

Driving the next generation of spatial standards: examples from hydro ontology

Torsten Hahmann

Dept. of Computer Science, University of Toronto

from Nov.: National Center for Geographic Information & Analysis (NCGIA)
School of Computing and Information Science
University of Maine

October 17, 2013

Why bother standardizing geospatial terms?

① Badly needed

- ▶ Basis for data & information integration and exchange in the earth sciences (compare the goals of EarthCube)
- ▶ Supports other disciplines: much information has a spatial component

② Many intuitive conceptualization, concepts, and relations, each of which comes with various interpretations that slightly differ

- ▶ Conceptualizations: raster/grid- vs. vector- vs. graph/network-based, discrete vs. continuous, flat vs. spherical, ...
- ▶ Concepts: *boundary, surface, curve, region, hole, water body, lake, toponyms, ...*
- ▶ Relations: *in contact with, is part of, contains, ...*

The role of standards: reference + implementation guide

Reference that defines a reusable terminology with shared semantics

- **standardized terms**
- **standardized definitions of the terms' meanings**

Expressive formal ontologies can help achieve both:

- **Terminology:** concepts and relations in a logical language
- **Shared semantics:** axioms that constrain the interpretation of the terms and help disambiguate the terms

Ontologies are *heavy* vs. *light* analogous to *reference manual* vs. *user guide*

- A reference ontology is necessarily **heavy**: complete, formal, rigorous
- Implementation/user guide is usually light

Current geospatial standards

ISO/OGC Simple Features, OGC GeoSPARQL, Spatial Schema (ISO 19107), Ordnance Survey Spatial Relations, GML, hydro ontologies (GWML, INSPIRE, SWEET)

- Specified using UML, RDF Schema, or lightweight OWL (OWL-DL)
 - ▶ **Light: standardize the terms** (vocabulary)
 - ▶ **Don't formalize the terms' meaning:** *not a formal reference*
 - ▶ Only the beginning of exploiting the benefits of ontologies for standards
 - ▶ Even the expressive power of OWL-DL not fully exploited yet
- Relations (between concepts) are less emphasized than concepts
 - ▶ Relations tie concepts together: need relations to describe how certain concepts relate to one another (incl. behaviour)
 - ▶ OWL language is less expressive with respect to relations
- Many concepts and relations are already formalized
 - ▶ e.g. mereotopological relations (RCC and Egenhofer's 9-intersections) are included in GeoSPARQL, Simple Features, and Ordnance Survey Spatial Relations, but **neither use the known logical formalizations**

How can expressive ontologies help improve standards?

General idea:

- 1 Identify key concepts and relations (terminology)
- 2 Axiomatize them in an expressive logic (e.g. Common Logic)
 - ▶ Identify primitive vs. definable concepts and relations
 - ▶ Constrain primitive concepts/relations
 - ▶ Define definable concepts/relations
- 3 Extract concept and relation hierarchies and verify
 - ▶ Use automated theorem provers
 - ▶ Verify consistency and that concepts and relations can be non-empty
 - ▶ Extract, e.g., subclass and subproperty relationships
 - ▶ Extract `DisjointWith` and `DisjointUnionOf` conditions
 - ▶ Essentially extract a lightweight ontology

Will use examples from current work on hydro ontology as examples

Goal: Multiple consistent representations of a standard

Common Logic

- Full specification
- Axioms relating concepts/relations
- Primitive vs. definable concepts/relations

extract

OWL

- Concept hierarchy
- Property hierarchy
- Domain and range restrictions for properties

extract

RDF

- terminology

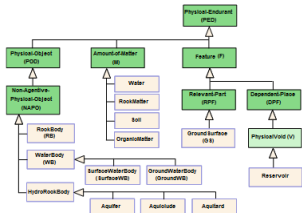
ground
& verify

classify

```

10
11 (cl-text multiaxis_herontology_codis/cont_bcont.clif
12
13 (cl-imports multiaxis_herontology_codis/cont.clif)
14 (cl-imports multiaxis_herontology_codis/definitions/has_max_dim.clif)
15 (cl-imports multiaxis_herontology_codis/definitions/has-obj.clif)
16 (cl-imports multiaxis_herontology_codis/definitions/sg.clif)
17
18 (cl-imports multiaxis_herontology_codis/theorems/in_theorem.clif)
19 (cl-imports multiaxis_herontology_codis/theorems/ax_theorem.clif)
20
21 (cl-statement 'RC-31: entities is the boundary are at least of a dimension lower than the bounded entity')
22
23 (forall (x y)
24   (if
25     (BCont x y)
26     (and
27       (Cont x y)
28       (Dgt x y)
29     )
30   )
31 )
32
33 (cl-statement 'RC-32: if a is contained in two superficially connected entities x and y embedded in an
entity v of the dimension of a, that is x has locally a cont of v, then z is contained in
the boundary of a')
34
35 (forall (x y v z)
36   (if
37     (and
38       (BCont x y)
39       (Btan x)
40       (P x v)
41       (Cont y v)
42       (Cont z x)
43       (Cont z y)
44     )
45     (BCont z x)
46   )
47 )
  
```

ground



Concepts:

PED, POB, M, F, RPF,
DPF, V, WaterBody,
RockBody

Relations (Properties):

Cont (contains), inside,
surrounds, hosts, hosts-v,
DK (constitutes), P
(part_of), PO (overlap),
Inc (incidence), BCont
(boundary contained),
etc.

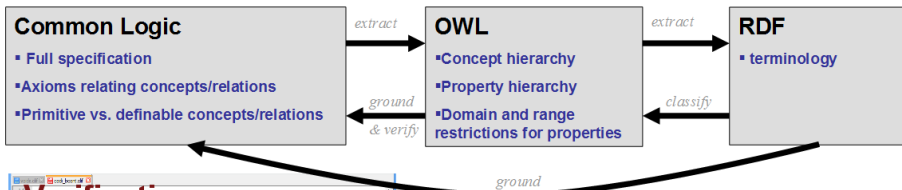
Concept hierarchy:

SubclassOf Reservoir V
SubclassOf V DPF

Property hierarchy:

SubPropertyOf hosts hosts-v
SubPropertyOf inside phys-contains

Goal: Multiple consistent representations of a standard

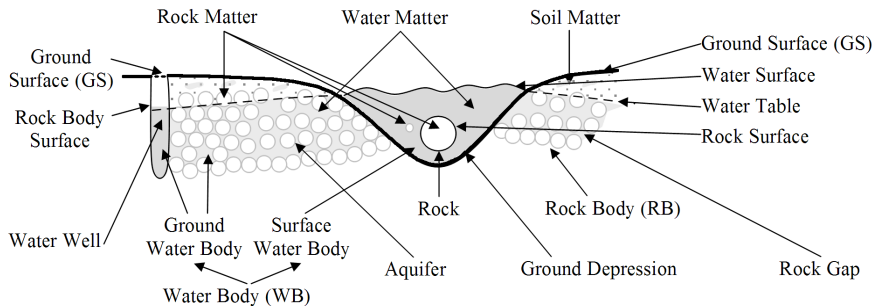


```
Verification:
13 (c) reports multidim_mereotopology_cod/odi_bif
14
15 2013-10-16 01:08:22,820 src.ClifModule INFO CREATED LADR TRANSLATION:
16 /stl/torsten/git/macleod/ga/multidim_mereotopology_dim/conversions/dim_basic.p9
17
18 2013-10-16 01:08:23,170 src.ClifModuleSet INFO USING 3 REASONERS: ['prover9', 'vampire', 'paradox']
19
20 ..
21
22 2013-10-16 01:08:24,594 src.process INFO FOUND MODEL: paradox
23
24 2013-10-16 01:08:27,914 src.ClifModuleSet INFO TERMINATED SUCCESSFULLY (1): paradox
25
26 2013-10-16 01:08:27,915 src.ClifModuleSet INFO CONSOLIDATED RESULT: 1
27
28 2013-10-16 01:08:27,915 src.ClifModuleSet INFO CONSISTENT (return value = 1): multidim_mereotopology_cod
29
30 imports = set([multidim_mereotopology_codi/definitions/point_region, multidim_mereotopology_cont/cont_ext, multidim_mereotopology_codi/definitions/areal_region, multidim_mereotopology_codi/definitions/min_max_in_dim, multidim_mereotopology_codi/definitions/sc, multidim_mereotopology_codi/definitions/ep, multidim_mereotopology_dim/definitions/eq_dim, multidim_mereotopology_dim/dim_prime_linear_unbounded, multidim_mereotopology_codi/definitions/ep, multidim_mereotopology_dim/definitions/dim_basic_defs, multidim_mereotopology_codi/definitions/c, multidim_mereotopology_codi/codi_linear, multidim_mereotopology_dim/definitions/curve, multidim_mereotopology_codi/codi_basic, multidim_mereotopology_dim/dim_prime_linear_unbounded])
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
```

Concepts:

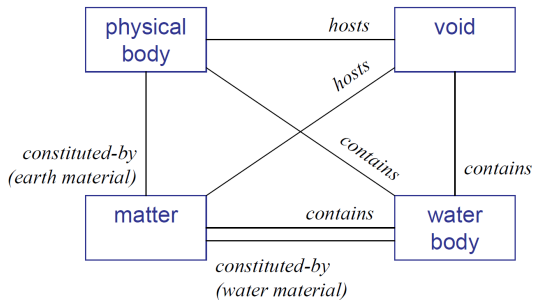
<u>SubclassOf</u> Reservoir V	<u>SubPropertyOf</u> hosts hosts-v
<u>SubclassOf</u> V DPF	<u>SubPropertyOf</u> inside phys-contains

Our work: Develop **formally grounded** hydro ontology



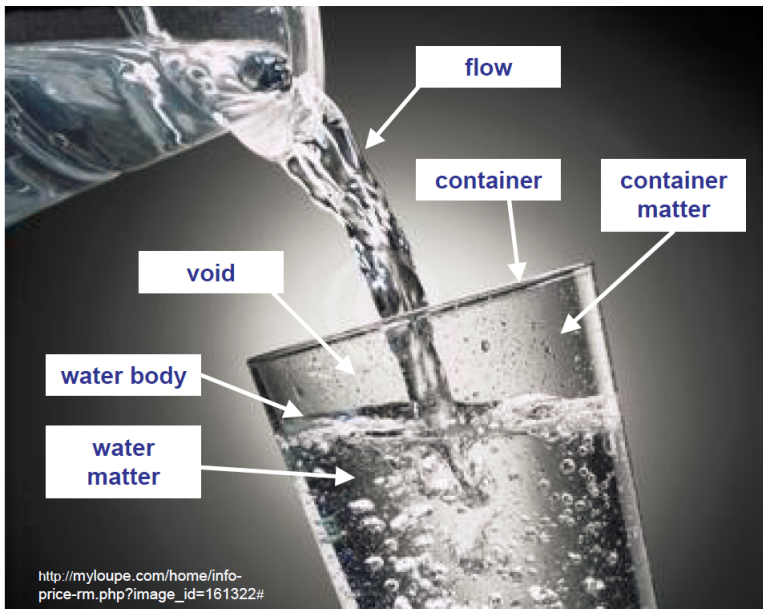
- 1 Rigorous formalization in Common Logic
- 2 Verification: assisted by first-order theorem provers; partly automated
 - ▶ prove consistency
 - ▶ prove coverage: exhaustiveness of concepts/relations
 - ▶ prove intuitive intended relationships ('theorems')
- 3 Extract taxonomies to extend OWL version of DOLCE upper ontology

Starting point: Basic elements of a hydro ontology



- Develop a rigorous formalization of these concepts and relations in a formal logic & extract a consistent lightweight vocabulary

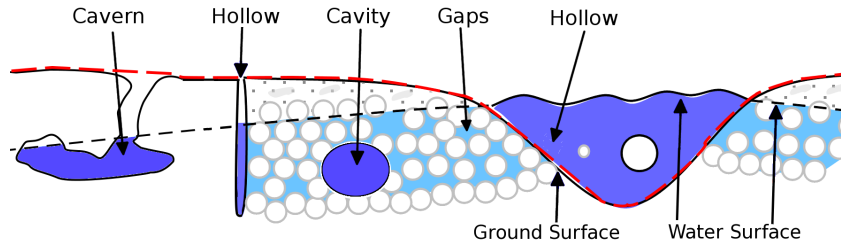
Basic elements of a reference hydro ontology: Analogy



Water and physical bodies (Hahmann & Brodaric, 2012)

Define water bodies by their physical containers' voids

- Lake or River WB: in a hollow of the ground surface
- Water Well WB: in a hollow below the ground surface
- Aquifer WB: in gaps in the rock matter and in holes below the ground surface



Example axioms: Water and rock bodies

A WaterBody may only be constituted by water if it has constituents:

$$WB(x) \rightarrow NAPO(x) \wedge \forall y[DK_1(y, x) \rightarrow Water(y)]$$

A RockBody is constituted by rock matter and only by rock matter:

$$RB(x) \equiv NAPO(x) \wedge \exists y[DK_1(y, x)] \wedge \forall y[DK_1(y, x) \rightarrow RockMatter(y)]$$

GS denotes a ground surface (not fully defined):

$$GS(gs) \rightarrow RPF(gs) \wedge \exists o[NAPO(o) \wedge hosts(o, gs)]$$

WB, RB, GS, Water, RockMatter Domain theory (Hydrogeology)

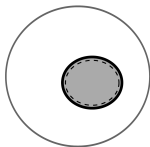
NAPO, RPF, DK₁, hosts DOLCE concepts/relations

Voids (Hahmann & Brodaric, 2012)

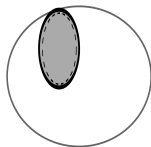
Holes vs. Gaps: based on whether the host is internally self-connected

Cavities vs. Tunnels vs. Depressions: based on the void's opening

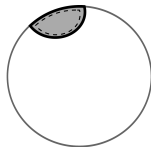
Opening to the outside vs. opening to other voids only



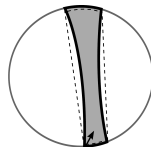
no opening:
Internal Cavity



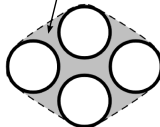
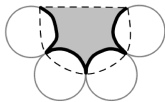
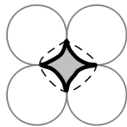
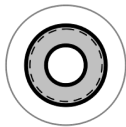
point-opening:
Tangential Cavity



single Icon opening
to the outside:
External Hollow



multiple openings:
Tunnel (System)



Example axioms: Water and rock bodies (contd.)

Surface- vs. Ground-WaterBody:

$$\text{SurfaceWB}(wb) \rightarrow \text{WB}(wb) \wedge \exists gs[\text{hol}_e(wb, gs) \wedge \text{GS}(gs)]$$

$$\text{GroundWB}(wb) \rightarrow \text{WB}(wb) \wedge \exists rb, gs[\text{RB}(rb) \wedge \text{hosts}(rb, gs) \wedge \text{GS}(gs) \wedge \\ r(wb) \subseteq \text{voidspace}(rb) \wedge \forall v[\text{hol}_e(rb, v) \rightarrow \neg \text{PO}(wb, v)]]$$

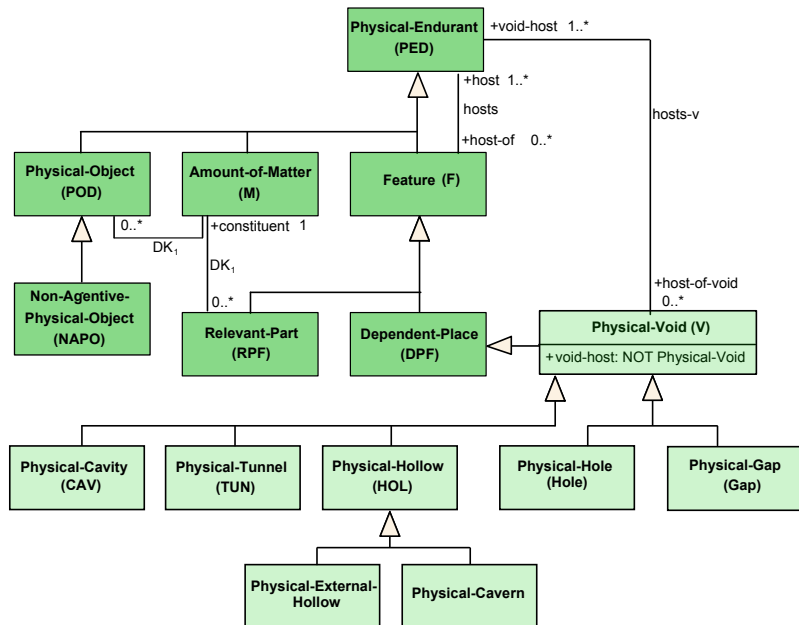
A HydroRockBody consists of a RockBody and a GroundWaterBody with the GroundWaterBody located in Voids of the RockBody:

$$\text{HydroRockBody}(aq) \rightarrow \text{NAPO}(aq) \wedge \exists rb, wb[r(aq) = r(rb) + r(wb) \wedge \\ \text{RB}(rb) \wedge \text{GroundWB}(wb) \wedge \\ r(wb) \subseteq \text{con-voidspace}(rb)]$$

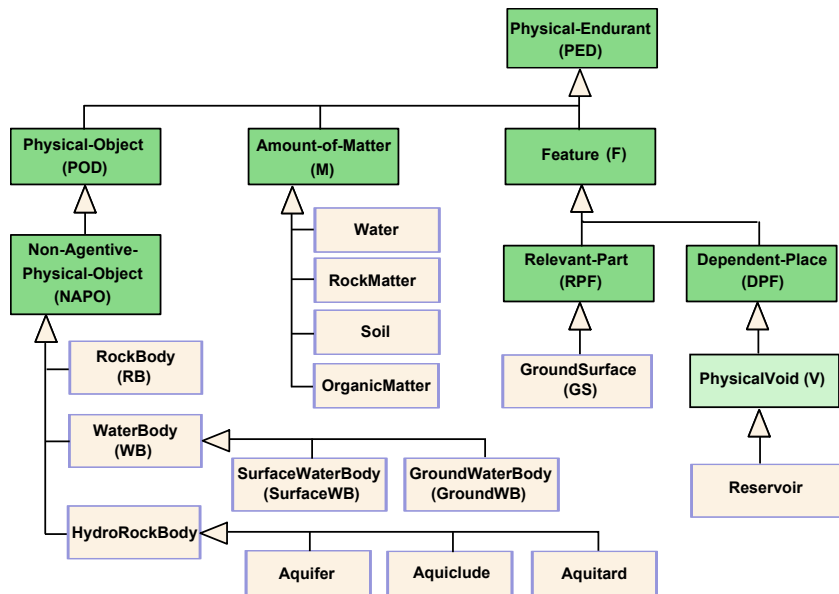
A Reservoir is the voidspace of some RockBody:

$$\text{Reservoir}(wr) \equiv \mathbf{V}(wr) \wedge \exists rb[\text{RB}(rb) \wedge r(wr) = \text{voidspace}(rb)]$$

DOLCE with voids – OWL version

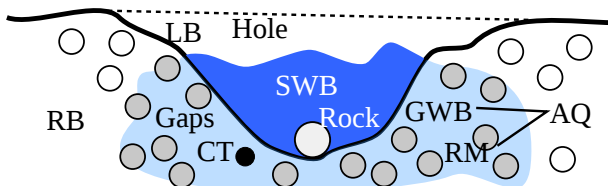


DOLCE with hydrogeology concepts – OWL version



Containment relations (Hahmann & Brodaric, 2013)

Relate voids, water bodies, and other physical bodies through containment relations



openly-surrounds-mat(RB, SWB)
hosts-v(RB, Hole)
mat-inside(Rock, Hole)
materially-contains(AQ, GWB)
encloses-mat(AQ, CT)
mat-inside(Gaps, GWB)
encloses-mat(GWB, CT)

openly-surrounds-mat(RB, Rock)
mat-inside(SWB, Hole)
openly-surrounds-mat(SWB, Rock)
materially-contains(AQ, RM)
hosts-v_{any}(AQ, Gaps)
mat-inside(Gaps, CT)

Containment relations: Heavy approach first

- Precise definitions based on topological-geometric containment relations, physical constraints and DOLCE concepts:

$$\text{fully-phys-contains}(y, x) \leftrightarrow PED(x) \wedge PED(y) \wedge P(r(x), ch(y)) \wedge [\neg mat(y) \rightarrow P(r(x), r(y))]$$

- Classify physical containment relations based on
 - ① whether container and containee are in a **physical dependency**
 - ② whether the **container** is a material or a void endurant
 - ★ **inside** (a void) vs. **surrounded** (by a material endurant)
 - ③ whether the **containee** is a material or void endurant
 - ④ other spatial relations: enclosure, contact, spatial parthood
- The resulting “leaf” relations are exhaustive and pairwise disjoint

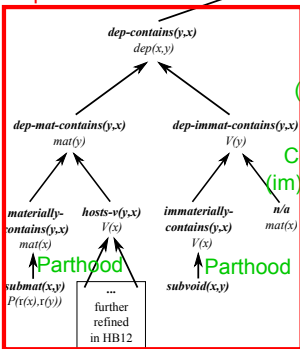
Containment relations: Light version follows

fully-phys-contains(y,x)
 $P(r(x),ch(y))$
 $\neg mat(y) \rightarrow P(r(x),r(y))$

Dependent containment

Dependence

Detachable containment

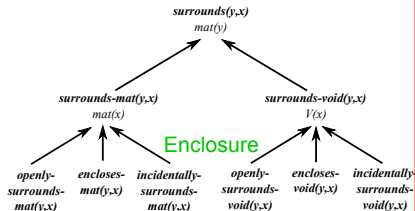
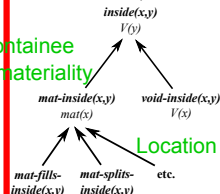


Container
(im)materiality

Containe
(im)materiality

Location

Enclosure



Taxonomy expressible in OWL using subproperty and DisjointWith relationships; cannot express exhaustiveness in OWL

Conclusions

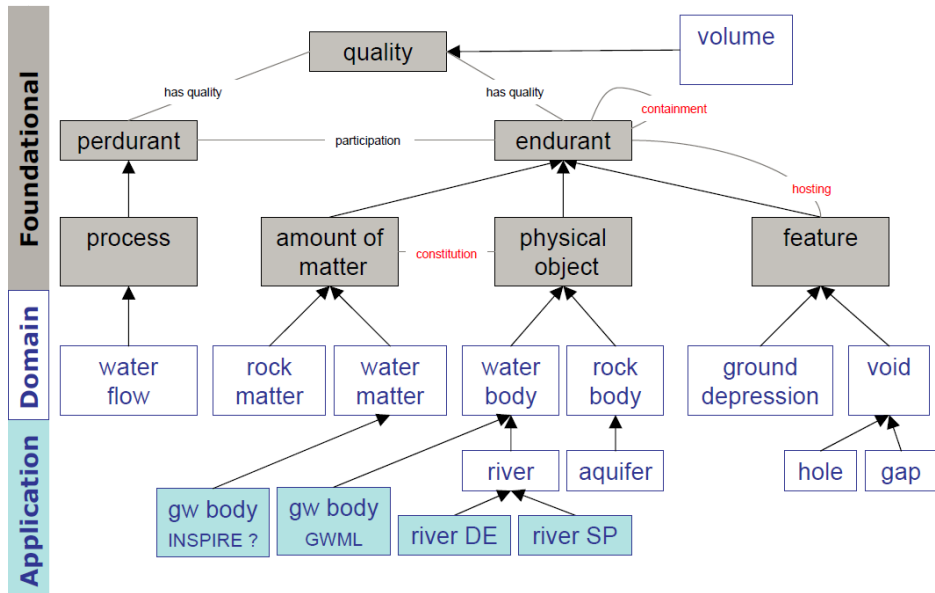
Critical to ground any lightweight implementation representation (“user guide”) in a formal reference (“technical specification”)

- Formally grounds and disambiguates geospatial concepts
- Serves as basis for **(semi-)automated** extraction of lightweight versions (OWL, RDF) that can be used as terminological reference for annotation or implementation in a triple store
- Formal specification helps automated verification

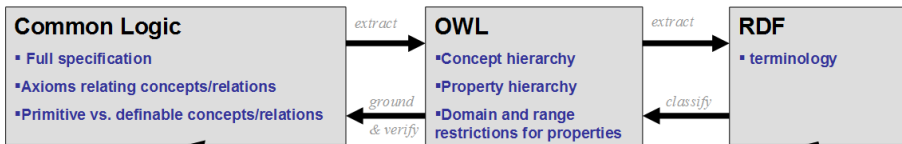
Standards should be flexible in two ways

- ① Amendable to various applications or domains, i.e., not too specific
- ② Offer various degrees of formality
 - ▶ Most formal: for reference, verification, and heavyweight reasoning
 - ▶ Least formal: as terminology for annotating data (‘Linked Data’)

1) Formally grounded tiered standards



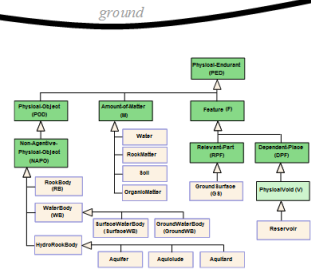
2) Various degrees of formality



```

10 (cl-text multiaxis_herpetology_codis/cont_bcont.clif
11
12 (cl-imports multiaxis_herpetology_codis/codis.clif)
13 (cl-imports multiaxis_herpetology_codis/definitions/has_max_dim.clif)
14 (cl-imports multiaxis_herpetology_codis/definitions/has_min_dim.clif)
15 (cl-imports multiaxis_herpetology_codis/definitions/has_min_dim.clif)
16 (cl-imports multiaxis_herpetology_codis/definitions/has_min_dim.clif)
17
18 (cl-imports multiaxis_herpetology_codis/theorems/ax_theorem.clif)
19 (cl-imports multiaxis_herpetology_codis/theorems/ax_theorem.clif)
20
21 (cl-statement 'RC-31: entities in the boundary are at least of a dimension lower than the bounded entity')
22
23 (forall (x y)
24   (if
25     (BCont x y)
26     (and
27       (Cont x y)
28       (Cont y y)
29       (Cont x x)
30     )
31   )
32 )
33
34 (cl-statement 'RC-32: if a is contained in two superficially connected entities x and y embedded in an
35 entity v of the dimension of a, that is x has locally a codimension of zero, then z is contained in
36 the boundary of a')
37
38 (forall (x y v z)
39   (if
40     (and
41       (BCont x y)
42       (BCont x v)
43       (BCont y v)
44       (Cont z x)
45       (Cont z y)
46     )
47     (BCont z v)
48   )
49 )
50

```



Concepts:
 PED, POB, M, F, RPF,
 DPF, V, WaterBody,
 RockBody

Relations (Properties):
 Cont (contains), inside,
 surrounds, hosts, hosts-v,
 DK (constitutes), P
 (part_of), PO (overlap),
 Inc (incidence), BCont
 (boundary contained),
 etc.

Concept hierarchy:
 SubclassOf Reservoir V
 SubclassOf V DPF

Property hierarchy:
 SubPropertyOf hosts hosts-v
 SubPropertyOf inside phys-contains

Two general observations

- Much work on upper ontologies, less on the middle layer
 - ▶ upper ontologies can only be formalized to a certain degree
 - ▶ narrow application-specific ontologies are often too tedious to formalize
 - ▶ missing the middle layer: can be formally standardized
 - ★ specific enough but also not too many concepts and relations
 - ★ that's the level where information integration and exchange happens
- Relations are often still neglected: less understood?
 - ▶ Relations define how concepts relate to one another

Publications

T. Hahmann, B. Brodaric: **The Void in Hydro Ontology**. In: Proc. of the 7th Int. Conference on Formal Ontology in Information Systems (FOIS-2012), 2012. IOS Press.

T. Hahmann, B. Brodaric: **Kinds of Full Physical Containment**. In: Proc. of the 11th Int. Conference on Spatial Information Theory (COSIT-2013), 2013, Springer.

T. Hahmann: **Reconciliation of Logical Theories of Space: from Multidimensional Mereotopology to Geometry**, PhD thesis, University of Toronto. Feb. 2013.

Full formalizations of the ontologies (in progress), COLORE repository, <http://stl.mie.utoronto.ca/colore/org.html>, Section “Space”

- ontologies discussed here are named `multidim_mereotopology_XXX` and `multidim_space_XXX`

Acknowledgements: Boyan Brodaric, Michael Gruninger

Thank you!