



Data Quality, Provenance and IPR Management services: their role in empowering geospatial data suppliers and users

Keiran Millard

HR Wallingford, United Kingdom (k.millard@hrwallingford.com)

This paper looks at current experiences of geospatial users and geospatial suppliers and how they have been limited by suitable frameworks for managing and communicating data quality, data provenance and intellectual property rights (IPR). Current political and technological drivers mean that increasing volumes of geospatial data are available through a plethora of different products and services, and whilst this is inherently a good thing it does create a new generation of challenges.

This paper consider two examples of where these issues have been examined and looks at the challenges and possible solutions from a data user and data supplier perspective. The first example is the IQmulus project that is researching fusion environments for big geospatial point clouds and coverages. The second example is the EU Emodnet programme that is establishing thematic data portals for public marine and coastal data.

IQmulus examines big geospatial data; the data from sources such as LIDAR, SONAR and numerical simulations; these data are simply too big for routine and ad-hoc analysis, yet they could realise a myriad of disparate, and readily useable, information products with the right infrastructure in place. IQmulus is researching how to deliver this infrastructure technically, but a financially sustainable delivery depends on being able to track and manage ownership and IPR across the numerous data sets being processed. This becomes complex when the data is composed of multiple overlapping coverages, however managing this allows for uses to be delivered highly-bespoke products to meet their budget and technical needs.

The Emodnet programme delivers harmonised marine data at the EU scale across seven thematic portals. As part of the Emodnet programme a series of 'check points' have been initiated to examine how useful these services and other public data services actually are to solve real-world problems. One key finding is that users have been confused by the fact that often data from the same source appears across multiple platforms and that current 19115-style metadata catalogues do not help the vast majority of users in making data selections.

To address this, we have looked at approaches used in the leisure industry. This industry has established tools to support users selecting the best hotel for their needs from the metadata available, supported by peer to peer rating. We have looked into how this approach can support users in selecting the best data to meet their needs.