The CGI: Advancing International Geoscience Data Interoperability

John Broome

- CGI Council
- Earth Sciences Sector, Natural Resources Canada

IUGS

CGI
Commission for the Management and Application of Geoscience Information

Outline

1) Geological Data Interoperability
2) The CGI and its goals
3) The CGI Interoperability Working Group
4) GeoSciML
5) Next steps
6) Lessons learned

IUGS

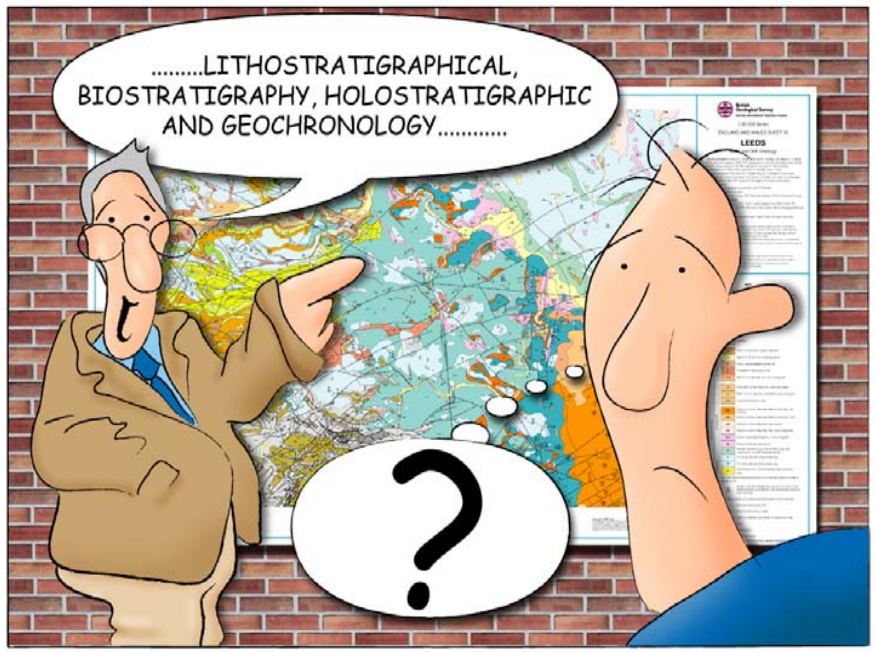
Natural Resources Canada | Ressources naturelles Canada

CGI
Commission for the Management and Application of Geoscience Information

Geology in Society

Access to geological data is often crucial in protecting human life, health and assets, and sustaining our environment and resources, but ….

Geological data are often not used because they are not accessible in a standard form that can be understood by non-specialists.



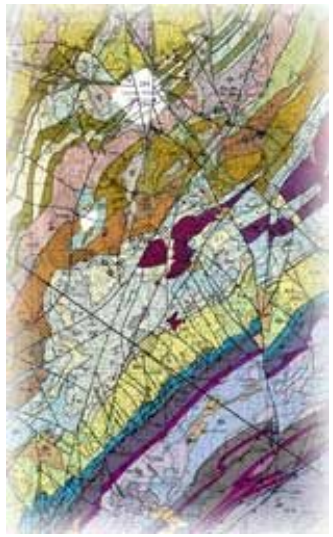
IUGS

Natural Resources Canada | Ressources naturelles Canada

CGI
Commission for the Management and Application of Geoscience Inform...

The Geology Interoperability Problem

- Geology is plagued by a lack of internationally-accepted data standards.
- Geological map data are a complex combination of information describing lithology, age, physical structure, chemical properties, and genesis.
- Geoscience varies with regional or national context.



Result: Geological data and information cannot be easily exchanged, integrated, and understood.

IUGS

Natural Resources Canada | Ressources naturelles Canada

CGI
Commission for the Management and Application of Geoscience Information

The CGI

In 2004, the International Union of Geological Sciences (IUGS) created the "Commission for the Management and Application of Geoscience Information"

- The CGI has 150 members in 50 nations (2006)
- The CGI is operated by a small but enthusiastic international Council.
- Visit the CGI website: http://www.bgs.ac.uk/cgi_web/

IUGS

Natural Resources Canada | Ressources naturelles Canada

CGI

Commission for the Management and Application of Geoscience Information

CGI Goals

1) Provide the means for transferring knowledge on geoscience information and systems.

2) Stimulate international dissemination of best practice in geoscience information.

3) Stimulate and support initiatives which are developing standards.

4) Establish and occupy an accepted position in the international geoscience information community and represent IUGS on geoscience information matters.

CGI
Commission for the Management and Application of Geoscience Information

CGI Initiatives

- Interoperability Working Group
- Multi-lingual Thesaurus Working Group
- Metadata standards
- African Outreach Workshop on Geoscience Information
- Data Management “Super-session” at International Geological Congress (2008)
- Regional Working Groups

IUGS

Natural Resources Canada | Ressources naturelles Canada

CGI
Commission for the Management and Application of Geoscience Information

CGI Interoperability Working Group

- **Interoperability Working Group**
- Multi-lingual Thesaurus Working Group
- Metadata standards
- African Outreach Workshop on Geoscience Information
- Data Management “Super-session” at International Geological Congress (2008)
- Regional Working Groups

IUGS

Natural Resources Canada | Ressources naturelles Canada

CGI
Commission for the Management and Application of Geoscience Information

Interoperability & CGI

1) Provide the means for transferring knowledge on geoscience information and systems.

2) Stimulate international dissemination of best practice in geoscience information.

3) Stimulate and support initiatives which are developing standards.

4) Establish and occupy an accepted position in the international geoscience information community and represent IUGS on geoscience information matters.

IUGS

Natural Resources Canada | Ressources naturelles Canada

CGI
Commission for the Management and Application of Geoscience Information

CGI Approach to Interoperability

Organizations have invested in customized data management systems that support local requirements, the implementation of new standards is intrusive and costly.

Recognizing this problem, the CGI approach to interoperability is based on development of core standards integrated with standard XML/GML interchange formats.

This non-intrusive approach allows retention of local systems but achieves interoperability through translation of data to a standard CGI interchange format (GeoSciML) that is compatible with OGC-standard services.

CGI
Commission for the Management and Application of Geoscience Information

Interoperability 101

Recipe for Data/Information Interoperability:

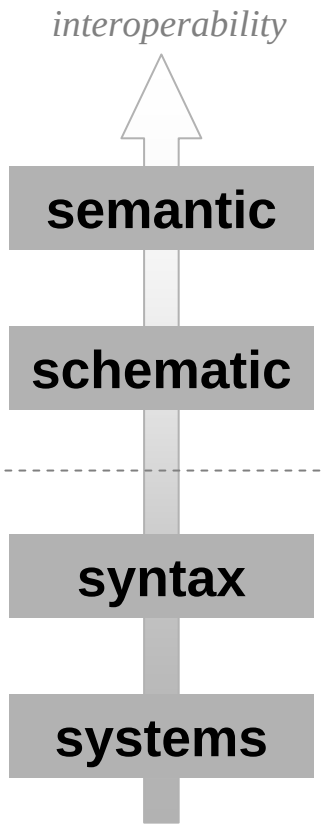
1) Context: Conceptual model – classification/ontology

2) Content: Terminology Standards – dictionaries/thesauri

3) Structure: schematic model - schema

OGC & Interoperability

Commission for the Management and Application of Geoscience Information



Ontology/Vocabulary (data content - information)

GeoSciML (data structure)

GML (data language)

WFS, WMS, WCS, … (data systems)

Geoscience

OGC Standards

IUGS

Natural Resources Canada | Ressources naturelles Canada

CGI
Commission for the Management and Application of Geoscience Information

Interoperability WG Objectives

1. Develop a conceptual model of geological information drawing on existing data models (NADM).
2. Implement an agreed subset of this model in an agreed schema language.
3. Identify areas that require standardised classifications in order to enable interchange.
4. Implement an XML/GML encoding of the model subset
5. Develop testbeds based on OGC standards to illustrate the potential of the data model/schema for interchange.

IUGS

Natural Resources Canada | Ressources naturelles Canada

CGI
Commission for the Management and Application of Geoscience Information

What is GeoSciML?

GeoSciML is an IUGS standard being developed by the Interoperability Working Group to enable geologic map data interoperability.

Key international geologic agencies are involved (U.S.A., Canada, U.K., France, Sweden, Australia, etc.),

GeoSciML development activities:

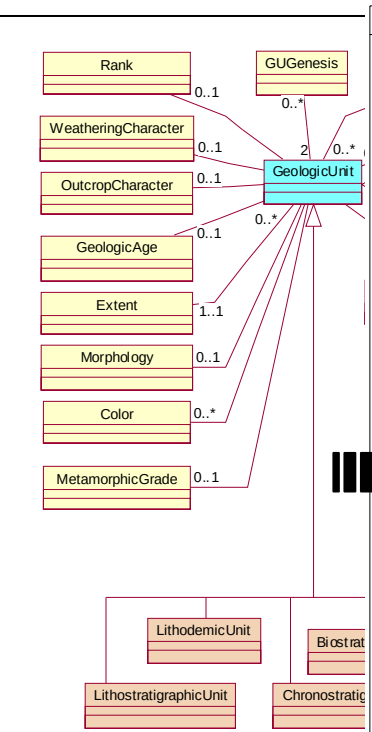
(1) Standard data interchange structure (schema)

(2) Standard data content,

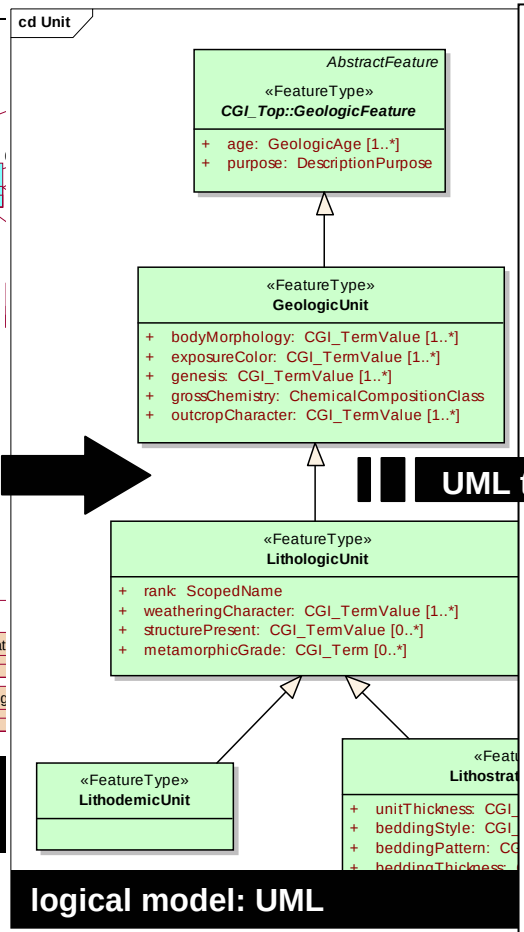
(3) A series of testbeds, in which geological agencies interchange information and test developments.

Conceptual⇗Logical⇗Physical Design

CGI
Commission for the Management and Application of Geoscience Information



conceptual model: no GML



logical model: UML

UML to GML

```
<LithodemicUnit gml:id="GSV53">
  <gml:description>Granite, syenite, volcanogenic sandstone,
   conglomerate, minor trachyte lava</gml:description>
  <gml:name>Mount Leinster Igneous Complex</gml:name>
  <purpose>typicalNorm</purpose>
  <age>
    <GeologicAge>
      <value>
        <CGI_TermRange>
          <lower>
            <CGI_TermValue>
              <value codeSpace="http://www.iugs-
                  cgi.org/geologicAgeVocabulary">Triassic</value>
            </CGI_TermValue>
          </lower>
          <upper>
            <CGI_TermValue>
              <value codeSpace="http://www.iugs-
                  cgi.org/geologicAgeVocabulary">Triassic</value>
            </CGI_TermValue>
          </upper>
        </CGI_TermRange>
      </value>
      <event>
        <CGI_TermValue>
          <value codeSpace="http://www.iugs-
            cgi.org/geologicAgeEventVocabulary">intrusion</value>
        </CGI_TermValue>
      </event>
    </GeologicAge>
  </age>
  <age>
```

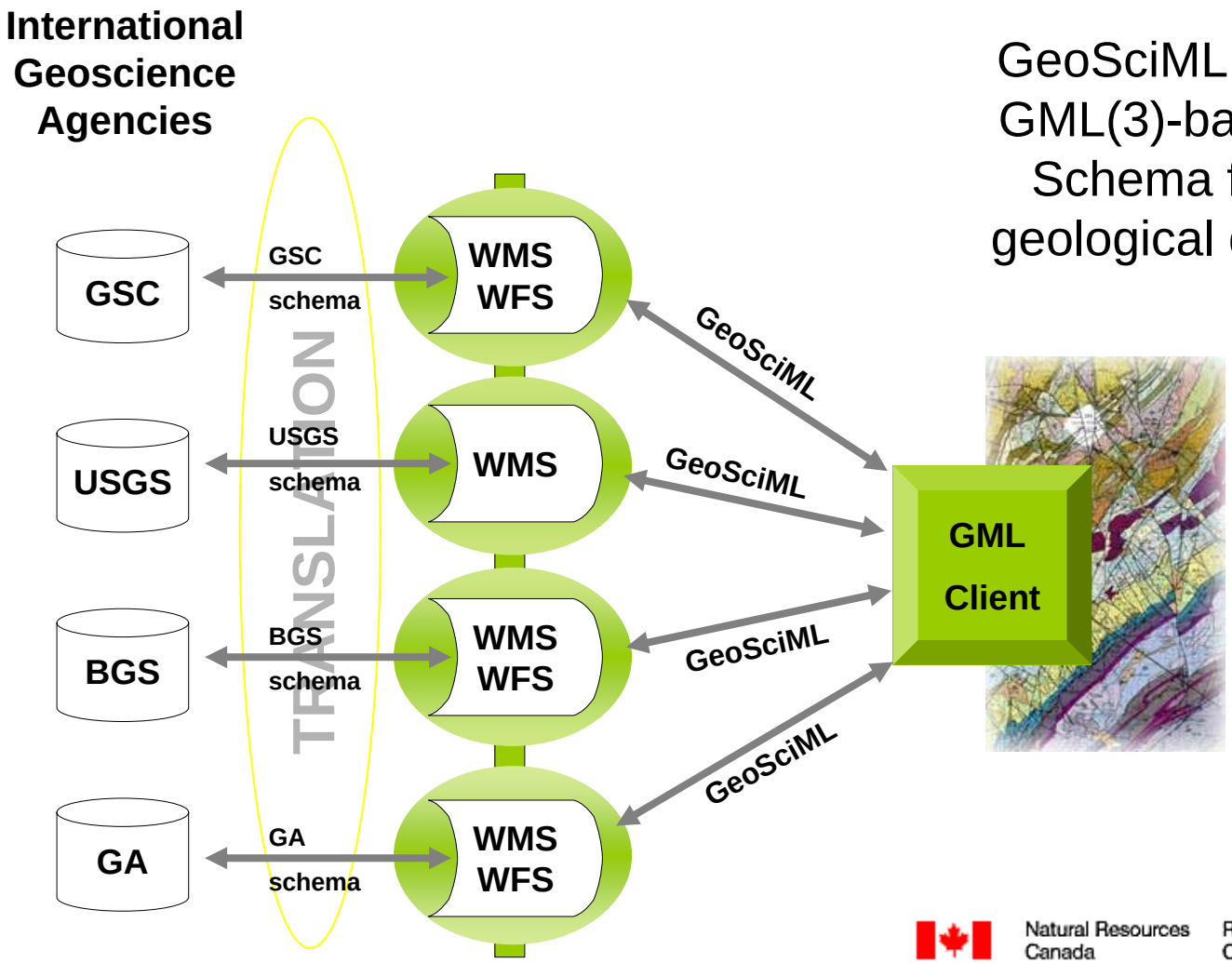
physical model: GML-XML

IUGS

Natural Resources Canada | Ressources naturelles Canada

CGI
Commission for the Management and Application of Geoscience Information

How does GeoSciML work?



GeoSciML is a GML(3)-based Schema for geological data

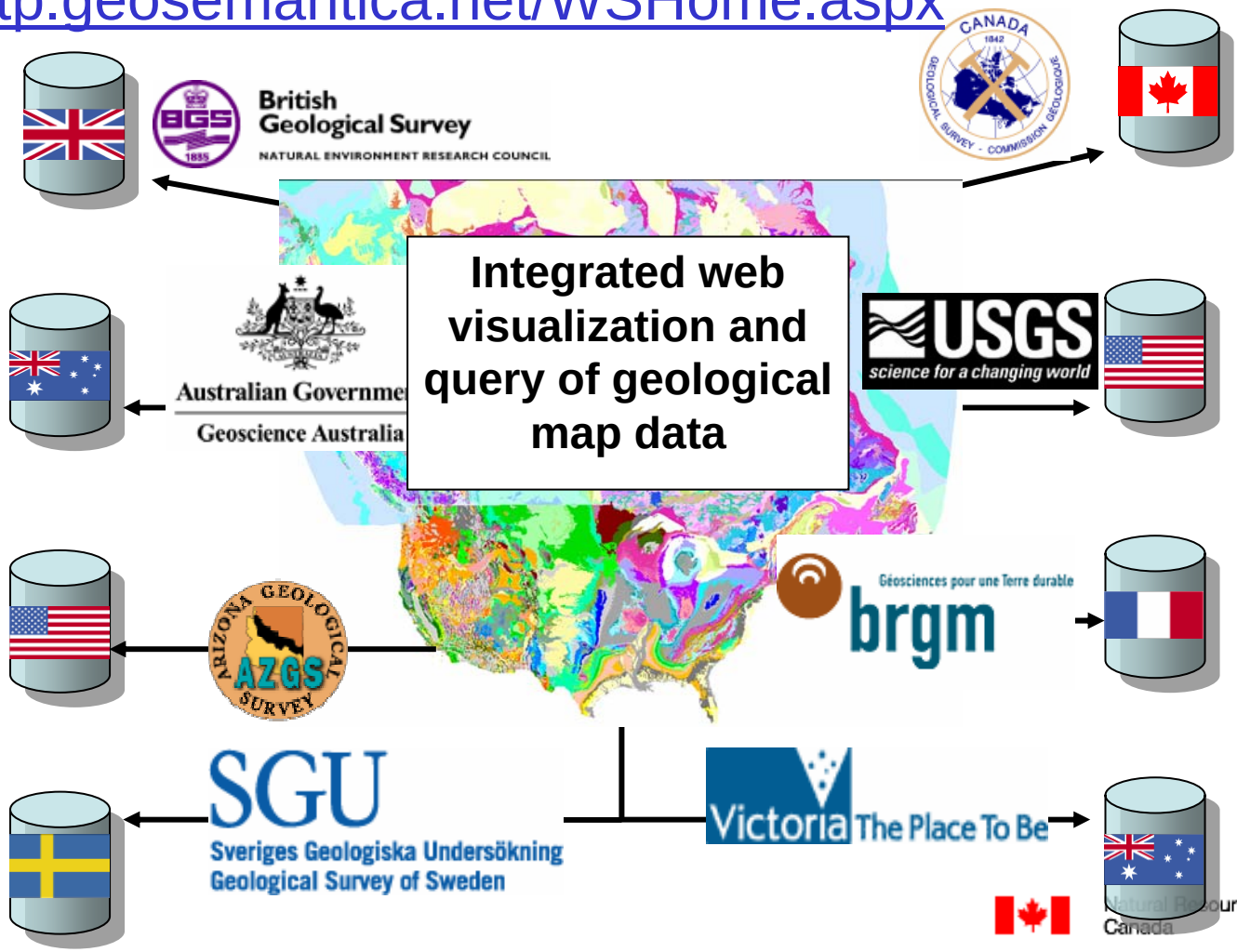
IUGS

Natural Resources Canada | Ressources naturelles Canada

GeoSciML Testbeds

Commission for the Management and Application of Geoscience Information

http://ctp.geosemantica.net/WSHome.aspx



IUGS

Natural Resources Canada
Ressources naturelles Canada

GeoSciML: Next steps

1) Separate development and operational streams.

2) Support operational implementation

3) Develop detailed documentation and user guides.

4) Investigate OGC certification of GeoSciML

CGI
Commission for the Management and Application of Geoscience Information

Lessons learned …

- Standards are easily developed but often not adopted.
- Standards must address a business need or solve a problem.
- Autonomous organizations are most likely to adopt standards if they are non-intrusive.
- Standards must be fully documented & supported with compliant tools.
- Obtaining resources for long-term support & maintenance is difficult.
- Design standards as extensions of core standards (ISO, OGC, etc.); implement using open source tools.
- New technologies and standards are typically not fully tested and will have problems.