Using a linked data approach to aid development of a metadata portal to support Marine Strategy Framework Directive (MSFD) implementation

Chris Wood

British Oceanographic Data Centre, National Oceanography Centre, Liverpool, United Kingdom (chwood@bodc.ac.uk)

Under the Marine Strategy Framework Directive (MSFD), EU Member States are mandated to achieve or maintain 'Good Environmental Status' (GES) in their marine areas by 2020, through a series of Programme of Measures (PoMs). The Celtic Seas Partnership (CSP), an EU LIFE+ project, aims to support policy makers, special-interest groups, users of the marine environment, and other interested stakeholders on MSFD implementation in the Celtic Seas geographical area. As part of this support, a metadata portal has been built to provide a signposting service to datasets that are relevant to MSFD within the Celtic Seas. To ensure that the metadata has the widest possible reach, a linked data approach was employed to construct the database. Although the metadata are stored in a traditional RDBS, the metadata are exposed as linked data via the D2RQ platform, allowing virtual RDF graphs to be generated. SPARQL queries can be executed against the end-point allowing any user to manipulate the metadata. D2RQ's mapping language, based on turtle, was used to map a wide range of relevant ontologies to the metadata (e.g. The Provenance Ontology (prov-o), Ocean Data Ontology (odo), Dublin Core Elements and Terms (dc & dcterms), Friend of a Friend (foaf), and Geospatial ontologies (geo)) allowing users to browse the metadata, either via SPARQL queries or by using D2RQ's HTML interface. The metadata were further enhanced by mapping relevant parameters to the NERC Vocabulary Server, itself built on a SPARQL endpoint. Additionally, a custom web front-end was built to enable users to browse the metadata and express queries through an intuitive graphical user interface that requires no prior knowledge of SPARQL. As well as providing means to browse the data via MSFD-related parameters (Descriptor, Criteria, and Indicator), the metadata records include the dataset's country of origin, the list of organisations involved in the management of the data, and links to any relevant INSPIRE-compliant services relating to the dataset. The web front-end therefore enables users to effectively filter, sort, or search the metadata. As the MSFD timeline requires Member States to review their progress on achieving or maintaining GES every six years, the timely development of this metadata portal will not only aid interested stakeholders in understanding how member states are meeting their targets, but also shows how linked data can be used effectively to support policy makers and associated legislative bodies.