Applying the RDF DataCube model to power data visualization and exploration dashboards for the Irish Wave and Weather Buoy Networks

Rob Thomas, Marine Institute (Ireland), rob.thomas@marine.ie
Donnchadh Ó Foghlú, Marine Institute (Ireland), donnchadh.ofoghlu@marine.ie
Kieran Lyons, Marine Institute (Ireland), kieran.lyons@marine.ie
Will Meaney, Marine Institute (Ireland), will.meaney@marine.ie

Motivation

Public administration frequently publishes Open Data in an ad-hoc manner. Yet society expects data-driven public services, not raw data. In order to ensure that public services address their needs society should be involved the design and production of such services. Within the software development world, these continuous interactions are termed "co-design" and "co-production".

There are challenges for the producers and consumers of open data, for example are the data small or large volume, how frequently are they updated, who is permitted access, how easy are they to access, etc. Within this context, data literacy (in terms of knowledge, experience and ability) of users is often overlooked when providing solutions.

OpenGovIntelligence (OGI) project

The Horizon 2020 funded OpenGovIntelligence project (OGI) planned to support and develop approaches to co-design and service co-production by providing software tools to aid decision making and better manage the complexities and precise nature of high quality statistical data. OGI sought to prove the effectiveness of the OGI approach through a diverse selection of pilot projects in six countries. The aim of the pilots was to develop services at both national and local levels to tackle challenges within society and Public Administration. The central model for data provision around which the OGI project planned to build tool was the RDF Data Cube model[1].

The RDF Data Cube model provides a vocabulary for publishing multi-dimensional data, such as statistics, on the web using the W3C RDF (Resource Description Framework) standard. The Data Cube model is compatible with the cube model on which the SDMX (Statistical Data and Metadata eXchange), an ISO standard for exchanging and sharing statistical data and metadata among organizations such as government statistics offices, is founded.

OGI Irish Pilot data dashboard - Irish Wave and Weather Buoy Networks

The Irish Wave and Weather Buoy Networks return metocean data at 5-60 minute intervals from 9 locations in the seas around Ireland. Outside of the Earth Sciences an example use case for these data is in supporting Blue Economy development and growth (e.g. renewable energy device development). Access to wind and wave energy data are valuable information in supporting decision making by the marine renewables community, specifically in selecting suitable locations in coastal and offshore zones for deployment as well as informing the design or selection of wave energy converters for use in these areas. While the full resolution and range of data collected are valuable, summary statistics (e.g. by season) are key to designing devices optimized to exploit the prevailing conditions for a location or to determine extreme event frequency to ensure structures are designed for such conditions.

The Marine Institute, as the operator of the buoy platforms, in partnership with the OGI project has published daily summary data from the Irish Wave and Weather Buoys using the RDF DataCube model.

---

[1] https://www.w3.org/TR/vocab-data-cube/

These daily statistics are available as Linked Data via a SPARQL endpoint API making these data semantically interoperable and machine readable. This API underpins a pilot dashboard for data exploration and visualization (Fig. 1). The dashboard[2] presents the user with the ability to explore the data and derive plots for the historic summary data, while interactively subsetting from the full resolution data behind the statistics.
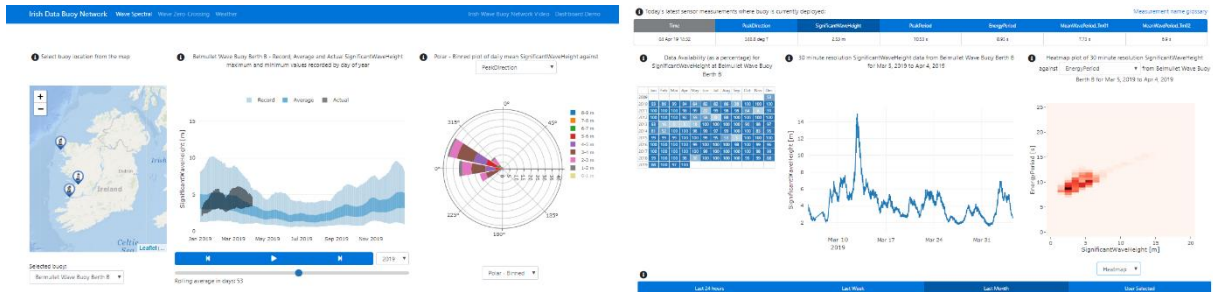

Figure 1: Irish Data Buoy Network data visualisation and exploration dashboard.

The pilot dashboard draws together summary statistics data from the SPARQL endpoint and complements these data with real-time data (using the Internet of Things MQTT protocol) and full resolution historic data from an instance of the NOAA ERDDAP data broker with a GraphQL API. GraphQL was developed by Facebook before being publicly released in 2015. Key features being the efficiency in returning results in the same structure as specified by the query and a query flexibility not available from REST services. Publishing environmental data with these technologies makes accessing environmental data available to developers outside those with Earth Science involvement and effectively lowers the entry bar for usage by providing data in self describing schemas for the RDF DataCubes and GraphQL API.

This presentation will highlight the benefits achieved by joining the project as a "pilot partner" and lesson learnt from the process. As well as discussing the technologies deployed and user feedback on from an end-user survey of the dashboard. The presentation will discuss how the dashboard is one of a range of access points that aim to cover the range of users' data literacy and interest while minimising the challenges faced by those wishing to consume data from the Irish Weather and Wave Buoy Network.

---

[2] https://vis.marine.ie/dashboards/