Geophysical Research Abstracts
Vol. 20, EGU2018-8054, 2018
EGU General Assembly 2018

Exposing the SeaDataNet metadata catalogues as Linked Data: from metadata to data on the web

Rob Thomas and Adam Leadbetter
Marine Institute, Ocean Science and Information Services, Ireland (rob.thomas@marine.ie)

Linked Data is a method for publishing structured data on the World Wide Web, and using Web addresses to provide connections between published data objects. Linked Data was proposed by Sir Tim Berners-Lee as a foundation to the "Web of Data". Until the emergence of the "Internet of Things", the "Web of Data" was less well known in contrast to the traditional "Web of Documents" that is well used and now integral to everyday life. Linked Data has gained ground in the marine science data management community over recent years, with projects in the United States (such as Rolling Deck to Repository and the Biological and Chemical Oceanography Data management Office) publishing records as Linked Data. Under the EU H2020 funded SeaDataCloud project, Linked Data representations of the SeaDataNet (SDN) metadata catalogues have been created with the aim to allow ease of interoperability at a global scale with these initiatives, and also allow a formal alignment with some of the INSPIRE Spatial Data Infrastructure requirements and the EU Data Portal.

By using a Linked Data approach, information relevant to more than one catalogue does not have to be replicated but can be inherited or inferred by linkages between the catalogues, connections between the metadata are more visible and with these linkages in place it is easier to allow the end user to browse from relevant information in one catalogue to the information they are ultimately seeking. Where possible this work has reused existing Linked Data standards for the publication of the SDN catalogues, in compliance with general World Wide Web data publishing best practices. Using domain specific vocabulary resources provided by the NERC Vocabulary Server (NVS) ties in with existing SDN infrastructure. This is consistent with the current SDN catalogue content and further integrates the catalogues into established Linked Data resources. Marine data interoperability is enhanced as NVS content has been used by the international marine informatics community by US, Canadian and Australian marine organisations within their Linked Data publications.

This presentation lays out the Linked Data patterns for publishing the metadata held within the SDN catalogues and the pre-existing standards these were aligned with:
- European Directory of Marine Organisations (EDMO) aligned with W3C Organization Ontology,
- European Directory of Marine Environmental Reseach Projects (EDMERP) aligned with W3C Provenance Ontology / DBPedia Research Project,
- European Directory of Integrated Observing Systems (EDIOS) aligned with INSPIRE Environmental Monitoring Facilities,
- European Directory of Marine Environmental Datasets (EDMED) aligned with W3C Data Catalogue (DCAT) / Google (schema.org) Science Datasets,
- Cruise Summary Reports (CSR) based on a proposed new ontology,
- Common Data Index (CDI) aligned with W3C Data Catalogue (DCAT) & ISO / OGC Observations & Measurements / OWL representation of ISO 19156 (Sampling Features model).

The presentation concludes by introducing how the SDN linked (meta)data can be leveraged through providing data via the cloud or APIs, such as ERDDAP, and in search engine results as "rich snippets".