



Newsletter

In This Issue

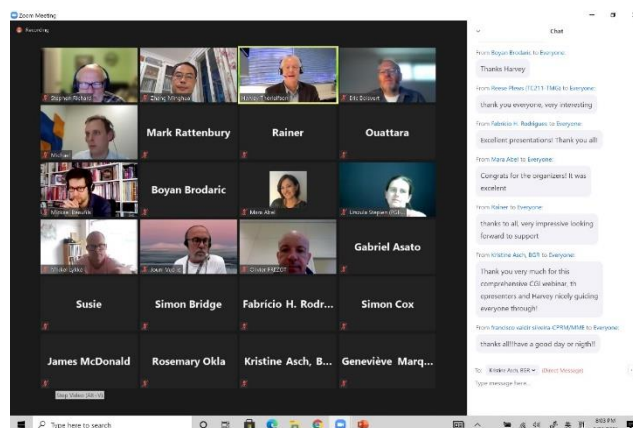
- Overview of CGI activities – Harvey Thorleifson
- OGC Geoscience Domain – Mickaël Beauflis
- GeoSciML – Éric Boisvert
- EarthResourceML – Michael Sexton
- CGI Geoscience Vocabularies – Mark Rattenbury
- Loop GeoScience Ontology – Boyan Brodaric and Stephen Richard
- DDE Standards Task Group – Zhang Minghua

Overview of CGI activities – Harvey Thorleifson

CGI is the International Union of Geological Sciences (IUGS) Commission for the Management and Application of Geoscience Information. The goal of CGI is to foster the interoperability and exchange of geoscience information, through community leadership, collaboration, education, as well as development and promotion of standards and best practices.



On 23 June 2021, from 10 AM to Noon UTC, CGI hosted a webinar attended by 91 experts from around the world.



CGI Webinar participants

Harvey Thorleifson, Director of the Minnesota Geological Survey in the USA and CGI Chair, indicated that the objective of the webinar was to share updates on the work that CGI is carrying out with partners, and to hear discussion. He thanked partners who were mentioned in the six presentations that followed. Discussion and feedback indicated that the webinar was a great success. Videos are available on [YouTube](https://www.youtube.com/watch?v=...).

Please visit the CGI web site at <https://cgi-iugs.org/>, and if you aren't a member, please click the **Join Now** button!

OGC Geoscience Domain – Mickaël Beaufils



Mickaël Beaufils of BRGM spoke as Chair of the Open Geospatial Consortium (OGC) Geoscience Domain Working Group (DWG), that partners with CGI and many other organizations in making spatial information findable, accessible, interoperable, and reusable (FAIR).

For findability, he cited persistent identifiers (PIDs), rich metadata, indexed data repositories, and PIDs in metadata; for accessibility, he cited standard communication protocols, open and free protocols, authentication where necessary, and metadata always available; for interoperability, he cited vocabularies, FAIR vocabularies, and linked metadata; and for reusability, he cited multiple attribute metadata, usage licenses, provenance, and community standards.

Key DWG goals are: interoperability, data definitions, formats, and services for publishing, search, and exchange; thematic/semantic coherence with related domains; interfaces with city/infrastructure and risk domains; showing value added by interoperability; best practices; and standards for spatial and temporal features, metadata, and other information.

Emphasis has been on the BhML borehole conceptual model, as well as (Geo)JSON(-LD), 3D-4D, BIM, geotechnics, ISO19156 for observations and measurements, OGC SensorThings API, OGC-API features, and SMART Data Loader.

The DWG has been active since being conceived at Dublin in 2016; the next meeting is in September.

GeoSciML – Éric Boisvert

Éric Boisvert of the Geological Survey of Canada spoke on behalf of the OGC Standard Working Group (SWG) for GeoSciML (Geoscience Markup Language), an XML- and GML-based machine-readable format for geological maps that has been stable for almost two decades since emanating from work such as XMML, NADM, and G-XML. Although V1 in 2005 was monolithic, V2 in 2009 had properties grouped, with boreholes and fossils added.



V3 in 2012 was packaged as 13 schemas, with themes added, OGC versions of some dependencies, custom elements dropped, external vocabularies, and explicit nil required. V4 in 2015, packaged as basic and extended, became OGC V4.1 in 2017, with reversion to optional properties, profiling, and Portrayal (Lite) integrated. The scope covers geology, including boreholes and specimens, with limited applicability to specialized domains, or symbology.

GeoSciML Lite, based on simple features and thus comparable to a SHP file, is widely compatible, although with limited expressiveness, and thus limited to simple applications. Extensions include GroundwaterML2, EarthResourceML, and geotechnics for civil engineering. OGC, and thus GeoSciML, is transitioning from XML to JSON (OGC API); future work will address knowledge encoding such as RDF and ontologies.

EarthResourceML – Michael Sexton



Michael Sexton of Geoscience Australia provided an overview of EarthResourceML (ERML), a markup language for the delivery of mineral occurrence, deposit, mining, and resource

data that was initiated in 2006 as a collaboration between Australian state and federal agencies. The name was adopted and a CGI working group formed in 2010.

ERML is an extension of GeoSciML, with a mineral occurrence being a subtype of Geologic Feature; the mining feature types of ERML are not part of GeoSciML.

The earth resource concept encompasses geology, economics, environment, and society. Through international collaboration, ERML has a minimum set of features and attributes for describing and communicating information about earth resources in a standardized way. Nineteen CGI vocabularies include mine status, commodity, and reserves; data delivery and implementation are well accommodated.

ERML v3, which is near completion, has an improved resources model to represent resources as reported, addition of processing plants, and removal of mandatory attributes and relationships. ERML v3 will have GeoServer feature templating and a smart data loader, with JSON outputs.

On September 9, 2021, at 10 UTC, we will hold an informal technical seminar on geoscience ontology and knowledge graph, to be attended by members and friends of relevant working groups of CGI and partner organisations. If you wish to attend, please contact the CGI Chair at thorleif@umn.edu.

CGI Geoscience Vocabularies – Mark Rattenbury

Mark Rattenbury of GNS Science in New Zealand, Te Pū Ao, spoke for the CGI Geoscience Terminology Working Group, which has been separate from the GeoSciML group since 2014, and which maintain lists of terms for earth features, properties, quantities, techniques, and processes that are collaboratively agreed upon and are widely accessible, including to machines. The terminology conveys unambiguous meaning, offers translations, reduces confusion, utilizes natural hierarchies, and supports data models.



The Working Group supports interoperability, develops and enables vocabularies, has 30 members in 16 countries, collaborates with partner groups, has had six face-to-face meetings, and has had much online interaction. A formal vocabulary development and adoption process is used. The vocabularies are somewhat FAIR, but there are many competing vocabularies. Ideally, CGI vocabularies will be better known and more widely used, because they are international, they support data models, and they are supported through enduring governance. Working Group priorities are to publish remaining vocabularies identified by working groups, update published vocabularies, add multilingual terms, and promote application of CGI vocabularies, especially in academic and industry contexts.

Challenges include aligning to international projects, finding time to make progress, and pandemic-related restrictions on travel.

Loop GeoScience Ontology – Boyan Brodaric and Stephen Richard

Boyan Brodaric of the Geological Survey of Canada, and Stephen Richard of the USGIN Foundation in USA, reported on GeoScience Ontology (GSO) research being conducted in association with Loop, an integrated and interoperable platform enabling 3D stochastic geological modeling. Knowledge management at present is a very active topic, due to the activities of Google, Amazon, Microsoft, and others. The Google Knowledge Graph initially released in 2016, the basis for their infobox, is built on a schema.org ontology, and a query API such as SPARQL.



The GSO incorporates a framework with entities, relations, and rules, as well as an OWL2 language. The GSO builds on NADM, GeoSciML, and DOLCE Rocks, and work in several ongoing projects. In association with GeoSciML, the activity includes, for example, an OWL conceptual schema and re-conceptualization, UML logical schema, XSD application schema, and XML instance documents.

Development will have to be broad and deep, modular, and standalone. The modular organization will be a layercake ontology, from common, to geology, to modules. Context dependency is an important concept, and relations include fundamental, spatial, temporal, and geological. Next steps include public feedback, knowledge repository testing, and integration with Loop 3D modeling code.

DDE Standards Task Group – Zhang Minghua



Zhang Minghua of the China Geological Survey, Co-Secretary General of CGI, described the work of the Deep-time Digital Earth Standards Task Group (DDE-STG). DDE is an IUGS big science

program focused on data-driven discovery regarding the evolution of life, matter, geography, and climate. DDE will be an international consortium that will link databases, make data FAIR, unify earth science, and promote research.

The kickoff meeting at Beijing in February 2019 was attended by representatives of 40 geoscience organizations, and founding organizations signed an accord. The DDE roadmap encompasses the 2019 accord, formal launch in 2020, midterm reporting in 2024, and final reporting in 2028. For the geoscience knowledge system, DDE has adopted CGI terminology, and knowledge system graphs have been in development since 2019.

The DDE-STG, now consisting of 28 leading professionals who held their first in-person meeting in January 2020, will now work in close collaboration with CGI and partners such as CODATA, to coordinate standards identification and development, survey stakeholders, develop the metadata standard based on ISO19115-2014(E), oversee review of the semantic knowledge system, provide advice to DDE working groups, provide training, and maintain the portfolio of standards.