Geophysical Research Abstracts
Vol. 16, EGU2014-5958-7, 2014
EGU General Assembly 2014

Connecting geoscience systems and data using Linked Open Data in the Web of Data

Bernd Ritschel (1), Günther Neher (2), Toshihiko Iyemori (3), Yukinobu Koyama (3), Akiyo Yatagai (4), Yasuhiro Murayama (5), Ivan Galkin (6), Todd King (7), Shing F. Fung (8), Steve Hughes (9), Ted Habermann (10), Mike Hapgood (11), and Anna Belehaki (12)

(1) Helmholtz Centre Potsdam - GFZ German Research Centre for Geosciences, Potsdam, Germany, (2) University of Applied Sciences, Faculty of Information Sciences, Potsdam, Germany, (3) Kyoto University, World Data Center for Geomagnetism, Kyoto, Japan, (4) Nagoya University, Nagoya, Japan, (5) National Institute of Information and Communications Technology, Tokyo, Japan, (6) University of Massachusetts, Lowell, MA, USA, (7) University of California, Los Angeles, CA, USA, (8) NASA Goddard Space Flight Center, Greenbelt , MD, USA, (9) Jet Propulsion Laboratory, Pasadena, CA, USA, (10) HDF Group, Champaign, IL, USA, (11) STFC Rutherford Appleton Laboratory, RAL Space, Harwell Oxford, United Kingdom, (12) National Observatory of Athens, Athens, Greece

Linked Data or Linked Open Data (LOD) in the realm of free and publically accessible data is one of the most promising and most used semantic Web frameworks connecting various types of data and vocabularies including geoscience and related domains. The semantic Web extension to the commonly existing and used World Wide Web is based on the meaning of entities and relationships or in different words classes and properties used for data in a global data and information space, the Web of Data. LOD data is referenced and mash-uped by URIs and is retrievable using simple parameter controlled HTTP-requests leading to a result which is human-understandable or machine-readable. Furthermore the publishing and mash-up of data in the semantic Web realm is realized by specific Web standards, such as RDF, RDFS, OWL and SPARQL defined for the Web of Data. Semantic Web based mash-up is the Web method to aggregate and reuse various contents from different sources, such as e.g. using FOAF as a model and vocabulary for the description of persons and organizations -in our case- related to geoscience projects, instruments, observations, data and so on. On the example of three different geoscience data and information management systems, such as ESPAS, IUGONET and GFZ ISDC and the associated science data and related metadata or better called context data, the concept of the mash-up of systems and data using the semantic Web approach and the Linked Open Data framework is described in this publication. Because the three systems are based on different data models, data storage structures and technical implementations an extra semantic Web layer upon the existing interfaces is used for mash-up solutions. In order to satisfy the semantic Web standards, data transition processes, such as the transfer of content stored in relational databases or mapped in XML documents into SPARQL capable databases or endpoints using D2R or XSLT is necessary. In addition, the use of mapped and/or merged domain specific and cross-domain vocabularies in the sense of terminological ontologies are the foundation for a virtually unified data retrieval and access in IUGONET, ESPAS and GFZ ISDC data management systems. SPARQL endpoints realized either by originally RDF databases, e.g. Virtuoso or by virtual SPARQL endpoints, e.g. D2R services enable an only upon Web standard-based mash-up of domain-specific systems and data, such as in this case the space weather and geomagnetic domain but also cross-domain connection to data and vocabularies, e.g. related to NASA's VxOs, particularly VWO or NASA's PDS data system within LOD.

LOD - Linked Open Data
RDF - Resource Description Framework
RDFS - RDF Schema
OWL - Ontology Web Language
SPARQL - SPARQL Protocol and RDF Query Language
FOAF - Friends of a Friend ontology
ESPAS - Near Earth Space Data Infrastructure for e-Science (Project)
IUGONET - Inter-university Upper Atmosphere Global Observation Network (Project)
GFZ ISDC - German Research Centre for Geosciences Information System and Data Center
XML - Extensible Mark-up Language
D2R - (Relational) Database to RDF (Transformation)
XSLT - Extensible Stylesheet Language Transformation
Virtuoso - OpenLink Virtuoso Universal Server (including RDF data management)

NASA - National Aeronautics and Space Administration
VOx - Virtual Observatories
VWO - Virtual Wave Observatory
PDS - Planetary Data System