



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 7.3.2 Second Standardization Report

| | |
|--------------------------------|--|
| Dissemination Level | Public |
| Due Date of Deliverable | Month 32, 31/07/2015 |
| Actual Submission Date | Month 33, 31/08/2015 |
| Work Package | WP7: Dissemination, Community Building, Exploitation & Standards |
| Task | T7.3 |
| Type | Report |
| Approval Status | Final |
| Version | 1.0 |
| Number of Pages | 9 |
| Filename | D7.3.2_Second_Standardisation_Report.pdf |

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

The information in this document reflects only the author's views and the European Community is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/ her sole risk and liability.



History

| Version | Date | Reason | Revised by |
|---------|------------|--|----------------------|
| 0.1 | 7/7/2015 | Collection of Standardisation Activities | Claus Stadler |
| 0.2 | 10/07/2015 | Ontos Standardisation Activities | Daniel Hladky |
| 0.3 | 04/08/2015 | Athena Standardisation Activities | Giorgos Giannopoulos |
| 0.4 | 16/08/2015 | OpenLink Standardisation Activities | Hugh Williams |
| 0.5 | 17/08/2015 | Consolidation of Contributions | Claus Stadler |
| 1.0 | 30/08/2015 | Finalization | Jens Lehmann |

Author List

| Organization | Name | Contact Information |
|--------------|----------------------|------------------------------------|
| InfAI | Claus Stadler | cstadler@informatik.uni-leipzig.de |
| InfAI | Jens Lehmann | lehmann@informatik.uni-leipzig.de |
| Ontos | Daniel Hladky | daniel.hladky@ontos.com |
| Athena | Giorgos Giannopoulos | giann@imis.athena-innovation.gr |
| OpenLink | Hugh Williams | hwilliams@openlinksw.com |

Executive Summary

This deliverable reports on standardization activities until near the end of the GeoKnow project. Many of the activities are continuations of those listed in the prior report (D7.3.1). As several of GeoKnow's tools are maturing, so is their standards compliance, such as Sparqlify's R2RML, and Virtuoso's GeoSPARQL support. In regard to standardization activities at the intersection of GIS and Linked Data, the Spatial Data on the Web Working Group was recently formed. This effort can be seen as a continuation of the efforts initiated with the Linked Geospatial Data (LGD) workshop in early 2014, to which GeoKnow contributed and continues to do so. An effort that benefits from the experiences and outcomes of GeoKnow is SHARE-PSI, a network for exchange of experience and ideas around implementing open data policies in the public sector.

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 Sparqlification Mapping Language |
| EDI | Electronic Data Interchange |

Table of Contents

| | | |
|----------|---|----------|
| 1 | INSPIRE | 5 |
| 2 | Work on RDB2RDF | 5 |
| 3 | Geospatial Semantic Web Community Group | 6 |
| 4 | RDF Conversion of EDI | 7 |
| 5 | Joint Standardisation Activities of GIS and Linked Data | 7 |
| 6 | GeoSPARQL support | 7 |
| 7 | Linked Data Stack and Standardisation across EU Projects | 8 |
| 8 | Open Government Data and Standardisation activities across EU Projects | 9 |

1 INSPIRE

The INSPIRE Directive by the European Commission¹ sets the legal and technical foundations toward interoperable Spatial Data Infrastructures (SDIs) across Europe. Until very recently, the INSPIRE ecosystem was 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 April 2015, the Joint Research Centre (JRC) of the European Commission has initiated a working group that will define an extension of the DCAT application profile for data portals in Europe (DCAT-AP)² for describing geospatial datasets, data series, and services. This extension (GeoDCAT-AP) will provide an RDF syntax binding for the union of metadata elements of the INSPIRE Metadata Regulation, the INSPIRE Regulation on interoperability of spatial datasets and services and the ISO 19115:2003 Core. Its basic use case is to make spatial datasets, data series, and services searchable on general data portals, thereby making geospatial information better searchable across borders and sectors. GeoDCAT-AP is being developed in the context of Action 1.1 – Improving semantic interoperability in European eGovernment systems of the European Commission's Interoperability Solutions for European Public Administrations (ISA) programme. GeoKnow has taken part in this working group, which has led to a proposed revision of the DCAT-AP that is currently under public review (until 30 August 2015). In another development, W3C³ and OGC⁴ have started in 2015 a joint collaboration with two working groups on standardization for Spatial Data on the Web⁵. Their mission is to determine integration of spatial information with other data on the Web; discovery by people and machines of different facts about 'places' at different levels of granularity; and assessing existing methods and tools to create a set of best practices. With respect to the GeoKnow project, T2.7 has contributed to INSPIRE by providing means to convert INSPIRE data and metadata to RDF, and also access its services via (Geo)SPARQL. With the support of the Joint Research Centre (JRC) of the European Commission, a methodology was developed towards repurposing INSPIRE SDIs as high-quality Linked Data sources according to the RDF and GeoSPARQL standards. Thanks to a suite of related open-source software tools (TripleGeo, TripleGeo-CSW) developed in GeoKnow, INSPIRE Catalogue Services, data, and metadata can be accessible with Semantic Web technologies, thus facilitating discovery, transformation, reusability, and exchange of INSPIRE features. These tools have been showcased with real-world examples in the workshop "INSPIRE & Linked Data: Bridging the gap"⁶, organized during the INSPIRE - World GeoSpatial Forum 2015 in Lisbon, Portugal (25-29 May 2015). This community meeting has also confirmed that several research projects (like SmartOpenData⁷, SDI4Apps⁸) and practitioners have tested or even made use of these GeoKnow tools for exposing INSPIRE data and metadata on the Semantic Web.

2 Work on RDB2RDF

In the prior report, we introduced our work on the the *Sparqlification Mapping Language* (SML)⁹ for mapping relational databases to RDF. The user study of SML, which indicated that it is more intuitive

¹<http://inspire.ec.europa.eu/>

²<https://joinup.ec.europa.eu/node/137964>

³<http://www.w3.org/>

⁴<http://www.opengeospatial.org/>

⁵<http://www.w3.org/2015/spatial/>

⁶http://www.geospatialworldforum.org/workshop.asp?Sp_Department=INSPIRE%20Linked%20Data

⁷<http://www.smartopendata.eu/>

⁸<http://sdi4apps.eu/>

⁹<http://sml.aksw.org>

and easier to write than the current W3C standard *R2RML*¹⁰, was recently published and presented at the Linked Data on the Web (LDOW) workshop¹¹. In addition, this work also comprised a tool for converting SML and R2RML mappings into each other¹². The tool covers most features of R2RML with the notable exception of the *rr:join* property, which is ongoing work. The Sparqlify tool¹³ was enhanced by integrating the conversion tool for featuring native R2RML support.

3 Geospatial Semantic Web Community Group

This W3C community group targeted at users and partners interested in the Geospatial Semantic Web domain¹⁴ was launched by GeoKnow at the beginning of the project. The mission of *Geospatial Semantic Web Community Group* is to make information seeking easier by allowing exploration, editing and inter-linking 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. Since the prior standardization report, the number of members in the group increased from 55 to 71. Table 1 lists all posts published by GeoKnow up to the time of writing this deliverable.

| Title | Author | Date |
|---|--------------------------------|-------------------|
| OSMRec – A tool for automatic annotation of spatial entities in OpenStreetMap | Giorgos Giannopoulos (Athena) | June 10, 2015 |
| Geospatial-semantic Exploration on the Move Linked Map | Uros Milosevic (IMP) | January 24, 2015 |
| W3C – Ontos Event about GeoKnow Generator | Francisco J. Lopez-Pellicer | August 24, 2014 |
| GeoKnow Tutorials available | Daniel Hladky (Ontos) | April 25, 2014 |
| GeoKnow Workshop at EDF 2014 | Daniel Hladky (Ontos) | March 20, 2014 |
| Geospatial Data User Survey Results | Daniel Hladky (Ontos) | March 1, 2014 |
| A Survey for Geospatial Data Users | Alejandra Garcia Rojas (Ontos) | May 6, 2013 |
| Overview Article | Jon Jay Le Grange (Ontos) | March 13, 2013 |
| Geoknow Blog | Jens Lehmann | December 21, 2012 |
| | Jon Jay Le Grange (Ontos) | December 21, 2012 |

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

¹⁰<http://www.w3.org/TR/r2rml/>

¹¹<http://events.linkeddata.org/ldow2015>

¹²<https://github.com/AKSW/sml-converters>

¹³<https://github.com/AKSW/Sparqlify>

¹⁴<http://www.w3.org/community/geosemweb/>

4 RDF Conversion of EDI

Electronic data interchange (EDI) is a family of standards, which when implemented act as common interface between two or more computer applications in terms of understanding the documents transmitted. Overall, progress on this topic is limited, as the main activities carried out as part of WP5 by *BROX* were mostly finished at the time of the first standardization report. This work was concerned the creation of an ontology for EDI messages together with a tool for RDFizing EDI data. Recently, the EDI event generator of the supply chain dashboard was turned into a stand-alone tool, which simplifies the generation of datasets corresponding to simulations of supply chain networks.

5 Joint Standardisation Activities of GIS and Linked Data

With Linked Data having gained significant importance over the past decade, the standards bodies W3C and OGC started exploring the mutual benefits of combining their technologies. During the Linked Geospatial Data Workshop 2014 in London¹⁵ 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. Thereby, ideas about *de-silo-fication* of data and thus interoperability between linked and spatial data systems were at the center of the considerations. With Phil Archer as the driving force, this led to shaping the *Spatial Data on the Web* working group charter¹⁶ and the foundation of the working group¹⁷ itself. The mission of the working group includes determining best practices on the integration, interlinking and discovery of spatial data on the Web and complete standardization of informal technologies already in wide-spread use. At this year's GeoLD workshop¹⁸, Frans Knibbe from the Spatial Data on the Web working group will hold the keynote on this topic and GeoKnow will present selected results.

6 GeoSPARQL support

There have been some minor improvements in Virtuoso GeoSPARQL support to resolve issues reported by users and customers of both the open source and commercial products, thus improving GeoSPARQL compliance and datatypes supported. Following on from the GeoSpatial Work shop in 2014 whereby a new W3C working group was proposed to be setup to foster collaboration between the OGC and W3C to create or recommend standards that work across both communities, the Data Shapes working group¹⁹ has now been established . OpenLink is a member of this working group and regularly attends meetings both f2f and tele conference calls. The LDBC FP7 Project that OpenLink was a partner in ended in March 2015 but continues as the LDBC Council, a not for profit organisation of which OpenLink is also a member, have held discussions proposed by OpenLink to add a geo spatial components to the Semantic Publishing Benchmark²⁰, hopefully in the next release. In the Fusepool P3 FP7 project use is made of the Geospatial data support in virtuoso for the Large-scale Technical Validation of Fusepool P3 Platform²¹ which use the Virtuoso LDP²² implementation and the German National Library dataset²³ as base data loaded for

¹⁵<http://www.w3.org/2014/03/lgd/>

¹⁶<http://www.w3.org/2015/spatial/charter>

¹⁷http://www.w3.org/2015/spatial/wiki/Main_Page

¹⁸http://geold.geoknow.eu/?page_id=64

¹⁹<http://www.w3.org/2014/data-shapes/charter>

²⁰<http://ldbncouncil.org/benchmarks/spb>

²¹<https://github.com/fusepoolP3/p3-wp5-deliverables/blob/master/d5.4-deliverable-report.md>

²²<http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/VirtLDP>

²³<http://www.dnb.de/EN/Standardisierung/GND/gnd.html>

performing this validation. An OSM to RDF Transformer²⁴ has been created for the platform enabling the transformation of OSM to RDF by the Fusspool P3 platform.

7 Linked Data Stack and Standardisation across EU Projects

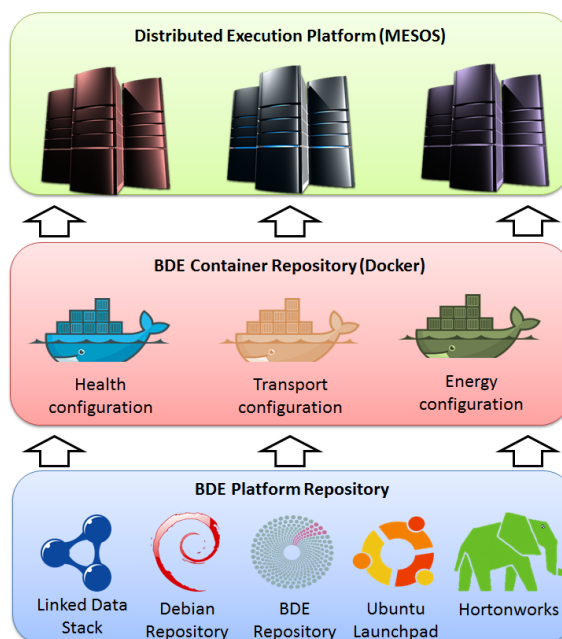


Figure 1: Overview of the deployment of the Big Data platform of the Big Data Europe project

The *Linked Data Stack*²⁵ (LDS) comprises Debian, RPM and Maven repositories featuring tools, libraries and datasets for fulfilling phases of the Linked Data Lifecycle. The stack was first presented at the Linked Data Europe Workshop²⁶ co-located with the European Data Forum 2014²⁷, and later at the GeoLD 2014 workshop²⁸ co-located with the SEMANTiCS²⁹ conference. GeoKnow will present the current state of the Linked Data Stack at this year's iteration of these events³⁰, fostering exchange between EU projects and the use of potential standard repositories like the Linked Data Stack. At present, the stack offers a total of 160 Debian packages (including *technical items*, such as libraries and utility packages), corresponding to around 40 tools and 60 datasets. Nightly, testing and stable versions are maintained in respective sub-repositories³¹.

Maintenance and enhancement of the Linked Data Stack will continue in several partner projects, namely Big Data Europe (BDE; InfAI)³² *Linked Data based crime analysis* (LiDaKrA; BROX)³³ *Semantic Analysis of Complex Events* (SAKE; Ontos, InfAI)³⁴ In addition, initial third-party contributions to the stack

²⁴<https://github.com/fusepoolP3/p3-transformer-configs/blob/master/openstreetmap/osm2rdf.xsl>

²⁵<http://stack.linkeddata.org/>

²⁶<http://www.linkeddataeurope.eu>

²⁷<http://2014.data-forum.eu>

²⁸http://geold.geoknow.eu/?page_id=64

²⁹<http://www.semantics.cc/>

³⁰<http://2015.data-forum.eu/>

³¹<http://stack.linkeddata.org/download/repo.php>

³²<http://www.big-data-europe.eu>

³³<http://eis.iai.uni-bonn.de/Projects/LiDaKrA.html>

³⁴<http://www.sake-projekt.de>

were made by the Media in Context (MICO)³⁵ project. For example, as can be seen from the depiction of BDE's deployment outline in [Figure 1](#), the LDS will play an important role by providing essential tooling and infrastructure packages.

8 Open Government Data and Standardisation activities across EU Projects

Since February 2014, PUPIN Institute is a part of the SHARE-PSI network - European network for the exchange of experience and ideas around implementing open data policies in the public sector. The main activity of the network is to organise a series of workshops examining different aspects of the revised European Directive on the Public Sector Information³⁶ and use the results of the workshops to propose best practices for the implementation of the Directive. SHARE-PSI, thus, complements the work of the W3C's Data on the Web Best Practices Working Group³⁷ aimed at providing best practices related to the publication and usage of data on the Web and facilitating the interaction between publishers and consumers. PUPIN team actively participated in the work of the programme committees for the following SHARE-PSI workshops

- Uses of open data within government for innovation and efficiency³⁸, Samos, 30 June to 1 July, 2014.
- Open Data Priorities and Engagement³⁹, Timișoara: 16 to 17 March, 2015.

Additionally, GeoKnow (Jens Lehmann, InfAI and Valentina Janev, PUPIN) and SHARE-PSI (Phil Archer, W3C Data Activity Lead, W3C/ERCIM) organized a Special session on: Open Data and Geographical Information System applications (ODaGIS) during the 5th International Conference on Information Society and Technology (ICIST 2015). Currently, PUPIN is involved in defining the SHARE-PSI best practices⁴⁰ that will be taken into consideration when writing the localized guidelines in SHARE-PSI partner countries⁴¹.

³⁵<http://www.mico-project.eu/>

³⁶<http://ec.europa.eu/digital-agenda/en/legal-rules>

³⁷<http://w3c.github.io/dwbp/bp.html>

³⁸<http://www.w3.org/2013/share-psi/workshop/samos/>

³⁹<http://www.w3.org/2013/share-psi/workshop/Timisoara/>

⁴⁰https://www.w3.org/2013/share-psi/wiki/Best_Practices

⁴¹https://www.w3.org/2013/share-psi/wiki/Localised_Guides