



> Semantic Web Use Cases and Case Studies

Case Study: Semantic Web Technology at Ordnance Survey

Catherine Dolbear, Ordnance Survey

March 2007



General Description

Ordnance Survey is Great Britain's national mapping agency. We create and update the definitive mapping data of England, Scotland and Wales. From this we produce and market a wide range of digital information and paper maps for business, leisure, educational and administrative use. Every day, we make more than 5,000 updates to our central database: the largest vector geospatial database in the world. These revisions reflect house extensions, major new buildings and other natural and man-made changes.

By maintaining and enhancing one of the world's most detailed geographic referencing frameworks, we provide an important resource for the nation. Our geographic information is extensively used throughout the public-sector. Specific uses include land and property registration; locating suitable derelict sites for house building; identifying areas of deprivation; planning new access to the countryside; controlling the flow of urban traffic; helping the police monitor crime patterns and catch offenders and a wide range of environmental applications. And use within the private sector is equally wide ranging: customer profiling; calculating insurance risk; managing property portfolios; transport logistics systems; and so on. Our data is an integral component of web directories, in-car navigation systems and many mobile phone applications. In fact, independent estimates show that Ordnance Survey's geographic information underpins £100 billion worth of economic activity in Britain.

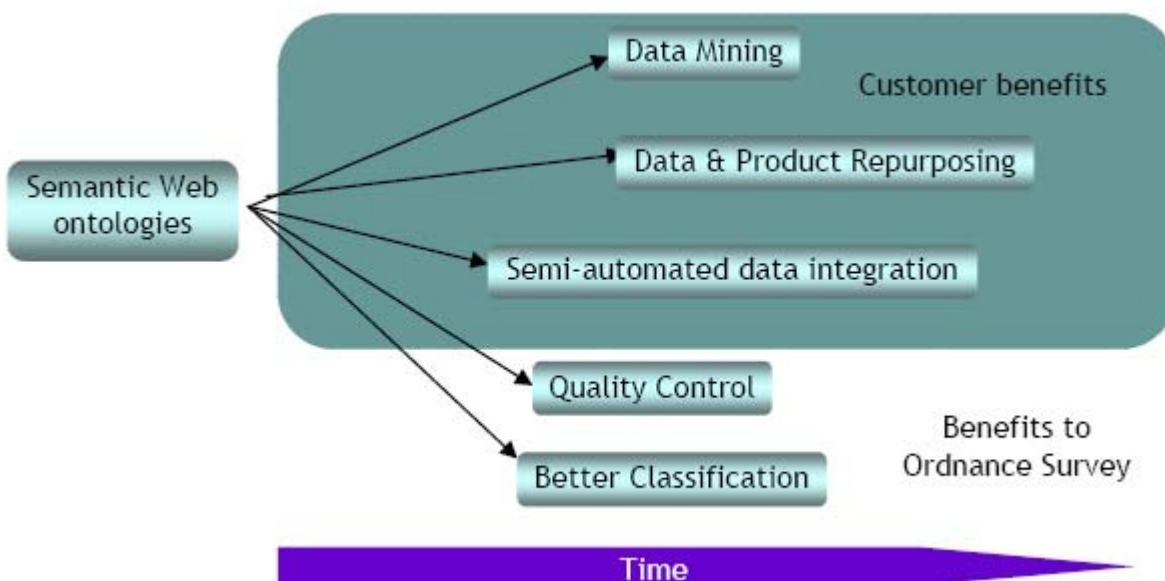


Figure 1: Organisational and customer benefits of semantic research

Key Benefits of Using Semantic Web Technology

Ordnance Survey has developed a new generation of highly detailed and structured digital data in a series of high precision layers – Topography, Addressing, Imagery and Integrated Transport Network™, comprising more than half a billion unique features such as buildings, roads, fences and so on. Our geographical data is rarely used in isolation, but is combined with the customer's own data and knowledge to solve their problems. Its role is to provide an underpinning framework for other data and therefore is central to many tasks that require data from different sources to be integrated.

We are investing in semantic technology research to cut the cost and improve the accuracy of **data integration**. Data integration is widely acknowledged to be expensive and problematic. When such activities involve the inclusion of complex datasets such as OS MasterMap® the challenge becomes that much greater. The problem is more than a syntactic one and cannot be solved by everyone adhering to one file format or standard. Rather it is also a problem of understanding what the data means: is the Ordnance Survey definition of a field the same as the way the Government Department for the Environment, Food and Rural Affairs (DEFRA) would understand it? Our definition is spatially delineated by barriers such as fences or ditches, while DEFRA might distinguish a field's extent by the kind of crops that were grown, as part of their task of calculating farmers' subsidies. By providing a machine-readable description known as an *ontology* of Ordnance Survey's data, these semantic differences are made explicit, the impact on the customer's application can be seen and we can begin to use ontologies to provide bridges between these different world views. As the semantic web community develops techniques for ontology merging, and ontology to database mapping, it will become easier to semi-automatically integrate our data with our customers.

Another benefit of semantic technology is that of **data repurposing**. Our customers have their own terminology for the things they use day to day: and they may not be the same words that Ordnance Survey uses. For example, the Environment Agency uses the term "Flood Defence" – but that doesn't appear as an attribute in our data, although we do capture the Weirs and Flood Walls and so on that are used as flood defences. By authoring an ontology to describe Ordnance Survey data, and one describing the terms a customer does use, we have already shown that it's possible to query an Ordnance Survey database for terms like "Flood Defence".

There are also internal business advantages for Ordnance Survey: with an explicit, machine readable description of the meaning of our data, we can improve **quality control** and **classification** decision-making.

Ordnance Survey's Semantic Work

To date we have built a domain ontology for Hydrology, along with reusable modules for some spatial and mereological relations. We are publishing an Administrative Geography for Great Britain in RDF and are currently working on an ontology for Buildings and Places, with the aim of eventually producing an ontology covering the whole of topography. Efficient querying of our data, either in a relational model via an ontology, or directly in RDF is an important topic for Ordnance Survey, given the scale of our database. We have also invested time in the development of tools and techniques for ontology authoring by domain experts – those skilled in their fields, but with little or no knowledge of the semantic web. Allowing people to write ontologies in natural-sounding English enables us to capture knowledge direct from the source, and takes some of the mystery out of semantic web ontologies. Equally importantly, it enables other domain experts to understand the ontologies, so they can either validate or reuse the ontology to assist them in tasks such as data integration.

As the national mapping agency of Great Britain we are taking a lead in semantic web technology to facilitate the future we see beyond maps: where our geospatial knowledge, not just data, can be queried, understood and integrated by our customers more easily, in a more accurate and cost effective way.

All Ordnance Survey ontologies are available at <http://www.ordnancesurvey.co.uk/ontology>

© Crown Copyright 2007, Ordnance Survey