DOMAIN REFERENCE ONTOLOGIES VS. DOMAIN ONTOLOGIES: WHAT'S THE DIFFERENCE?
Lessons from the Water Domain

Torsten Hahmann, UMaine, NCGIA
Shirly Stephen, UMaine
Boyan Brodaric, GSC
1. **Motivation**
   - The Water Domain
   - State-Of-the-Art in Hydro Ontologies
   - The Role of a Domain Reference Ontology

2. **Use Case**
   - Objective & Approach
   - Hydro Foundational Ontology (HyFO)
   - Groundwater Markup Language (GWML2)
   - Results

3. **Takeaway**
THE WATER DOMAIN

Water Features (surface, subsurface, atmospheric) & their location and spatial configuration (rock bodies, surfaces, depressions) & Flow between them
<table>
<thead>
<tr>
<th>Generality</th>
<th>Scope &amp; Level of Detail</th>
<th>Format</th>
<th>Foundational Grounding</th>
</tr>
</thead>
<tbody>
<tr>
<td>WaterML2.0</td>
<td>Application-specific standard (for integrating water observations)</td>
<td>XML</td>
<td></td>
</tr>
<tr>
<td>RiverML</td>
<td>Application-specific standard</td>
<td>XML, builds on WaterML</td>
<td></td>
</tr>
<tr>
<td>SWEET</td>
<td>Hydro domain ontology</td>
<td>OWL</td>
<td></td>
</tr>
<tr>
<td>INSPIRE</td>
<td>Hydro domain ontology</td>
<td>UML (terms fully defined only in a glossary)</td>
<td></td>
</tr>
<tr>
<td>HY_Features</td>
<td>Hydro domain ontology</td>
<td>UML</td>
<td></td>
</tr>
<tr>
<td>hydrOntology</td>
<td>Application ontology (for interoperability among data sources of IGN-E)</td>
<td>XML</td>
<td></td>
</tr>
<tr>
<td>Surface Water Ontology (NHD)</td>
<td>Hydro domain ontology</td>
<td>OWL axioms RDF triples with a SPARQL endpoint</td>
<td>BFO</td>
</tr>
<tr>
<td>Surface Water Schema</td>
<td>Hydro domain ontology</td>
<td>OWL DL</td>
<td></td>
</tr>
<tr>
<td>GWML</td>
<td>Application-specific (built using GeoSciML and O&amp;M schemas)</td>
<td>UML conceptual schema, XML physical schema (terms fully defined only in a glossary)</td>
<td></td>
</tr>
</tbody>
</table>
Existing ontologies for the water domain—and for most scientific domains—rely on conceptual models (UML diagrams) and lightweight ontologies (RDF or subclass hierarchies specified in OWL).

Semantic nuances in usage of terms is not formally expressable in these languages.
SEMANTIC NUANCES: WHAT IS AN AQUIFER?

GWML2

A hydrogeological unit that potentially stores groundwater.

INSPIRE

An aquifer is a rock body, but does not capture the notion of voids in it.

NWIS (USGS)

Geological formation or structure that supplies water to wells/springs.

• These subtle differences are currently not formalized
  ➔ typically only available in narrative form
  ➔ prevents use of automated integration (e.g., ontology alignment) techniques

• How do we reconcile these semantic differences?
Vision: Leveraging a domain reference ontology (here: the hydro reference ontology HyFO) to analyze and revise geoscience knowledge representations (here: hydro data models) to increase semantic interoperability among them.

Exhibits many characteristics of foundational ontologies, but more specific: they are foundational for their domain.
What does a Domain Reference Ontology look like?

- Exhibits many characteristics of foundational ontologies: *foundational for their domain*
  1. Foundationally grounded
  2. Broad coverage on the highest level in the domain: focuses on the *key concepts and relations* in the domain; but does not aim to capture the domain comprehensively
  3. Specified in a highly expressive and fully machine-interpretable ontology language
- Provides “neutral” language to express semantic differences; Purpose is not to directly define the scientific terms (e.g. aquifer), but ontological helper concepts and relations
Developing Domain Reference Ontologies

Significant time needs to be invested into their careful development

- Identify key concepts and relations in the domain
- Need to be developed by ontologists, though with input from domain experts
- Shortcuts just create another domain ontology, but not a reference for the domain
- Payoff: Once finished, can be used to analyze and improve multiple domain and application ontologies and increase their semantic interoperability
We tested whether a domain reference ontology can be successfully used to help clarify the semantics and improve their formalization of existing domain ontologies.

- Role of Domain Reference Ontology: provide ontological guidance and formal language to analyze other domain ontologies
- Domain Reference Ontology: **Hydro Foundational Ontology (HyFO)**
- Analyzed Domain Ontology: **Groundwater Markup Language, v2 (GWML2)**, specified as UML diagrams with English explanations

**Step 1:** Formalize all aspects of GWML2 semantics in first-order logic using HyFO terminology

**Step 2:** Identify and resolve semantic ambiguities and inconsistencies in GWML2

**Step 3:** Verify that the resulting GWML2-FOL is a logically consistent extension of HyFO
**Hydro Foundational Ontology (HyFO)**

- Under development (Hahmann, Brodaric, Gruninger) since 2011
- Progress presented in a number of papers (voids, containment relations, constitution/dependency, water features)

**Container Object**
The solid object where water can be located; e.g., the channel of a river, the rock body in an aquifer.

**Void**
Space(s) in the container that can be filled with water; e.g., pores in an aquifer, depression in a channel.

**Matter**
Material that constitutes a container or water object; e.g., solid rock matter, water matter.

**Water Object**
The liquid object located in the container/void; e.g. the spatio-temporal object that encompasses all the water in an aquifer.

**Diagram**

- Container Solid Body (CSB) hosts Hydro Void (HV) contains Matter (M) constitutes Water Body (WB) contains
HYFO RELATIONS

- Spatial relations
  - Spatial part
  - Spatial overlap
  - Different kinds of spatial connection

- Physical relations:
  - Material-spatial dependence
  - Containment
  - Hosting a void
  - Different kinds of physical parts
Groundwater Markup Language (GWML2)


- UML schemas with narrative semantic descriptions.
- Extends GeoSciML (a markup language for geosciences) to describe hydrogeological concepts.
- Extends OGC/ISO Observations and Measurements schema to describe concepts and properties relevant to flow of groundwater.
**RESULTS 1: FORMALIZATION OF GWML2 IN FIRST-ORDER LOGIC (GWML2-FOL) WITH CLARIFIED AND REFINED SEMANTICS USING HYFO’S TERMINOLOGY**

A HydroVoid that is hosted by the Solid Body that is a submaterial of a Hydrogeo Unit or a Well.

\[ \text{HGV}(x) \equiv \text{HV}(x) \land \exists y,z \ [\text{SB}(y) \land \text{hosts-v}(y,x) \land \text{submat}(y,z) \land (\text{HGU}(z) \lor W(z))] \]

Water Body that consists of all water in an Aquifer and resides in a Hydrogeo Unit.

\[ \text{AWB}(x) \equiv \text{WB}(x) \land \exists y,v \ [\text{A}(y) \land \forall z \ [(\text{intra-const}(z,y) \land \text{WM}(z) \rightarrow \text{submat}(z,x)] \land \text{HGV}(v) \land P(r(x),r(v))] \]

Specialization of Matter

\[ \text{EM}(x) \rightarrow \text{M}(x) \]

A NAPO that is constituted of non-fluid Earth Material.

\[ \text{SB}(x) \equiv \text{NAPO}(x) \land \exists y \ [\text{M}(y) \land \text{intra-const}(y,x) \land \forall z \ [\text{intra-const}(z,x) \rightarrow \neg \text{FM}(z)]] \]

A specialization of Hydro Rock Body.

\[ \text{HGU}(x) \rightarrow \text{HRB}(x) \]
**Results 1:** Formalization of GWML2 in first-order logic (GWML2-FOL) with clarified and refined semantics using HYFO’s terminology

A HydroVoid that is hosted by the Solid Body that is a submaterial of a Hydrogeo Unit or a Well.

Water Body that consists of all water in an Aquifer and resides in a Hydrogeo Unit.

**Full Definition (GWML2 Aquifer):**

\[ A(x) \equiv AU(x) \land \\
\neg \exists y[AU(y) \land x \neq y \land \text{submaterial}(y, x)] \land \\
\exists s, a[\text{CSB}(s) \land \text{submaterial}(s, x) \land \text{AWB}(a) \land \\
\text{submaterial}(a, x) \land \neg \text{ZEX}(\text{con-voidspace}(s)) \land \\
\forall w[\text{intragranular-constituent}(w, a) \\
\rightarrow P(r(w), \text{con-voidspace}(s)))]] \\

\]
Results 2: Layering (modularization) of GWML2 concepts into ones that are groundwater-specific or more general

Groundwater module encapsulates GWML2’s unique focus and terminology but is also a (consistent) extension of HyFO for alignment with other hydro ontologies in the future.

[1] Foundational Concepts from DOLCE

[2] Solid and Fluid (Geo-)Physics


Introduces several new classes to GWML2 at appropriate layer to separate groundwater-specific concepts from general water concepts.

- Number of revisions to GWML2 terminology (disambiguation)
HydroRockBody captures physical things that are partly solid, partly fluid, and partly space ➔ captures idea of a “water feature” (aquifer, river, lake)
**Summary of Results from Use Case**

1. **Formalization of GWML2 in first-order logic (GWML2-FOL) with clarified and refined semantics using HyFO’s terminology**
   - Logical extension of HyFO
   - Available in Common Logic format from COLORE
   - Tested consistency using theorem provers

2. **Layering (modularization) of GWML2 concepts into groundwater-specific concepts and into more general ones**
   - Modules that encapsulates GWML2 unique focus and terminology but is consistent with HyFO and thus other hydro ontologies aligned to it in future
   - Completion of the hierarchy

3. **Number of revisions to GWML2**
   - Concept disambiguation
   - Introduction of several new classes to GWML2 to separate groundwater-specific concepts from general water concepts

4. **Successfully tested HyFO for its suitability as a domain reference ontology for the water domain**
   - Proposed extending HyFO with a fifth core concept
**Takeaway, Next Steps & Open Questions**

- **Domain Reference Ontologies are not just another standard**
  - Different purpose that domain ontologies: Analysis, refinement & integration of domain ontologies
  - Not intended for direct use in applications

- **Apply same methodology to another hydro ontology to actually test/prove that the resulting ontologies are semantically interoperable: HY_Features?**

- **How transferable is it to other domains?**
  - Nothing really specific to the water domain
  - Requires identifying well-scoped domains

- **General guidelines on developing domain reference ontologies?**
  - Requires broad skills and knowledge: formal logic, philosophical principles of upper ontologies and domain expertise
  - Seem to be the hardest ones to develop: really require close collaboration between ontologists and domain experts
  - There is no free lunch here 😞
Selected References


