Collaborative Project

GeoKnow - Making the Web an Exploratory for Geospatial Knowledge

Project Number: 318159 **Start Date of Project:** 2012/12/01 **Duration:** 36 months

Deliverable 7.3.1
First Standardization Report

Dissemination Level	Public
Due Date of Deliverable	Month 16, 31/03/2014
Actual Submission Date	Month 19, 19/06/2014
Work Package	WP7: Dissemination, Community Building, Exploitation & Standards
Task	T7.3
Type	Report
Approval Status	Final
Version	1.0
Number of Pages	8
Filename	D7.3.1_First_Standardisation_Report.pdf

Abstract: In this deliverable, we report on standardisation activities within the first half of the GeoKnow project, in particular participation in relevant groups and communities, organisation of events and dissemination and implementation of standards.

Project funded by the European Commission within the Seventh Framework Programme (2007 - 2013)

History

Version	Date	Reason	Revised by
0.1	20/10/2013	Collection of Standardisation Activities	Jens Lehmann
0.2	08/11/2013	More details on activities	Mohamed Sherif
0.3	10/01/2014	Extensions and Clean-Up	Jens Lehmann
0.4	09/03/2014	Ontos Standardisation Activities	Daniel Hladky
0.5	11/04/2014	OpenLink Standardisation Activities	Hugh Williams
0.9	30/05/2014	Extensions and Clean-Up	Jens Lehmann
1.0	05/06/2014	Final	Claus Stadler

Author List

Organization	Name	Contact Information
InfAI	Jens Lehmann	lehmann@informatik.uni-leipzig.de
InfAI	Mohamed Sherif	sherif@informatik.uni-leipzig.de
Ontos	Daniel Hladky	daniel.hladky@ontos.com
OpenLink	Hugh Williams	hwilliams@openlinksw.com
InfAI	Claus Stadler	cstadler@informatik.uni-leipzig.de

Executive Summary

In this deliverable, we report on standardisation activities within the first half of the GeoKnow project, in particular participation in relevant groups and communities, organisation of events, dissemination, and implementation of standards. Precisely, our new developments include RDF conversions for INSPIRE and EDI data, and a novel language for expressing RDB2RDF mappings, which has been shown to be easier to use than the current W3C standard R2RML. We engage in community work through our W3C community group, and event participation, such as the Linking Geospatial Data (LGD) workshop, where the OGC and W3C communities were brought together, and the European Data Forum. Furthermore, we contribute to the adoption of existing (partly recent) standards by providing implementations with increasing coverage.

Table of Contents

1 INSPIRE

The INSPIRE Directive[1] by the European Commission sets the legal and technical foundations toward interoperable Spatial Data Infrastructures (SDIs) across Europe. EU member states are already providing such services for several geospatial data themes (e.g., transportation networks, administrative units). Unfortunately, the INSPIRE ecosystem is currently disjoint from the Semantic Web, without any means to repurpose existing SDIs as high-quality data sources, and thus multiply their value through interlinking, reasoning and inferencing. In Task 2.7 of the project, GeoKnow contributes to supporting the INSPIRE Directive by providing means to convert INSPIRE data to RDF and/or access its services via SPARQL. To achieve this, XSLT transformations and the TripleGeo tool have been developed, which are both publicly available[2]. Furthermore, there has also been a meeting with the Joint Research Council in Italy as they are one of the main driving forces behind INSPIRE. Those community activities also ensure that the results in GeoKnow are widely disseminated and that semantic interoperability of INSPIRE data and meta data will be a target of future developments and standards.

2 Work on SML

The combination of the advantages of widely used relational databases and semantic technologies has attracted significant research over the past decade. In particular, mapping languages for the conversion of databases to RDF knowledge bases have been developed and standardized. *The Sparqlification Mapping Language* (SML) provides an intuitive way to declare those mappings based on SQL and SPARQL construct queries.

The study of SML[3] indicates that it is more intuitive and easier to write than the current W3C standard. Also, normally the process of SML mappings generation not only involves less errors than other mapping languages but also SML mapping is much easier for debugging. SML has been implemented in the Sparqlify tool[4], whose development advanced in the course of GeoKnow[5]. In parallel, Sparqlify's R2RML support is being improved for the purpose of providing better interoperability and thus contributing to the adoption of R2RML.

3 Geospatial Semantic Web Community Group

One initiative of the GeoKnow project was the creation of a community group in the W3C to target more potential users/partners interested in the Geospatioal Semantic Web domain. This group is available at http://www.w3.org/community/geosemweb/ since the beginning of the project. The mission of *Geospatial Semantic Web Community Group* is to make information seeking easier by allowing exploration, editing and interlinking of heterogeneous information sources with a spatial dimension. The scope of the work is to bring together scientists, GIS users, Linked Data users, data consumers and providers, interested in the exploitation of linked geospatial data.

Meanwhile the group consists of 55 participants who are interested in geospatial semantic data. At GeoKnow, we have used this group to share surveys that targets geospatial data consumers and providers,

[1] http://inspire.ec.europa.eu/
[2] https://github.com/GeoKnow/TripleGeo
[3] http://sml.aksw.org
[4] https://github.com/AKSW/Sparqlify
[5] http://svn.aksw.org/projects/GeoKnow/Public/D2.2.1_Integration_of_Geospatial_Databases.pdf

Title	Author	Date
W3C – Ontos Event about GeoKnow Generator	Daniel Hladky (Ontos)	April 25, 2014
GeoKnow Tutorials available	Daniel Hladky (Ontos)	March 20, 2014
GeoKnow Workshop at EDF 2014	Daniel Hladky (Ontos)	March 1, 2014
Geospatial Data User Survey Results	Alejandra Garcia Rojas (Ontos)	May 6, 2013
A Survey for Geospatial Data Users	Jon Jay Le Grange (Ontos)	March 13, 2013
Overview Article	Jens Lehmann	December 21, 2012
Geoknow Blog	Jon Jay Le Grange (Ontos)	December 21, 2012

Table 1: Posts in the Geospatial Semantic Web Community Group created by the GeoKnow team

and GIS users interested in having an integrated web of geospatial data. Also we have also shared some activities carried out in the consortium. Table 1 presents the list of posts created by the GeoKnow team.

Besides communicating the project activities, we have also contributed to the creation of a wiki[6] with relevant pointers in the Geospatial and Semantic Web domain, such as projects, vocabularies, and datasets.

4 RDF Conversion of EDI

Electronic data interchange (EDI) is a document standard, which when implemented acts as common interface between two or more computer applications in terms of understanding the document transmitted. It is commonly used by big companies for e-commerce purposes, such as sending orders to warehouses or tracking their order. It is more than mere e-mail; for instance, organizations might replace bills of lading and even cheques with appropriate EDI messages. It also refers specifically to a family of standards.

Conversion of EDI to RDF is done as part of WP5 by *BROX*. In this task, techniques for making the information from EDI transactions identifiable and accessible on the Web of Data. An identification scheme was devised for each of the relevant EDI message types, Also, the prevalent data models in the supply chain application is mapped to RDF.

5 Standardisation Activities at the Intersection of GIS and Linked Data

The Linked Geospatial Data workshop in London[7] brought together a broad range of geospatial interest groups from the Linked Data and geoinformatics worlds, with the goal of determining how the established world of Geoinformatics can be integrated with the Linked Data Web.

There is a clear need to bridge the chasm between the Linked Data and geoinformatics domains as both have much to offer for each other. Integration of Linked Data with geoinformatics will aid in the "desolification" of data warehouses, such as those built on SQL relational databases, by enabling them to be dereferencable via HTTP URIs and hence become part of the Linked Data Web. Linked

[6] http://www.w3.org/community/geosemweb/wiki/Main_Page
[7] http://www.w3.org/2014/03/lgd/

Data technologies have sufficiently matured to now provide the necessary performance and scalability for driving Linked GeoData solutions. There are a number of geospatial vocabularies, such as *WGS84 Geo Positioning*[8], *NeoGeo*[9], *Core location*[10], and *GeoSPARQL*[11], of which GeoSPARQL is becoming dominant in the Linked Data space. There was interest in GeoSPARQL supporting the *GeoJSON*[12] encoding which has been around for sometime now and already in use in many GIS systems. Futhermore, an extension of the *JSON-LD*[13] specification for the geo domain, referred to as GeoJSON-LD, was considered a worthwhile effort.

6 GeoSPARQL support

GeoSPARQL is a standard for representation and querying of geospatially linked data for the Semantic Web from the Open Geospatial Consortium (OGC). The definition of a small ontology based on well-understood OGC standards is intended. This will provide a standardized exchange basis for geospatial RDF data which can support both quantitative and qualitative spatial reasoning and querying with the SPARQL database query language.

Support for geospatial queries, both SQL and SPARQL, that were previously only available in the Virtuoso commercial product, has been made available in the open source version for use in the GeoKnow project. Initially, this extension only covered point geometry types. As of today, a vast set of additional features has been added, that are in line with the recommendation of the GeoKnow Market Research Overview deliverable[14] for the adoption of the OGC GeoSPARQL standard. Specifically, the following geometry types and their associated functions are now supported: points, linestrings, multilinestings, multi-polygons, boxes, polygons, multipolygons (polygon with holes), and geometry collections.

Enhancements will continue over the course of the project to achieve full GeoSPARQL compliance, which hopefully will eventually become a W3C standard.

The Virtuoso query optimiser has been enhanced to improve query execution speeds for geo spatial query types.

7 Linked Data Stack and Standardisation across EU Projects

During the European Data Forum 2014[15], one of the most important events in Europe about Open Data, the GeoKnow team presented the GeoKnow project at a booth and in the Linked Data Europe Workshop[16]. Tools and applications in the Linked Data Stack[17] as result of the GeoKnow project were shown to the audience. The talks focused on the GeoKnow research objectives related to geospatial linked open data, the GeoKnow Generator and various tools that help to fulfill the Linked Data Lifecycle with geospatial data.[18]

[8] http://www.w3.org/2003/01/geo/wgs84_pos
[9] http://geovocab.org/doc/neogeo.html
[10] https://joinup.ec.europa.eu/asset/core_location
[11] http://www.opengeospatial.org/standards/geosparql
[12] http://geojson.org/geojson-spec.html
[13] http://json-ld.org/spec/latest/
[14] http://svn.aksw.org/projects/GeoKnow/Public/D2.1.1_Market_and_Research_Overview.pdf
[15] http://2014.data-forum.eu
[16] http://www.linkeddataeurope.eu
[17] http://stack.linkeddata.org/
[18] A subset of the slides are available at http://www.linkeddataeurope.eu/wp-content/uploads/2014/01/GeoKnow_Short.pdf.

The GeoKnow consortium, in particular the partners InfAI, Athena and BROX were involved in the organisation of the European Data Forum and the Linked Data Europe Workshop. Both events fostered exchange between EU projects and the use of potential standard repositories like the Linked Data Stack.

8 W3C Members of the GeoKnow Consortium

8.1 Ontos

With the start of the GeoKnow project Ontos established the W3C Business and Community Group "Geospatial Semantic Web" under link http://www.w3.org/community/geosemweb/ . Besides regular posts Ontos was involved in establishing connections to the W3C activity group "Linking Geospatial Data" leading to the participation of GeoKnow team members at two workshops. Under the lead of the W3C and the project http://www.smartopendata.eu Ontos established connections to this project in order to exchange ideas on how to establish a joint working group at W3C for the topic of geospatial data. With the first release of the GeoKnow Generator, Ontos is organising in joint cooperation with W3C Switzerland an event in May 2014 where results of the GeoKnow project will be shown. In June 2014 Ontos will participate at the W3C world-wide meeting and at the face-to-face office meeting in order to establish other connections and investigate how to promote the project results into other countries using the W3C office network.

8.2 OpenLink

OpenLink Software as a W3C member is actively involved with the standardisation activities related to RDF and SPARQL. OpenLink regularly attends and contributes the SPARQL 1.1 working group ensuring Virtuoso implements all the SPARQL 1.1 feature as soon as they are ratified. In recent months OpenLink have been heavily involved in the W3C Linked Data Platform working group and already has an LDP Client implementation in Virtuoso based on the third working draft specification[19] and is currently nearing completion of a new LDP implementation based on the fourth work draft specification[20], and will be used as the RDF Linked Data API in the Fusepool P3 FP7 project[21] OpenLink is a partner in.

During the recent Linked Geospatial Data Workshop 2014 in London[22] it was proposed that a joint working group be setup between the OGC and W3C to create or recommend standards that work across those communities. To this end the W3C and OGC are in advanced and very positive discussions towards an MoU for operating the working group, a first draft of the charter of the joint working group is taking shape at http://www.w3.org/2014/05/geo-charter. Phil Archer (W3C) will be the driving force behind this "Geo Data on the Web" working group charter.

[19] http://www.w3.org/TR/2013/WD-ldp-20130730/
[20] http://www.w3.org/TR/2014/WD-ldp-20140311/
[21] http://www.fusepool.eu/p3
[22] http://www.w3.org/2014/03/lgd

Abbreviations and Acronyms

OGC	Open Geospatial Consortium
SQL	Structured Query Language
RDF	Resource Description Framework
SPARQL	SPARQL Protocol and RDF Query Language
SML	Work on Sparqlifcation Mapping Language
IDE	Electronic Data Interchange