation: That the USGS apply predicates such as owl:sameAs to identify equivalent NHD feature instances when published by different groups within the USGS.

To enable other data providers to link to NHD data, it is important to advertise NHDPlus Linked Data to the general public. There are a number of semantic web search engines available on the World Wide Web. Two examples of such engines are Sindice[10] and the Datahub[11]. Sindice is a platform to build applications on top of this data. Sindice collects Web Data in many ways, following existing web standards, and offers Search and Querying across this data, updated live every few minutes. The Datahub is a data management platform from the Open Knowledge Foundation, based on the CKAN tool which has been designed for managing and publishing collections of data. Registration on the Datahub allows a Linked Data collection to be included in the Linked Open Data cloud that is illustrated in Figure 1.

Recommendation: That, upon producing an NHD Linked Data product, the USGS should aim to advertise theavailability of the product on the aforementioned catalogues and other similar engines.

---

[10] http://sindice.com/

[11] http://datahub.io/

Copyright © 2015 Open Geospatial Consortium.

The testbed results clearly show that there are several benefits of using RDF to publish NHD and Gazetteer data as Linked Data. First, the use of HTTP URIs makes it possible to generate globally unique data. Second, the ability to dereference URIs makes it possible for client applications to discover and retrieve the resources described by those URIs. Third, the simplicity of the triple model makes it usable by any application capable of associating data from different sources. Fourth, all (or most) of the information associated with an entity can be combined by merging into a single graph thereby providing a record of historical and current knowledge about that entity. Fifth, the ability to specify new predicates or reuse existing ones, offers significantly more flexibility than alternative approaches such as application schemas described as XML schema definition (XSD). Sixth, the use of controlled vocabularies such as OWL and SKOS, allows some structure to be applied within RDF-encoded data when necessary.

12 Conclusions

This engineering report has provided guidelines on the publication of hydrographic and hydrological data using Linked Data principles applied to technologies based on OGC standards. Also presented are the findings and lessons learnt from the experimentation conducted by Testbed 11. The engineering report concludes that OGC web services, supported by GeoSPARQL Servers can indeed be used to generate and publish USGS NHD and gazetteer data as Linked Data. Further, the engineering report also concludes that existing NHD identifiers can be used to provide the cross referencing required to link NHD features to one another, and also to link to non-NHD data.

12.1 Recommendations

This report makes the following recommendations:

Recommendation 1 — Registry for managing unique identifiers for Linked Data: The USGS should establish a registry of all web feature services offering NHD data to support the management of unique identifiers within a future Linked Data framework.

Recommendation 2 — Application of URI aliases: The USGS should apply predicates such as owl:sameAs to identify similar NHD feature instances when published by different groups within the USGS.

Recommendation 3 — Advertise future NHD Linked Data Products in Semantic Web Search Engines: Upon producing an NHD Linked Data product, the USGS should aim to advertise the availability of the product on the aforementioned catalogues and other similar engines.

Recommendation 4 — Establish a standard for a GeoSPARQL Server: The OGC should establish a Standard Working Group to develop a specification for a GeoSPARQL Server.

Copyright © 2015 Open Geospatial Consortium.

12.2 Revision history

Date	Release	Editor	Primary clauses modified	Description
11-Aug-2015	Final	Carl Reed	Various	Prepare for publication.

Copyright © 2015 Open Geospatial Consortium.