Ontology Reference Summit 2016: Semantic Integration in Engineering

Semantic Alignment of the Groundwater Markup Language with the Reference Hydro Ontology HyFO

Presenters: Shirly Stephen, Torsten Hahmann

Spatial Information Science and Engineering

University of Maine, Orono, ME 04469

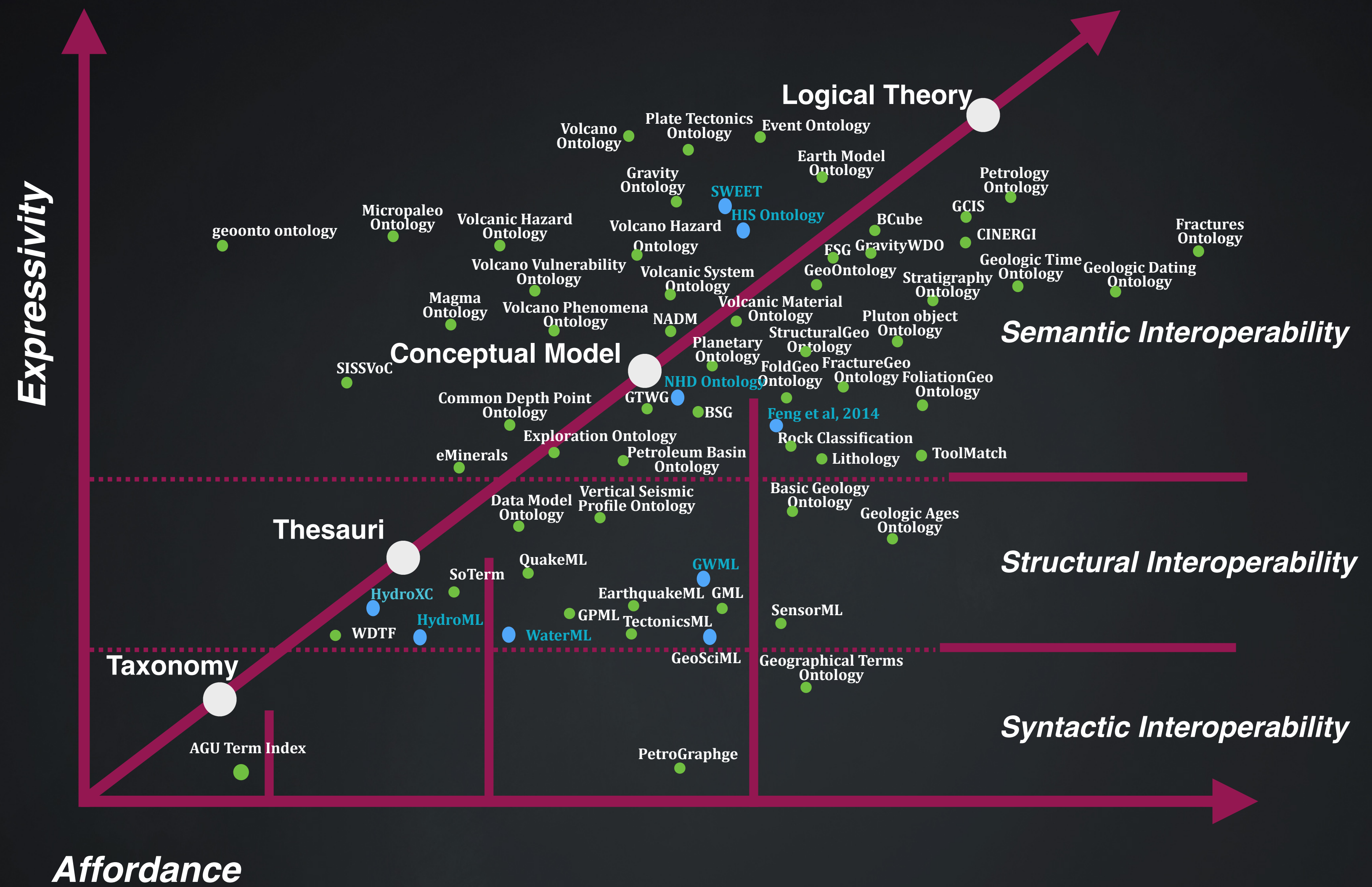
Joint work with Boyan Brodaric

Geological Survey of Canada

March 30th, 2016

THE UNIVERSITY OF MAINE

Geoscience Ontologies in the Semantic Spectrum



Adopted from: *An overview of semantic models in the geosciences – Brandon Whitehead*

Introduction

- Hydrologic information is described in multiple schemas, markup-languages and ontologies.
- Semantic integration is a challenge because these standards are, for the most part, heterogenous as they:
  - are *mostly fragmented and disconnected, describing either surface or groundwater.*
  - *lack foundational grounding.*
  - *use the same or similar terms but with differences in semantics.*
  - *are described using different formal (or non-formal) languages.*
- We propose to achieve semantic interoperability via ontology integration.

	Generality	Ontological Richness	Formality and Expressivity	Foundational Grounding
WaterML2.0	*Application-specific standard (for integrating water observations)*	*For time-indexed hydrologic observations*	***XML***	✗
RiverML	*Application-specific standard*	*Models geometry/morphology of rivers*	***XML, builds on WaterML***	✗
SWEET	*Hydro domain ontology (not rich in relations aspect)*	*- surface and sub-surface domains* *- approximately 50 concepts with mostly taxonomic relations*	***OWL***	✗
INSPIRE	*Hydro domain ontology*	*- surface and subsurface domain* *- groundwater model based on GeoSciML* *- does not describe the concept of void*	***UML*** *(complete definitions in a glossary and not formalized)*	✗
HY_Features	*Hydro domain ontology*	*- emphasis on surface water features* *- does not contain and hydrologic relations*	***UML***	✗
hydrOntology	*Application ontology (for interoperability among data sources of IGN-E)*	*- 150 classes, 34 object properties, 66 data properties., 256 axioms*	***XML*** *concepts varying from global-continental to regional-local*	✗
Hydro	*Data-driven ontology (based on features in the NOMGEO)*	*- 51 classes* *- very comprehensive, but only relates generic names to feature type*	***OWL Full*** *-*	✗
Surface Water Ontology (NHD)	*Hydro domain ontology*	*- Based on The National Hydrology Dataset in the National Map* *- specifies taxonomic type of hydro features*	***OWL axioms*** ***RDF triples with a SPARQL endpoint***	***BFO***
Surface Water Schema	*Hydro domain ontology*	*Distinguishes between container (solid object) and water body in the surface domain*	***DL axioms and OWL***	✗
GWML	*Application-specific (built using GeoSciML and O&M schemas)*	*- For the exchange of groundwater related information* *- Contains hydrogeo units, voids and wells*	***UML conceptual schema, XML physical schema***	✗

Existing Semantic Integration Techniques

- Existing *ontology matching* and *alignment* techniques find similarities, equivalences and subsumption relations between two (or more) ontologies given that they,
  - *Are syntactically and schematically integrated.*
  - *Are of similar scope.*
  - *Are no more expressive than OWL.*
- (Whereas) semantic integration between existing hydrologic ontologies and schemas additionally requires:
  - *Translation between ontology languages.*
  - *More rigorous specification of the semantics in each ontology.*
- This can currently only be done by manual integration of the ontologies.

THE UNIVERSITY OF MAINE

Semantic data integration using a Reference Ontology

What is an aquifer ?

NWIS – *Geological formation/structure that supplies water to wells and springs*

GWML2 – *An aquifer is a hydrogeological unit that potentially stores ground water*

INSPIRE – *An aquifer is a rock body, but does not capture the notion of voids or water bodies inside it*

A reference ontology is not just another standard, but defines concepts in a level of detail such that other domain ontologies/standards can be expressed using this terminology.

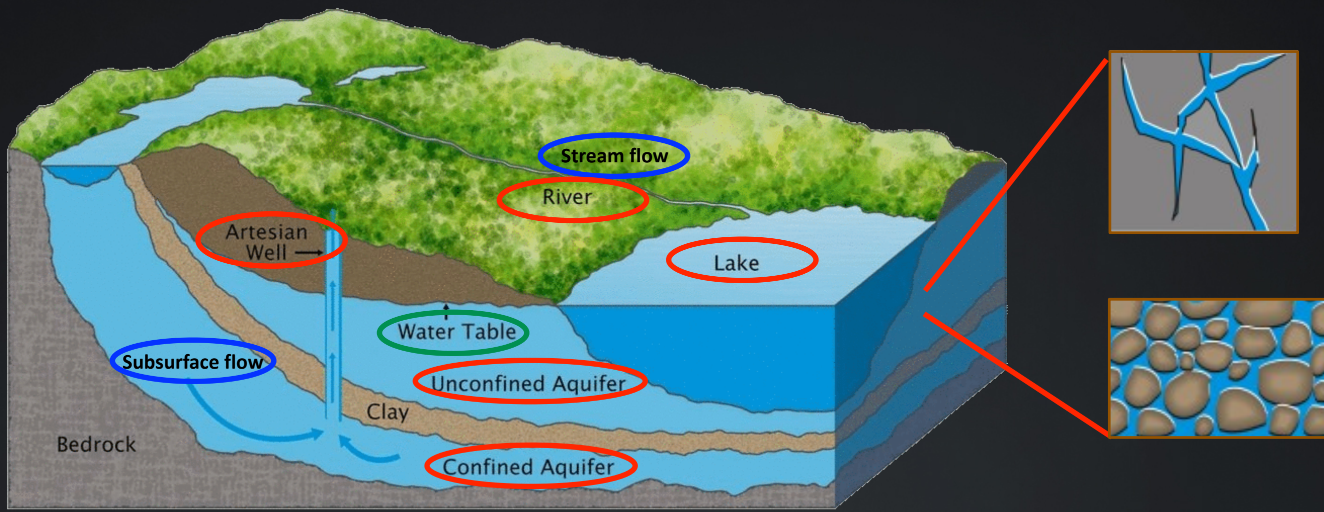
Need for a Reference Ontology for Hydrology Domain

- None of the existing hydro standards are good candidates for a domain reference ontology due to:
  - *insufficient domain coverage*
  - *insufficient level of detail*

- Criteria for a reference ontology:
  - *Foundational grounding ('ontologically sound')*
  - *Broad coverage of the entire hydro domain (both surface and subsurface water storage and flow)*
  - *Detailed, rigorous axiomatization of all semantics in a language that affords automated verification and reasoning*

THE UNIVERSITY OF MAINE

The Hydro Foundational Ontology (HyFO)



1 First Order Logic

- Rigorously axiomatized in FOL
- Logical consistency of the axioms verified to ensure they are free of contradictions

2 Hydro Ontological Square

- Includes a set of four common concepts that are central to surface and subsurface water storage
- Foundational kinds of physical endurants

3 Foundationally grounded in DOLCE upper ontology

- Distinguishes between endurants (that are wholly present at different points in time) from perdurants (that are not present at single point in time)

4 Physical Endurants and Spatial Regions

- Distinguishes between physical objects and its spatial region
- Uses qualitative spatial relations such as connected C, spatial overlap PO, parthood P, superficial contact SC, to express spatial invariants between objects in the domain

5 Voids as entities in their own right and at different granular levels

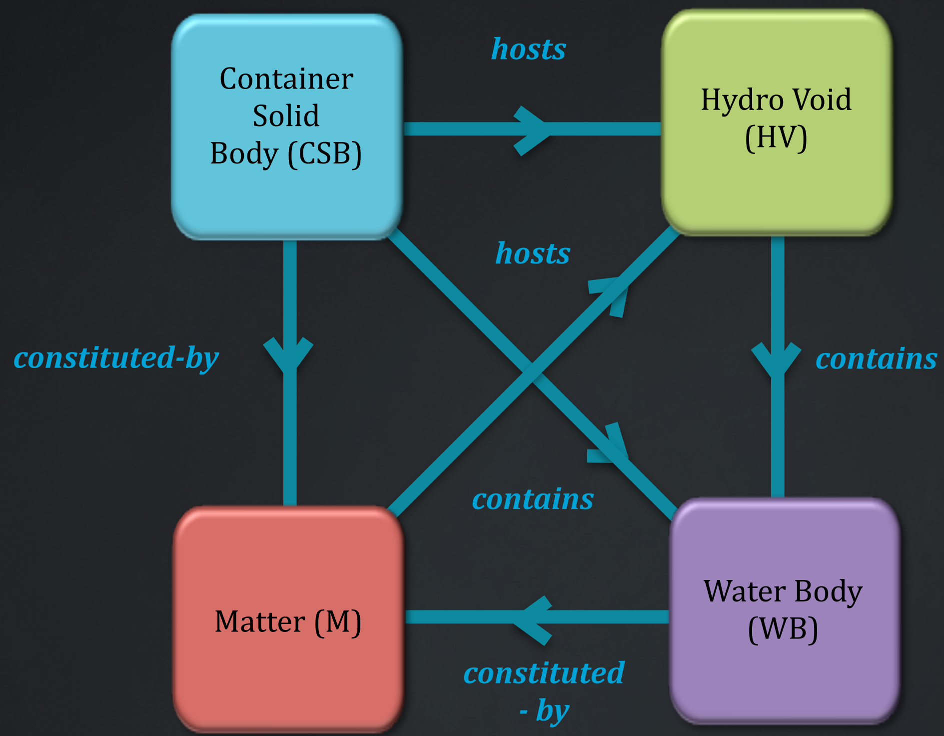
- Connectivity of the host: holes vs. gaps
- Granularity: macroscopic holes in an object vs. microscopic pores in the object's matter
- Openings of the voids: cavity (0), depression (1), tunnel (2)

6 Exhaustive Containment relations at different granular levels

- Detachable and dependent containment based on the material dependence between container and containee
- Intergranular and intragranular constitution depending on the granularity of constituting matter

Emerging Hydro Reference Ontology: Hydro Foundational Ontology (HyFO)

– *based on the Hydro Ontological Square*



REFERENCES:

- *Towards a foundational hydro ontology for water data interoperability - T Hahmann, B Brodaric.*
- *The Void in Hydro Ontology - T Hahmann, B Brodaric.*
- *Voids and material constitution across physical granularities - T Hahmann, B Brodaric.*
- *Kinds of full physical containment - T Hahmann, B Brodaric.*

The Semantic Integration Approach

- Test the viability of Hydro Foundational Ontology (HyFO) *as reference ontology*
  - *Map existing hydro ontologies to HyFO to increase their semantic precision, and integrate them with one another.*

- **Step 1 –** Test its coverage and generality with respect to groundwater concepts:
  - *Coverage: can it represent all relevant groundwater concepts in sufficient detail?*
  - *Generality: is it compatible (i.e., consistent) with existing groundwater ontologies?*

- **Step 2 –** (Future Work): Do the same test for surface water.

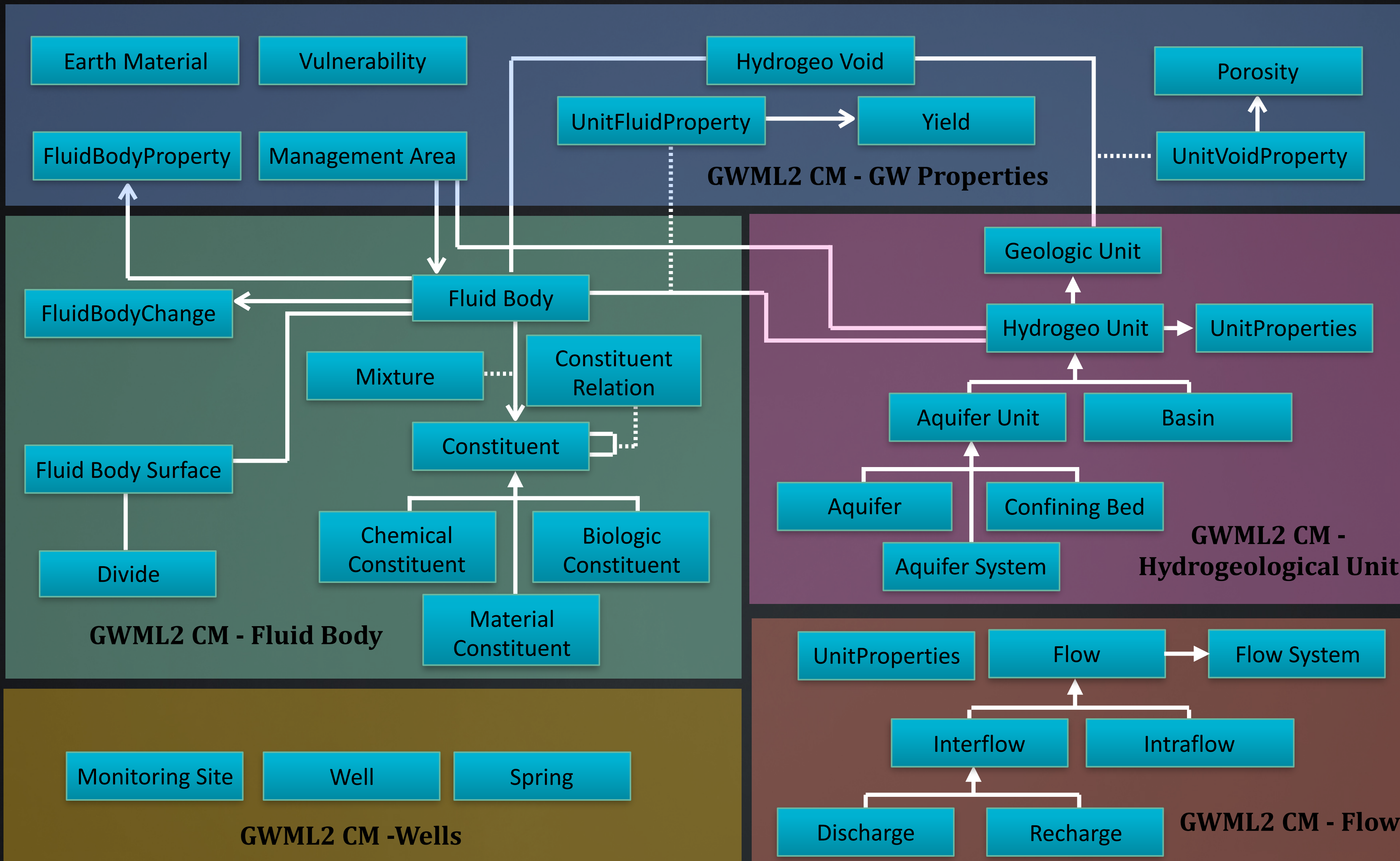
Step 1 – Ground Water Markup Language (GWML2)

GWML2 is a GML (Geographic Markup Language) application and an OpenGIS Consortium (OGC) standard for the exchange of groundwater information.

- Extends Observations and Measurements schema (an OGC/ISO standard) to describe concepts and properties relevant to flow of groundwater.

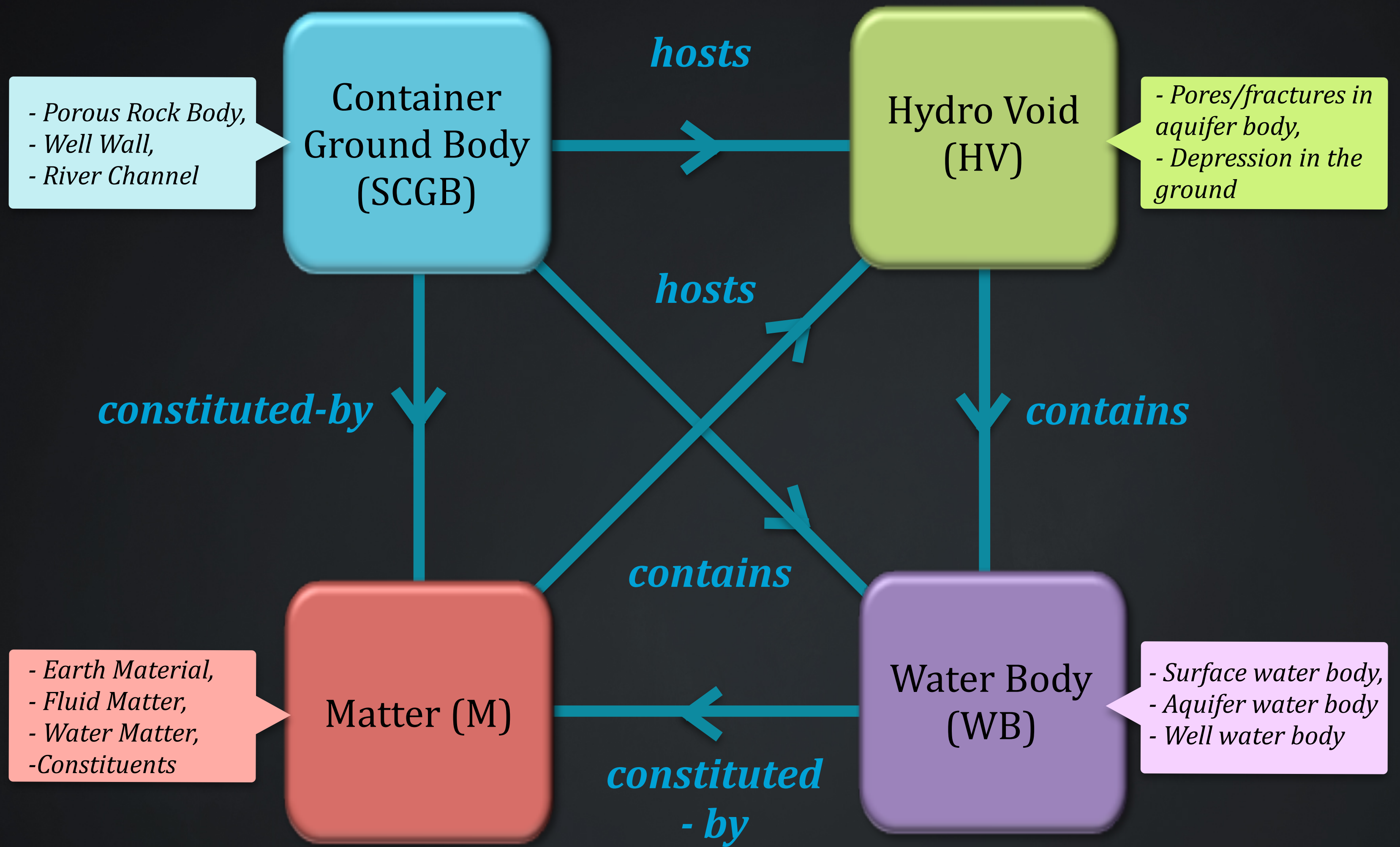
- Extends GeoSciML (a markup language for geosciences) especially *Geologic Unit* and *Earth Material* to describe hydrogeological concepts.

GWML2 – Overview

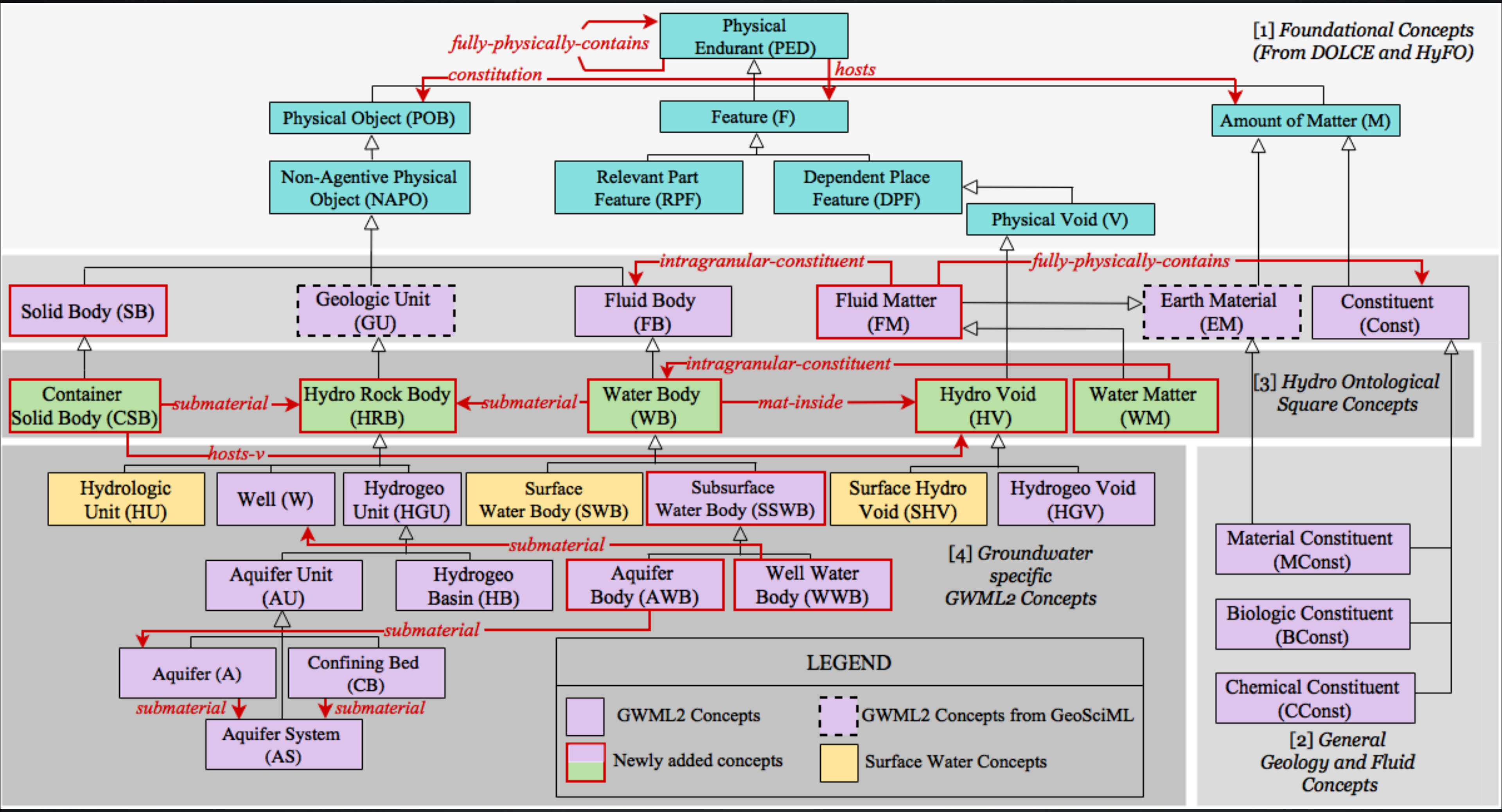


THE UNIVERSITY OF MAINE

GWML2 in HyFO terms



Subclass hierarchy of GWML2, HyFO and DOLCE

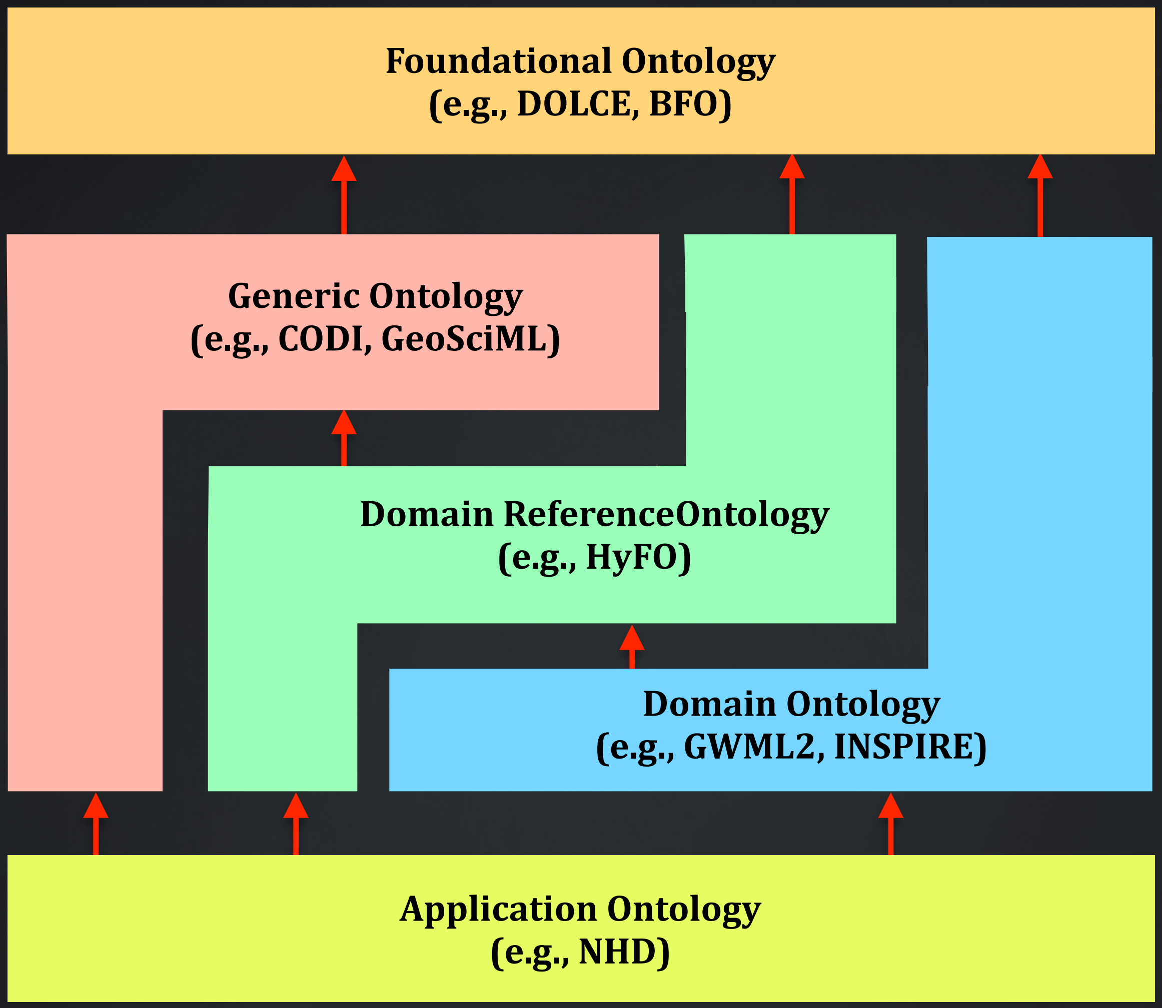


Differentiates GWML2 concepts that are generically applicable to both surface and subsurface water storage from those that are groundwater specific.

General Summary

- We create an axiomatic model of GWML2's core concepts (except flow concepts and properties) as a consistent extension of HyFO and DOLCE.
- Semantic ambiguities and other ontological obstacles that hinder integration of GWML2 with other ontologies are identified and resolved.
- We obtain a stratified subclass hierarchy of GWML2, HyFO and DOLCE.

Ontological Stratification for Geoscience Ontologies



THE UNIVERSITY OF MAINE

Broader Implications of This Work

The science of hydrology would be inherently simple if water were unable to penetrate the earth's surface
– Harold E. Thomas

- This is a first work that demonstrates the suitability of HyFO as a reference ontology for the hydro domain.
  - *Describes GWML2 concepts (except flow) as an extension of HyFO and DOLCE without any logical contradictions.*
- Helps extend HyFO with missing definable concepts that are needed to integrate GWML2.
- Use case of how to effectively utilize formal ontological analysis and rigorous axiomatizations in the development and integration of geoscience standards.

1865 THE UNIVERSITY OF MAINE

THANK YOU

If you have further questions or comments please contact:

torsten.hahmann@maine.edu

shirly.stephen@maine.edu