The Challenge of Delivering Geoscience Data for Public Use

John Broome
Earth Science Sector, NRCan, Ottawa

The world has adopted online sources as the preferred source for data and information. More recently there have been a dramatic increase in the availability of spatial data that can be visualized online. Geoscience data are also often online for visualization and download but these data are often of limited use to non-geoscientists due to their complexity and inconsistency. This paper will explore the challenge of delivering geoscience data online for non-specialist use and propose work that needs to be done to address this significant challenge.

Throughout the world, data spatial infrastructures are being built. This process has resulted in much improved understanding and methodology for the online integration and delivery of geospatial data. The foundation of spatial data infrastructures is open access to data combined with data and service standards that permit data to be discovered, visualized, and integrated online. In Canada, GeoConnections (http://geoconnections.org) is building the Canadian Geospatial Data Infrastructure, or CDGI.

There are many types of geoscience data but this paper will focus on how to deliver some of the information typically represented in a geological map or its equivalent digital coverage. The information in such a "map" is a complex mix of structure, lithology, genesis, and age typically delivered using a structure and classification that is specific to the organization, author, and scale.

It is clear that the utility of geoscience data will be maximized if they can be integrated with other CGDI spatial data through adoption of common standards and delivered using standard online data services. Given the objective of delivering data to non-specialists geology will ideally be delivered: a) seamlessly across Canada regardless of scale , b) described using consistent language appropriate for the user, who will often not be a geoscientist.

These are not trivial challenges considering the traditional approach to classification of geoscience data was not designed for data management. Delivering seamless geological coverage will require the delivery of data classified using a consistent taxonomy. The simplest technical solution would be for all organizations delivering geology to adopt common data standards but this is a significant cultural challenge. A more viable approach is for organizations to continue to manage data using their existing approach but to translate data to a commonly-agreed standard for dissemination.

Standards for general geospatial metadata (ISO) and data are already mature and a geological data exchange schema, or structure, for exchange of geological data in GML format called GeoSciML has been developed by an international team of data experts coordinated by the IUGS Commission for the Management and Application of Geoscience Data (CGI). GeoSciML is part of the solution but consistent classification and terminology appropriate for the intended use is also required for a complete solution. Considerable work is currently ongoing to address the classification/terminology issue. In Canada, the Canadian Lexicon of Geological Names is available on the CGKN website and, internationally, the CGI and the GeoSCIML Team are coordinating international efforts to establish the required standards.

Once geological data are accessible in standard form, existing technology based on OGC web services for discovery (WCS), visualization (WMS), and access (WFS, GML) can be used to

make the data accessible. The OneGeology initiative (http://onegeology.org) demonstrates how the WMS and WFS standards combined with GeoSCIML can be used to display geological coverage online. A number of OneGeology participants currently deliver their data to the portal using OGC web services and GeoSciML, providing a proof of concept for the proposed model. The map coverage currently displayed from these services is inconsistent because the data services have not adopted a standard classification.

Clearly the required technology and standards required to address the challenge of delivering seamless and consistent geology are becoming available. Organizations must recognize the importance of investing in establishing the standards, policies, and practices necessary to permit them to deliver consistent geological data. By pursuing this path, the use and utility of geology data will be dramatically enhanced in the future.