Geophysical Research Abstracts
Vol. 15, EGU2013-7958, 2013
EGU General Assembly 2013

Publishing high-quality climate data on the semantic web

Andrew Woolf (1), Armin Haller (2), Laurent Lefort (2), and Kerry Taylor (2)

(1) Climate & Water Division, Bureau of Meteorology, Canberra, Australia (a.woolf@bom.gov.au), (2) CSIRO ICT Centre, Canberra, Australia ({firstname.lastname}@csiro.au)

The effort over more than a decade to establish the semantic web [Berners-Lee et. al., 2001] has received a major boost in recent years through the Open Government movement. Governments around the world are seeking technical solutions to enable more open and transparent access to Public Sector Information (PSI) they hold. Existing technical protocols and data standards tend to be domain specific, and so limit the ability to publish and integrate data across domains (health, environment, statistics, education, etc.).

The web provides a domain-neutral platform for information publishing, and has proven itself beyond expectations for publishing and linking human-readable electronic documents. Extending the web pattern to data (often called Web 3.0) offers enormous potential. The semantic web applies the basic web principles to data [Berners-Lee, 2006]:

- using URIs as identifiers (for data objects and real-world 'things', instead of documents)
- making the URIs actionable by providing useful information via HTTP
- using a common exchange standard (serialised RDF for data instead of HTML for documents)
- establishing typed links between information objects to enable linking and integration

Leading examples of 'linked data' for publishing PSI may be found in both the UK (http://data.gov.uk/linked-data) and US (http://www.data.gov/page/semantic-web).

The Bureau of Meteorology (BoM) is Australia's national meteorological agency, and has a new mandate to establish a national environmental information infrastructure (under the National Plan for Environmental Information, NPEI [BoM, 2012a]). While the initial approach is based on the existing best practice Spatial Data Infrastructure (SDI) architecture, linked-data is being explored as a technological alternative that shows great promise for the future. We report here the first trial of government linked-data in Australia under data.gov.au.

In this initial pilot study, we have taken BoM's new high-quality reference surface temperature dataset, Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) [BoM, 2012b]. This dataset contains daily homogenised surface temperature observations for 112 locations around Australia, dating back to 1910. An ontology for the dataset was developed [Lefort et. al., 2012], based on the existing Semantic Sensor Network ontology [Compton et. al., 2012] and the W3C RDF Data Cube vocabulary [W3C, 2012]. Additional vocabularies were developed, e.g. for BoM weather stations and rainfall districts. The dataset was converted to RDF and loaded into an RDF triplestore. The Linked-Data API (http://code.google.com/p/linked-data-api) was used to configure specific URI query patterns (e.g. for observation timeseries slices by station), and a SPARQL endpoint was provided for direct querying. In addition, some demonstration 'mash-ups' were developed, providing an interactive browser-based interface to the temperature timeseries.

References

[Berners-Lee et. al., 2001] Tim Berners-Lee, James Hendler and Ora Lassila (2001), "The Semantic Web", *Scientific American*, May 2001.

[Berners-Lee, 2006] Tim Berners-Lee (2006), "Linked Data - Design Issues", W3C [http://www.w3.org/DesignIssues/LinkedData.html]

[BoM, 2012a] Bureau of Meteorology (2012), "Environmental information" [http://www.bom.gov.au/environment/]

[BoM, 2012b] Bureau of Meteorology (2012), "Australian Climate Observations Reference Network - Surface Air Temperature" [http://www.bom.gov.au/climate/change/acorn-sat/]

[Compton et. al., 2012] Michael Compton, Payam Barnaghi, Luis Bermudez, Raul Garcia-Castro, Oscar Corcho, Simon Cox, John Graybeal, Manfred Hauswirth, Cory Henson, Arthur Herzog, Vincent Huang, Krzysztof Janowicz, W. David Kelsey, Danh Le Phuoc, Laurent Lefort, Myriam Leggieri, Holger Neuhaus, Andriy Nikolov, Kevin Page, Alexandre Passant, Amit Sheth, Kerry Taylor (2012), "The SSN Ontology of the W3C Semantic Sensor Network Incubator Group", *J. Web Semantics*, 17 (2012) [http://dx.doi.org/10.1016/j.websem.2012.05.003]

[Lefort et. al., 2012] Laurent Lefort, Josh Bobruk, Armin Haller, Kerry Taylor and Andrew Woolf (2012), "A Linked Sensor Data Cube for a 100 Year Homogenised daily temperature dataset", *Proc. Semantic Sensor Networks 2012* [http://ceur-ws.org/Vol-904/paper10.pdf]

[W3C, 2012] W3C (2012), "The RDF Data Cube Vocabulary", [http://www.w3.org/TR/vocab-data-cube/]