# Linking Geospatial Data – BGS Position Paper

**John Laxton, BGS Edinburgh (jll@bgs.ac.uk)**

The British Geological Survey is a part of the Natural Environment Research Council (NERC) and is its principal supplier of national capability in geoscience. BGS aims to advance understanding of the structure, properties and processes of the solid Earth system through interdisciplinary surveys, monitoring and research for the benefit of society. BGS is the UK's premier provider of objective and authoritative geoscientific data, information and knowledge for wealth creation, sustainable use of natural resources, reducing risk and living with the impacts of environmental change. Traditionally the main means of making information available was through the publication of geological maps, originally paper and more recently digital as the product of GIS-based spatial databases.

Digital maps provide the opportunity to make explicit much underlying information about the geology that was implicit on traditional maps, but generally only understood by experts. OGC web services enable this information to be delivered to users but in order to maximise the benefit and enable interoperability between data providers it was necessary for the international geological community to develop a common conceptual information model and interchange specification. The result was GeoSciML (http://www.cgi-iugs.org/tech_collaboration/geosciml.html) which is underpinned by several established OGC and ISO standards, including Web Feature Service (WFS – ISO 19142), Geography Markup Language (GML – ISO 19136), Observations & Measurements (O&M – ISO 19156), and SWE Common. GeoSciML v4 is currently being developed as an OGC standard. The GeoSciML standard was used as the basis for developing the INSPIRE Geology Data Specification. The use of GeoSciML web services to meet INSPIRE requirements will allow INSPIRE data to be extended to other types of geosciences information in an internationally conformant way.

A range of vocabularies have been developed to constrain GeoSciML fields and these are available in SKOS RDF (http://resource.geosciml.org/). Some vocabularies are a simple list of defined concepts, but others are a more complex set of linked concepts. For example the litostratigraphic lexicon provides a description of stratigraphic rock units which includes information on age and lithology. These are in turn complex concepts where, for example, a particular rock lithology can be described in terms of its individual components.

This clearly lends itself to the use of linked data and BGS has set up a linked data service (http://data.bgs.ac.uk/home.html) to test the potential and get feedback from users. This initial service is centred on the BGS 1:625k geological map and provides links from spatial objects on the map to ontologies for the lithostratigraphic lexicon, geochronology, and earth material (lithology).

We are also in very early discussion with the Centre for Ecology and Hydrology (CEH) and the British Oceanographic Data Centre (BODC) on the potential of linking environmental chemistry data across NERC. i.e. exposing our GBASE (geochemical) data in linked data form using common vocabularies in order that they can be compared with chemical field sampling datasets from CEH and marine sea and sediment sampling from BODC, thus enabling tracking of contaminants from "source to sea". The main problem is not so much technological as scientific, in that the numbers themselves can't give any answers, they have to be used with understanding of the methods involved and in context of the regional picture or expected baseline concentration.

Another potential future application of linked data would be as a means of developing cross-domain linkages. Our current implementation, while making much more data accessible and explicit, still assumes geoscientific understanding in making sensible linkages. We would like to develop applications that enable professionals from other disciplines, or the general public, to query our data using terminology they understand. The first part of this would be to build ontologies embedding the scientific links between concepts to capture geoscientific understanding. Ontologies would then need to be developed linking across domains – a much bigger task. The effort required in developing rich ontologies is at present the main barrier to developing powerful linked data applications.

Geological maps are increasingly being replaced by digital geological spatial models which make the three-dimensional geometry of geological objects explicit. One of the objectives of this is to aid understanding of the geometry by non-experts who commonly find geological maps difficult to interpret. Delivering models over the web to end-users is more difficult than delivering maps, and as part of the FP7 EarthServer project (http://www.earthserver.eu/) we have been looking at delivering model surfaces as coverages using WCS. A service has been implemented (http://earthserver.bgs.ac.uk/) which delivers a range of different types of geosciences coverage, including model surfaces, and a GUI has been developed to provide a query interface using WCS and WCPS along with visualisation through appropriate clients, including a 3D client.

At present the surfaces just carry a name, but in order to be interrogated in a useful way they need much richer description. We are investigating using GeoSciML to link the model surfaces to the implicit rock bodies between the surfaces, and to provide the sort of information for these rock bodies, such as age and lithology, that would be available from a geological map. The best way to achieve this appears to be to put the GeoSciML into the coverage metadata and we have implemented this as a trial. Another EarthServer project partner is developing a combined xQuery WCPS query language which we aim to test against queries such as 'find those geological units comprised mainly of sand within 25m of the surface'. Queries of this type have wide potential application and their implementation through a query interface against a service delivering model surface coverages would be very useful.

BGS has also developed mobile geospatial applications (http://www.bgs.ac.uk/data/mobileApps.html):

- iGeology, for iPhone/iPad and Android, gives you an interactive geological map of the UK.

- iGeology 3D, for Android, uses augmented reality technology to 'paint' a geological map on the ground around you. It can be used alone or as a companion product to iGeology.
- *my*Soil, for iPhone/iPad and Android devices, gives you an interactive map of the soils of Britain.

One of the principal benefits of mobile apps is the ability to link a user's current location with a range of geospatial data. These apps rely on web services to deliver this information to the user's phone. Extending these web services to use information from linked ontologies would enhance the richness of the information provided by such smartphone apps, creating tools that provide interdisciplinary solutions to environmental problems.