Global Geoscience Data Transfer Standards - GeoSciML and EarthResourceML

Ollie Raymond¹, Jouni Vuollo², Steve Richard³, CGI Interoperability Working Group⁴

¹ Geoscience Australia, Canberra, Australia, oliver.raymond@ga.gov.au
² Geological Survey of Finland, Rovaniemi, Finland, jouni.vuollo@gtk.fi
³ Arizona Geological Survey, Tucson, USA, steve.richard@azgs.az.gov

BACKGROUND

The Interoperability Working Group of the International Union of Geological Sciences’ Commission for the Management and Application of Geoscience Information (IUGS-CGI) was formed in 2004 to develop international data transfer standards for geoscience information. Originally tasked with delivering a standard for data that is typically found on a geological map, the IWG developed GeoSciML, a UML data model and GML-based markup language based on the spatial data standards of the Open Geospatial Consortium (OGC).

GEOSCIML

Since the release of version 1 in 2006, GeoSciML has grown in scope to cover a wide range of geological feature and data types. The current release, GeoSciML v3 (http://www.geosciml.org, Raymond et al, 2012), covers geological units, geological structures (e.g., faults, folds, foliation), unit contacts, earth materials (e.g., rocks, minerals, unconsolidated materials), geological ages and events, boreholes, geomorphological features, petrophysical rock properties, geological specimens and sampling processes, and analytical metadata.

GeoSciML v3 makes use of recently upgraded OGC and ISO standards, including GML v3.2, SWECommon v2, and Observations & Measurements (O&M) v2. Links to the O&M standard are strong in relation to delivery of geological sampling and analytical data such as geochemistry and geochronology. Some previously used GeoSciML data types have been superseded in favour of data types provided by OGC’s SWECommon and GML data standards. The version 3 model has also been refactored from a single application schema in version 2 into a number of smaller, more manageable schema modules with individual namespaces (e.g., GeologicAge, EarthMaterial, PhysicalProperties). As a result, GeoSciML v3 is not backwardly compatible with previous versions.

GeoSciML v4 is being prepared as an Open Geospatial Consortium (OGC) standard. Development of GeoSciML v4 will occur within the OGC’s Earth Systems Science DWG under a collaborative IUGS-OGC arrangement. It is anticipated that version 4 of GeoSciML will address user concerns about the perceived complexity of the model, and allow easier implementation of certain core GeoSciML components.

EARTHRESOURCEML

EarthResourceML v2 (http://www.earthresourceml.org, Vuollo et al, 2012) is the latest version of the CGI-IUGS Interoperability Working Group (IWG) interchange standard for earth resources data such as mineral deposits and mining activities. The IWG has worked with the European Union’s INSPIRE project to modify the data standard to meet EU requirements. EarthResourceML now forms the basis of the INSPIRE Mineral Resources data specification.

In version 2, EarthResourceML has been extended to cover all solid earth resources. The single application schema contains two principal components: one centred on the EarthResource — i.e., the MineralOccurrence feature, along with its associated commodities and ore measures which describe the material of potential economic value; and the other centred on a MiningFeature, which uses Mine and Mining Activity features to describe the exploitation of the EarthResource. EarthResourceML also provides a formal structure for reporting resources and reserves that complies with national and international reporting codes. The EU INSPIRE requirements saw the addition of some mineral exploration and environmental aspects to the current version, such as description of exploration activity and mining waste.

EarthResourceML v2 is compatible with GeoSciML v3 and uses model and schema patterns and features common to GeoSciML. The major refactoring of GeoSciML between versions 2 and 3, along with the upgrade to OGC GML v3.2 and SWE Common Data Model v2 required a major version upgrade to EarthResourceML from v1 to v2.

PORTRAYAL STANDARDS
GeoSciML-Portrayal (Richard et al, 2012) is an OGC GML simple-feature application schema based on a small, highly simplified core of GeoSciML. It was developed to provide a schema for presentation of geologic map units, contacts, and shear displacement structures (i.e., fault and shear zones) in Web Map Services (WMS). The schema includes text attributes for human users to read when browsing a geologic map, a hyperlink to a full GeoSciML feature element (if available), and a symbol identifier field to enable a user-defined symbolization scheme in each map service. The schema establishes naming conventions for attributes commonly used to symbolize geological maps. The use of standard vocabularies with these attributes for geological age and lithology enables map symbolization using shared legends to achieve visual harmonization of maps provided by different services.

Deployment of Web Map and Web Feature Services from a single data source is relatively straightforward. Providing a simple-feature WFS that can be bound to the WMS allows clients to acquire basic text geologic feature descriptions that can be used in web mapping applications to construct custom legends using GeoSciML-Portrayal, as well as to identify and select features for further processing that are acquired as highly structured, information-rich, GeoSciML complex features. Similar GML simple-feature portrayal schemas are proposed for mineral occurrences, mines and mining activities (i.e., EarthResourceML features), boreholes, and geological specimens.

VOCABULARIES AND SUPPORTING INFRASTRUCTURE

The IUGS-CGI Interoperability Working Group has developed vocabularies for classification of lithologies, fault and contact types, and many other geoscience concepts to support data provided by the GeoSciML and EarthResourceML schemas. The IWG have established a vocabulary service at http://resource.geosciml.org for these vocabularies in RDF-SKOS and HTML formats that can be accessed by GeoSciML web services. The vocabulary service also includes RDF representations of terms for geological ages which are defined by the International Commission on Stratigraphy (ICS) Stratigraphic Chart. Further vocabulary development, notably to support the EarthResourceML data standard, will be undertaken by the new IUGS-CGI Geoscience Terminology Working Group. The new GeoSciML v3 and EarthResourceML v2 schemas do not include vocabularies in them. However, the data models recommend a standard pattern to reference the controlled vocabularies using HTTP-URI links.

The new GeoSciML and EarthResourceML releases will include example XML instances showing best practice encoding for delivering a range of geoscientific data such as geological units, structures, earth materials and geochemistry. In addition, an exemplar database has been developed to assist users with model testing and demonstration of web service implementation from a database environment. Schematron scripts have also been developed to assist in validation of GeoSciML WFS services and to support a higher degree of semantic interoperability and accreditation of GeoSciML services.

INTERNATIONAL USE OF IUGS-CGI STANDARDS

GeoSciML, GeoSciML-Portrayal and EarthResourceML have been adopted or recommended as the web service data exchange standards in key international interoperability initiatives. They include the global OneGeology project (http://www.onegeology.org), the European Union INSPIRE Directive, the US Geoscience Information Network (http://usgin.org), the Canadian Groundwater Information Network (http://www.gw-info.net), the Australian AuScope project (http://portal.auscope.org/portal/gmap.html), the Australia/NZ Government Geoscience Information Committee (http://www.geoscience.gov.au), the Visualising Victoria’s Groundwater portal (http://www.vvg.org.au) and the African-European Georesources Observation System (AEGOS, http://www.aegos-project.org). Major software manufacturers, such as ESRI, are working closely with some of these data sharing initiatives and are actively seeking to implement IUGS-CGI data standards.

REFERENCES


Ollie Raymond is Senior Information Geologist in the Continental Geology Section of Geoscience Australia, with 20 years experience of designing, compiling, and delivering digital geoscience data and is the current custodian of Geoscience Australia's national geological maps and databases. Ollie is also currently a CGI Council member and has been the chair of the IWG GeoSciML design task group since 2008.

Dr Jouni Vuollo is Chief Geologist for Data Management and Quality at the Geological Survey of Finland (GTK), overseeing the compilation, quality and delivery of Finland's national geological databases. He is a member of the technical working group for the European Union’s INSPIRE geoscience and mineral resources data specifications, and is the current chair of the IWG EarthResourceML design task group.

Dr Steve Richard is head of geoinformatics at the Arizona Geological Survey. Steve has over 20 years experience in geological mapping, data modelling, and geoscience data delivery. He is a founding member of the IWG GeoSciML design task group, is currently the interim chair of the CGI Geoscience Terminology Working Group, and is a driving force behind the technical specifications for the US Geoscience Information Network.

The Commission for the Management and Application of Geoscience Information (CGI) is a commission of the International Union of Geological Sciences, the peak international body for geological sciences, and was formed in 2004. The CGI’s aims to

- provide the means for transferring knowledge on geoscience information and systems
- stimulate international dissemination of best practice in geoscience information
- stimulate and support initiatives which are developing standards, and
- establish and occupy an accepted position in the international geoscience information community and represent IUGS on geoscience information matters.

The CGI’s Interoperability Working Group (IWG) currently includes sixteen geoscientists and information modellers from nine different countries. The IWG works to design and implement the GeoSciML and EarthResourceML data transfer standards, and to coordinate supporting infrastructure for the transfer of interoperable geoscience information.