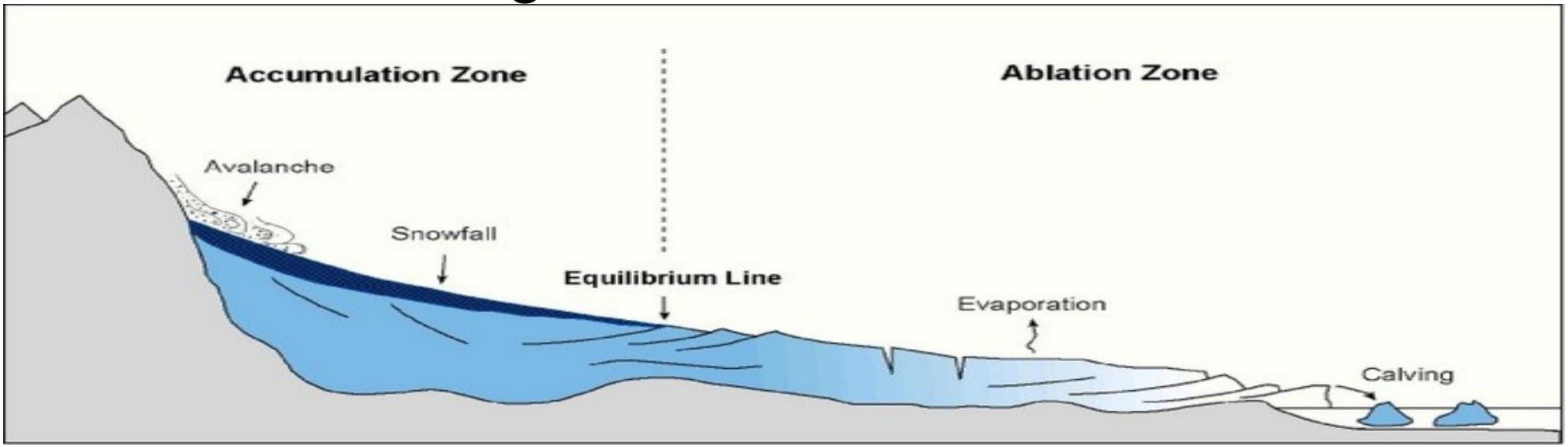
Semantic Harmonization

Illustrating harmonization levels -vocabularies, onologies and ODPs...

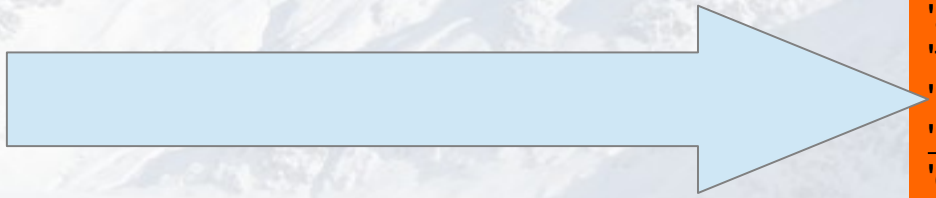


Gary Berg-Cross (consultant) & Charles Vardeman (Center for Research Computing (CRC), U. of Notre Dame)

Winter ESIP Meeting, January 2020

The Vocabulary Can Guide Ontology Refinement

Added & moved some sub-types
- ice mass
+ glacier
- ice cap
- icefall
- ice lens
+ anchor ice
+ floating ice mass



Some concepts differ by where ice is as well as its constituents (salt vs fresh water) like:

ice shelf - A floating ice sheet of considerable thickness....
attached to the coast.

'ice mass'
'anchor ice'
'floating ice mass'
'ice calf'
'land ice mass'
'glacier snout'
'ice cap'
'ice lens'
'ice shelf'
'ice tongue'
iceberg
icefall
'sea ice mass'
'water ice core'

Some Refinements are easier & direct
Updating EnvO Ice Shelf Example – more spatial relations

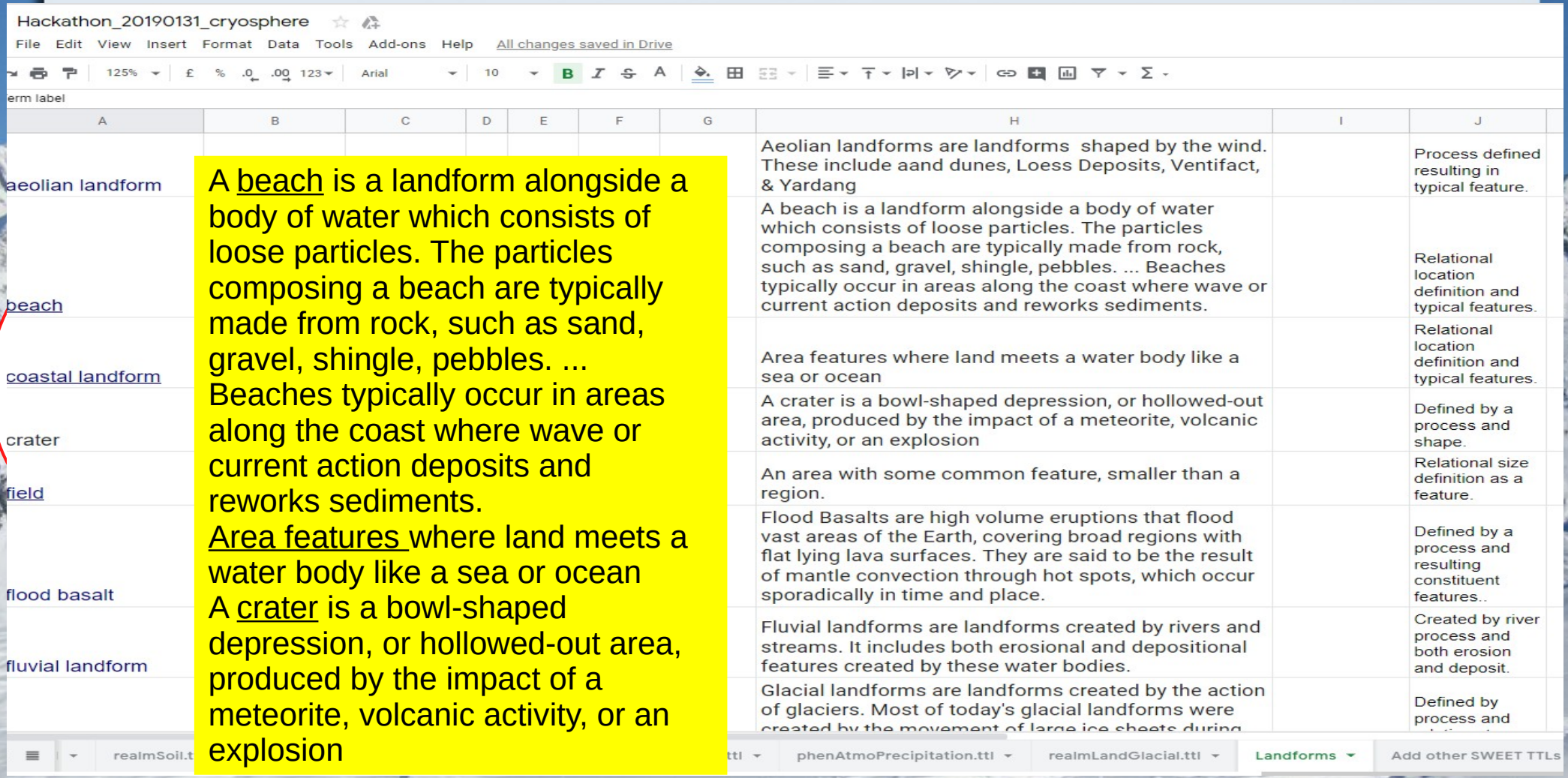
Old EnvO axioms for Ice Shelf

- formed as result of some glacial transport process
- overlaps some sea coast
- floating ice mass

Updated EnvO Ice Shelf

- 'adjacent to' some 'marine water body'
- 'attached to' some 'sea coast'
- 'formed as result of' some 'mass ice flow'
- 'formed as result of' some snowfall
- 'has quality' some buoyancy
- 'land ice mass'
- partially_surrounded_by some atmosphere'

Terrain Terms From SWEET



Hackathon_20190131_cryosphere

File Edit View Insert Format Data Tools Add-ons Help All changes saved in Drive

erm label

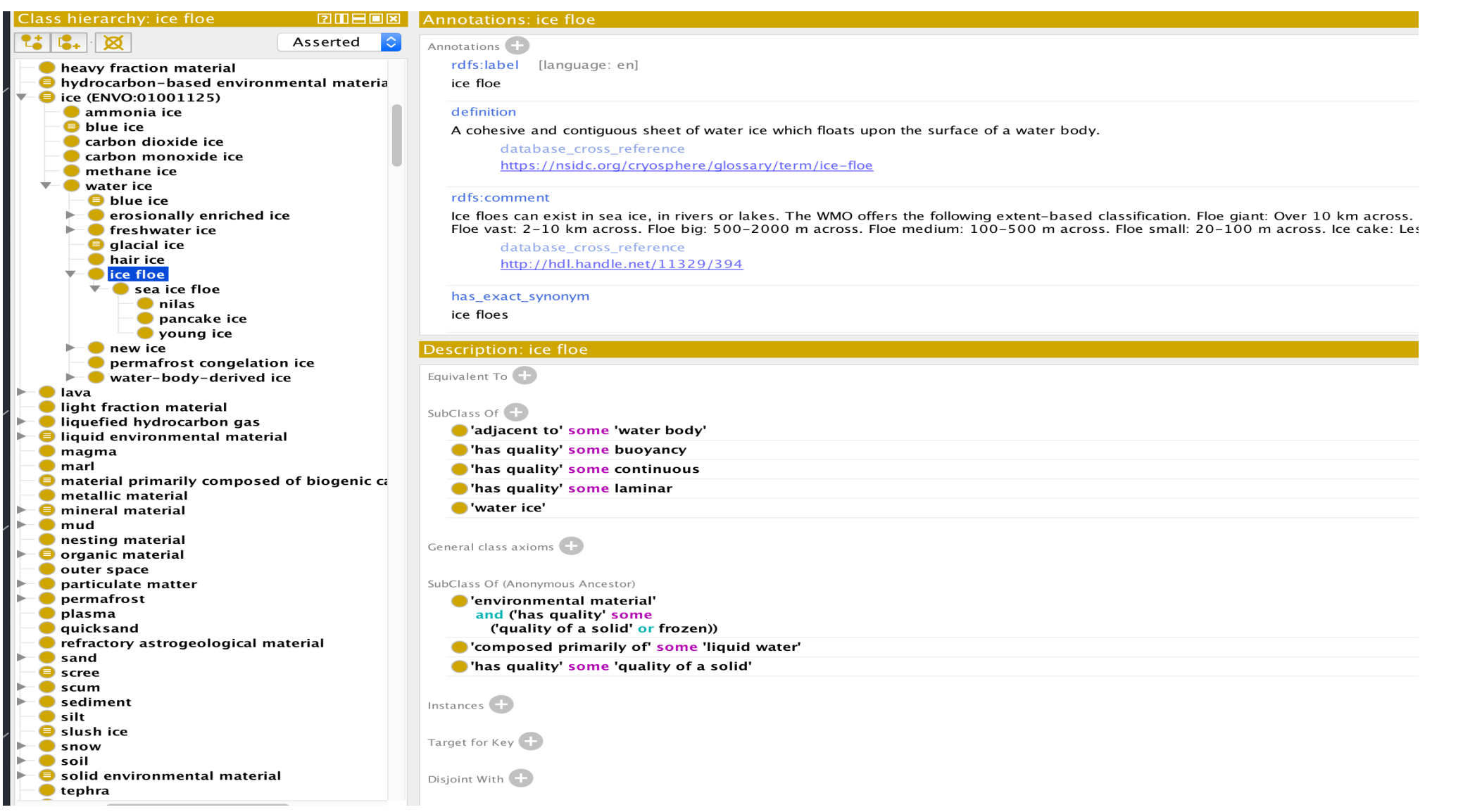
A	H	I	J
aeolian landform	Aeolian landforms are landforms shaped by the wind. These include aand dunes, Loess Deposits, Ventifact, & Yardang		Process defined resulting in typical feature.
beach	A beach is a landform alongside a body of water which consists of loose particles. The particles composing a beach are typically made from rock, such as sand, gravel, shingle, pebbles. ... Beaches typically occur in areas along the coast where wave or current action deposits and reworks sediments.		Relational location definition and typical features.
coastal landform	Area features where land meets a water body like a sea or ocean		Relational location definition and typical features.
crater	A crater is a bowl-shaped depression, or hollowed-out area, produced by the impact of a meteorite, volcanic activity, or an explosion		Defined by a process and shape.
field	An area with some common feature, smaller than a region.		Relational size definition as a feature.
flood basalt	Flood Basalts are high volume eruptions that flood vast areas of the Earth, covering broad regions with flat lying lava surfaces. They are said to be the result of mantle convection through hot spots, which occur sporadically in time and place.		Defined by a process and resulting constituent features..
fluvial landform	Fluvial landforms are landforms created by rivers and streams. It includes both erosional and depositional features created by these water bodies.		Created by river process and both erosion and deposit.
	Glacial landforms are landforms created by the action of glaciers. Most of today's glacial landforms were created by the movement of large ice sheets during		Defined by process and

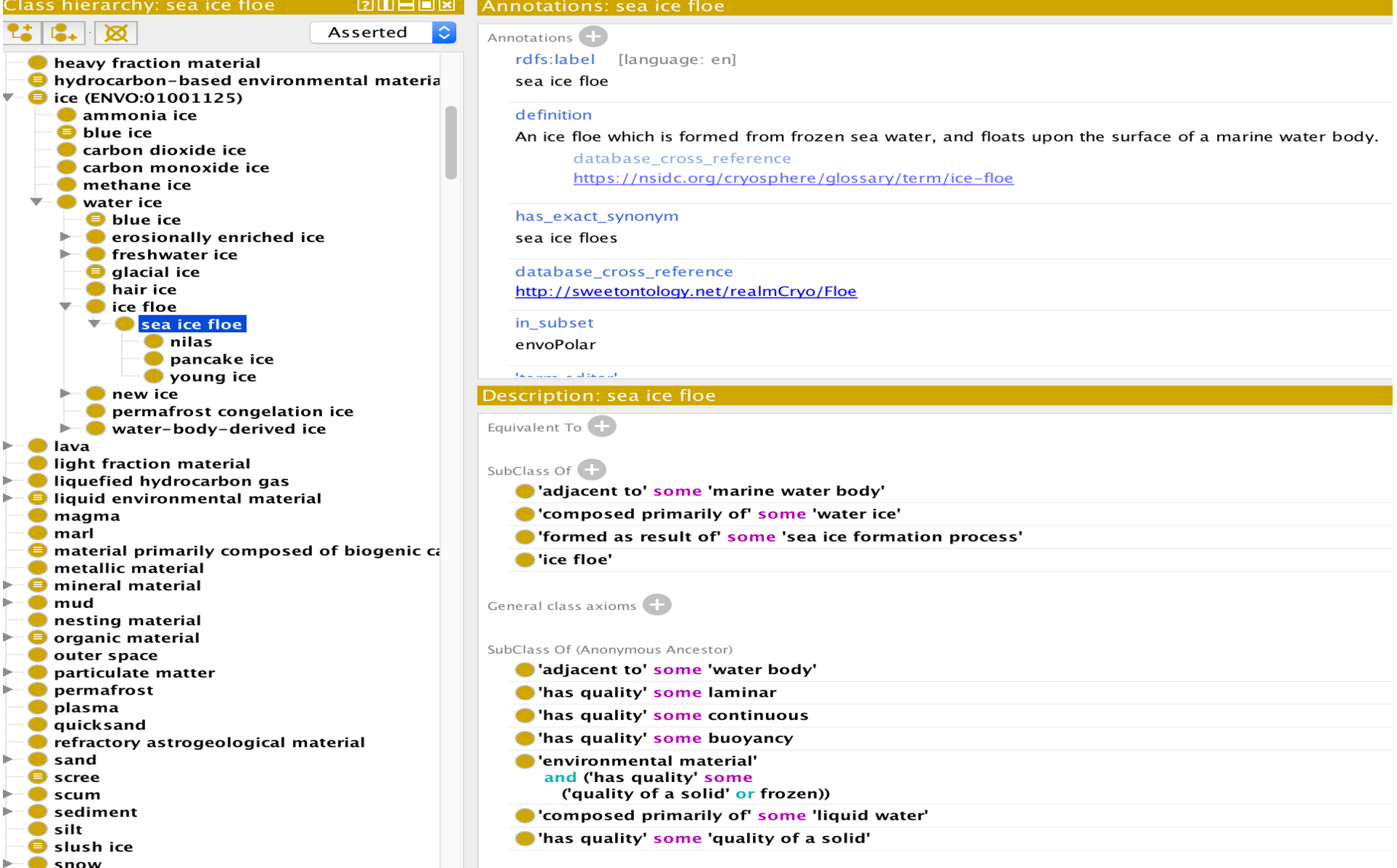
A beach is a landform alongside a body of water which consists of loose particles. The particles composing a beach are typically made from rock, such as sand, gravel, shingle, pebbles. ... Beaches typically occur in areas along the coast where wave or current action deposits and reworks sediments.
Area features where land meets a water body like a sea or ocean
A crater is a bowl-shaped depression, or hollowed-out area, produced by the impact of a meteorite, volcanic activity, or an explosion

realmSoil.t ... phenAtmoPrecipitation.ttl realmLandGlacial.ttl Landforms Add other SWEET TTLs

Terrain Vocabulary Sources

- Neuendorf, K.K.E., Mehl Jr., J.P., and Jackson, J.A. (ed.) 2005.
  - Glossary of geology, 5th Ed. American Geological Institute, Alexandria, VA. as gathered in the National Soil Survey Handbook (430-629-H, 1 st Ed., Amend. 25, Nov. 2017) which has a Glossary of Landform and Geologic Terms.
- Not all of the terms were there so wikipedia was my 2nd source.


Class hierarchy: ice floe
Asserted
heavy fraction material
hydrocarbon-based environmental materia
ice (ENVO:01001125)
ammonia ice
blue ice
carbon dioxide ice
carbon monoxide ice
methane ice
water ice
blue ice
erosionally enriched ice
freshwater ice
glacial ice
hair ice
ice floe
sea ice floe
nilas
pancake ice
young ice
new ice
permafrost congelation ice
water-body-derived ice
lava
light fraction material
liquefied hydrocarbon gas
liquid environmental material
magma
marl
material primarily composed of biogenic ca
metallic material
mineral material
mud
nesting material
organic material
outer space
particulate matter
permafrost
plasma
quicksand
refractory astrogeological material
sand
scree
scum
sediment
silt
slush ice
snow
soil
solid environmental material
tephra
Annotations: ice floe
Annotations
rdfs:label [language: en]
ice floe
definition
A cohesive and contiguous sheet of water ice which floats upon the surface of a water body.
database_cross_reference
https://nsidc.org/cryosphere/glossary/term/ice-floe
rdfs:comment
Ice floes can exist in sea ice, in rivers or lakes. The WMO offers the following extent-based classification. Floe giant: Over 10 km across. Floe vast: 2-10 km across. Floe big: 500-2000 m across. Floe medium: 100-500 m across. Floe small: 20-100 m across. Ice cake: Les
database_cross_reference
http://hdl.handle.net/11329/394
has_exact_synonym
ice floes
Description: ice floe
Equivalent To
SubClass Of
'adjacent to' some 'water body'
'has quality' some buoyancy
'has quality' some continuous
'has quality' some laminar
'water ice'
General class axioms
SubClass Of (Anonymous Ancestor)
'environmental material'
and ('has quality' some
('quality of a solid' or frozen))
'composed primarily of' some 'liquid water'
'has quality' some 'quality of a solid'
Instances
Target for Key
Disjoint With


Class hierarchy: sea ice floe
Asserted
heavy fraction material
hydrocarbon-based environmental materia
ice (ENVO:01001125)
ammonia ice
blue ice
carbon dioxide ice
carbon monoxide ice
methane ice
water ice
blue ice
erosionally enriched ice
freshwater ice
glacial ice
hair ice
ice floe
sea ice floe
nilas
pancake ice
young ice
new ice
permafrost congelation ice
water-body-derived ice
lava
light fraction material
liquefied hydrocarbon gas
liquid environmental material
magma
marl
material primarily composed of biogenic ca
metallic material
mineral material
mud
nesting material
organic material
outer space
particulate matter
permafrost
plasma
quicksand
refractory astrogeological material
sand
scree
scum
sediment
silt
slush ice
snow
Annotations: sea ice floe
Annotations
rdfs:label [language: en]
sea ice floe
definition
An ice floe which is formed from frozen sea water, and floats upon the surface of a marine water body.
database_cross_reference
https://nsidc.org/cryosphere/glossary/term/ice-floe
has_exact_synonym
sea ice floes
database_cross_reference
http://sweetontology.net/realmCryo/Floe
in_subset
envoPolar
Description: sea ice floe
Equivalent To
SubClass Of
'adjacent to' some 'marine water body'
'composed primarily of' some 'water ice'
'formed as result of' some 'sea ice formation process'
'ice floe'
General class axioms
SubClass Of (Anonymous Ancestor)
'adjacent to' some 'water body'
'has quality' some laminar
'has quality' some continuous
'has quality' some buoyancy
'environmental material'
and ('has quality' some
('quality of a solid' or frozen))
'composed primarily of' some 'liquid water'
'has quality' some 'quality of a solid'

```
http://sweetontology.net/realmCryo/Floe
:Floe rdf:type owl:Class ;
    owl:equivalentClass soreac:IceFloe ;
    rdfs:subClassOf soreac:SeaIce ;
    rdfs:label "floe"@en ;
    rdfs:comment "DBXref: http://purl.obolibrary.org/obo/ENVO_03000066"@en;
    skos:definition  [      You, a month ago • Adding definitions to relmCryo
      rdfs:comment  "An ice floe which is formed from frozen sea water, and floats upon the surface of a marine water body."@en ;
      dcterms:source <https://nsidc.org/cryosphere/glossary/term/ice-floe> ;
      dcterms:created "2019-12-10T08:54:28-08:00Z"^^xsd:dateTimeStamp ;
      dcterms:creator <https://orcid.org/0000-0003-4091-6059> ;
    ] .
```

```
Follow link (option + click)
###  http://sweetontology.net/realmCryo/IceFloe
soreac:IceFloe rdf:type owl:Class ;
               rdfs:label "ice floe"@en .
```

tps://github.com/ESIPFed/sweet/issu

charlesvardeman commented 27 days ago — Collaborator

During the SWEET-ENVO Cryosphere harmonization "Vocamps", database cross-references were captured in ENVO using the database_cross_reference pointing to the SWEET URI for the term. The reciprocal relation was not captured in SWEET because no mechanism currently exists. ENVO utilizes the <http://www.geneontology.org/formats/oboInOwl#hasDbXref > annotation property for capturing this information that is contained in the OBO Format metamodel. During the AGU Community Ontology Engineering Workshop several options were discussed to add the ability for cross-references in SWEET.

1. Directly reference the http://www.geneontology.org/formats/oboInOwl# prefix in SWEET and use hasDbXref annotation property.
2. Create a new class in SWEET that is essentially the equivalent to:

```
:DbXref a owl:Class ;
:hasDbXref a owl:AnnotationProperty .
:hasURI a owl:AnnotationProperty .
```

Where the :hasDbXref property takes the "http://www.w3.org/2001/XMLSchema#anyURI datatype". No obvious consensus was reached during the workshop as to where this class and property should be located in the SWEET hierarchy. It was discussed that a potential solution would be to create a utility module to contain these classes which would have the advantage of being extensible if other relations from oboInOwl need to be utilized (ex #replacedBy, #consider, etc).

cmungall commented 26 days ago — Collaborator

I recommend using skos

I think we should also include skos in the ENVO release. We use dbxref for historic reasons. See

https://douroucouli.wordpress.com/2019/05/27/never-mind-the-logix-taming-the-semantic-anarchy-of-mappings-in-ontologie/

cmungall commented 26 days ago — Collaborator

To be clear on the workflow:

the master location will be in ENVO, all edits will be made there, and as part of the sweet release process or as a periodic step there will be some kind of SPARQL UPDATE or similar to update SWEET?

Monkeying around with OWL

Musings on building and using ontologies, posts by Chris Mungall

HOME ABOUT

ONTOTIPS: A SERIES OF ASSORTED ONTOLOGY DEVELOPMENT GUIDELINES BIOLOGICAL KNOWLEDGE

← OntoTip: Single-inheritance principle considered dangerous

OntoTip: Clearly document your design decisions →

Never mind the logix: taming the semantic anarchy of mappings in ontologies

MAY 27, 2019 LEAVE A COMMENT

Mappings between ontologies, or between an ontology and an ontology-like resource, are a necessary fact of life when working with ontologies. For example, GO provides mappings to external resources such as KEGG, MetaCyc, RHEA, EC, and many others. Uberon (a multi-species anatomy ontology) provides mappings to species-specific anatomy ontologies like ZFA, FMA, and also to more specialized resources such as the Allen Brain Atlases. These mappings can be used for a variety of purposes, such as data integration – data annotated using different ontologies can be 'cross-walked' to use a single system.

https://douroucouli.wordpress.com/2019/05/27/never-mind-the-logix-taming-the-semantic-anarchy-of-mappings-in-

Option 2. Use oboInOwl hasDbXref property

Before there was OWL, there was OBO-Format. And lo, OBO-Format gave us the xref. Well not really, the xref was just an example of the long standing tradition of database cross-reference in bioinformatics. In bioinformatics we love minting new IDs. For any given gene you may have its ENSEMBL ID, it's MOD or HGNC ID, it's OMIM ID, it's NCBI Gene/Entrez ID, and a host of other IDs in other databases. The other day I caught my cat minting gene IDs. It's widespread. This necessitates a system of cross-references. These are rarely 1:1, since there are reasons for representations in different systems to diverge. The OBO-Format xref was for exactly the same use case. When GO started, there were already similar overlapping databases and classifications, including longstanding efforts like EC.

In the OWL serialization of OBO-Format (oboInOwl) this becomes an annotation assertion axioms using the *oboInOwl:hasDbXref* property. Many ontologies such as GO, HPO, MONDO, UBERON, ZFA, DO, MP, CHEBI, etc continue to use the xref as the primary way to express mappings, even though they are no longer tied to obo format for development.

https://douroucouli.wordpress.com/2019/05/27/never-mind-the-logix-taming-the-semantic-anarchy-of-mappings-in-

Option 3. Use SKOS vocabulary for mapping properties

In the traditional tale of Goldilocks and the three OWLs, Goldilocks tries three bowls of semantic porridge. The first is too strong, the second too weak, and the third one is just right. If the first bowl is OWL logical axioms, the second bowl is oboInOwl xrefs, the third bowl would be the Simple Knowledge Organization System (SKOS) mapping vocabulary.

This provides a hierarchy of mapping properties

- mappingRelation
  - closeMatch
  - broadMatch
  - narrowMatch
  - exactMatch

These can be used to link SKOS concepts across different concept schemes. The exactMatch property has the properties of transitivity and symmetry, but is still weaker than owl equivalence as it lacks the property of substitutibility. SKOS properties are axiomatized allowing entailment. Note that broad and narrow match are not transitive, but they both entail broader transitive properties transitiveBroadMatch and narrowBroadMatch.

Using skos mapping relations, we can map between an OBO ontology and MESH without worrying about the lack of OWL semantics for MESH. We can use exactMatch for 1:1 mappings, and closeMatch if we are less confident. We don't have to worry about injecting semantics, it's just a mapping!

https://douroucouli.wordpress.com/2019/05/27/never-mind-the-logix-taming-the-semantic-anarchy-of-mappings-in-

Expressing Mappings in OWL

Option 1. Direct logical axioms

OWL provides constructs that allow us to *unambiguously* state the relationship between two things (regardless of whether the things are in the same ontology or two different ones). If we believe that GO:0000010 (trans-hexaprenyltranstransferase activity) and RHEA:20836 are equivalent, we can write this as:

```
GO:0000010 owl:equivalentClass RHEA:20836
```

https://douroucouli.wordpress.com/2019/05/27/never-mind-the-logix-taming-the-semantic-anarchy-of-mappings-in-

OBO Foundary as Example

RELATION TO TIME / GRANULARITY	CONTINUANT				OCCURRENT
	INDEPENDENT		DEPENDENT		
ORGAN AND ORGANISM	Organism (NCBI Taxonomy)	Anatomical Entity (FMA, CARO)	Organ Function (FMP, CPRO)	Phenotypic Quality (PaTO)	Organism-Level Process (GO)
CELL AND CELLULAR COMPONENT	Cell (CL)	Cellular Component (FMA,GO)	Cellular Function (GO)		Cellular Process (GO)
MOLECULE	Molecule (ChEBI, SO, RnaO, PrO)		Molecular Function (GO)		Molecular Process (GO)

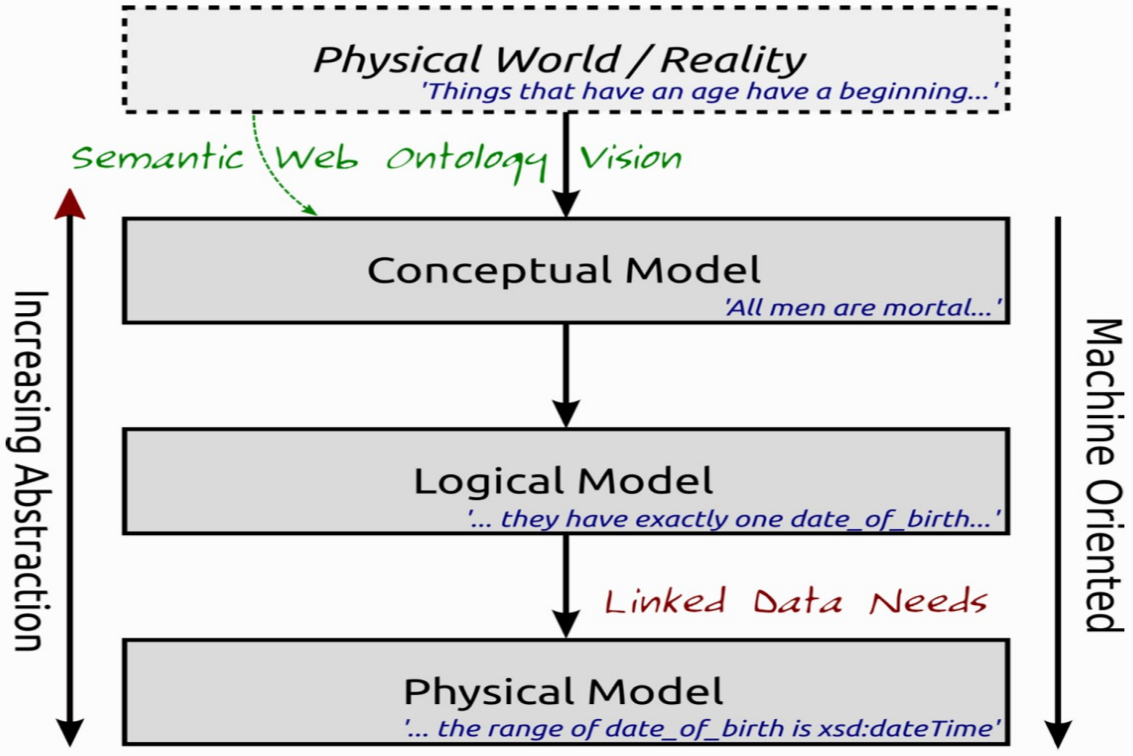
Ontology Engineering: A View from the Trenches

WOP 2015 Keynote

Krzysztof Janowicz

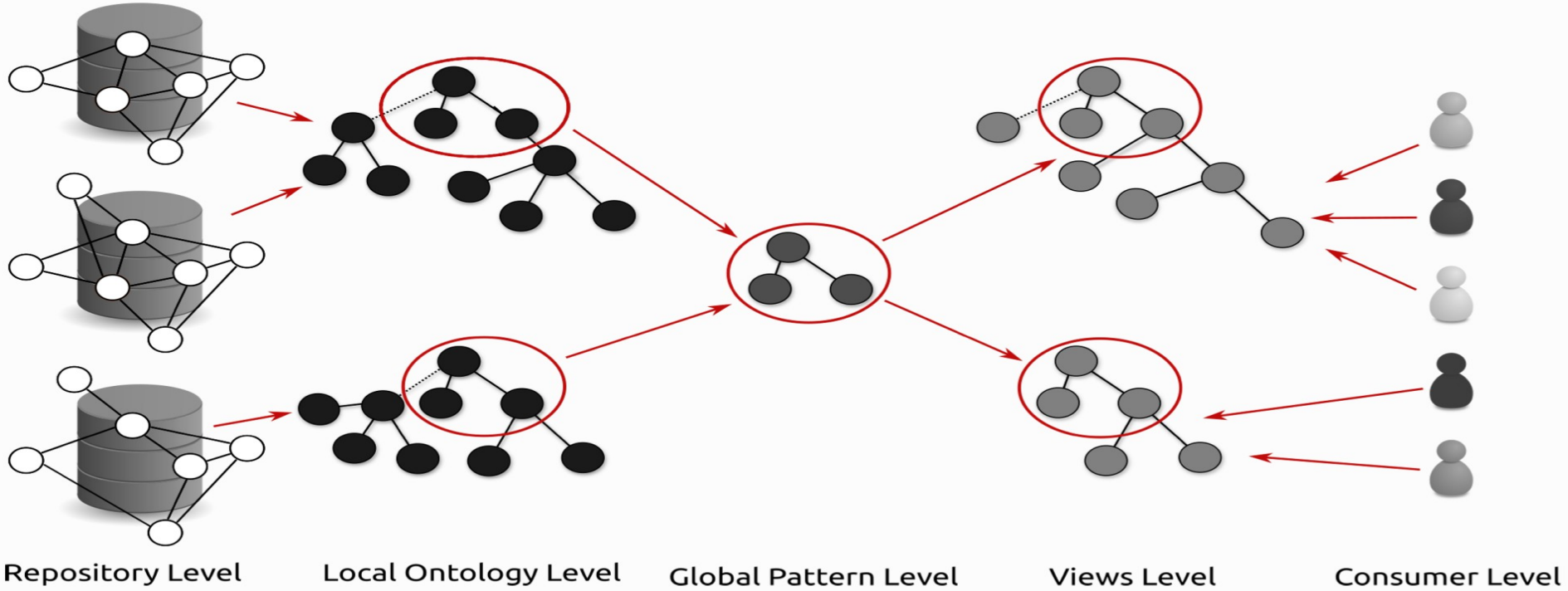
STKO Lab, University of California, Santa Barbara, USA

Is There a Common Cause to These Problems?



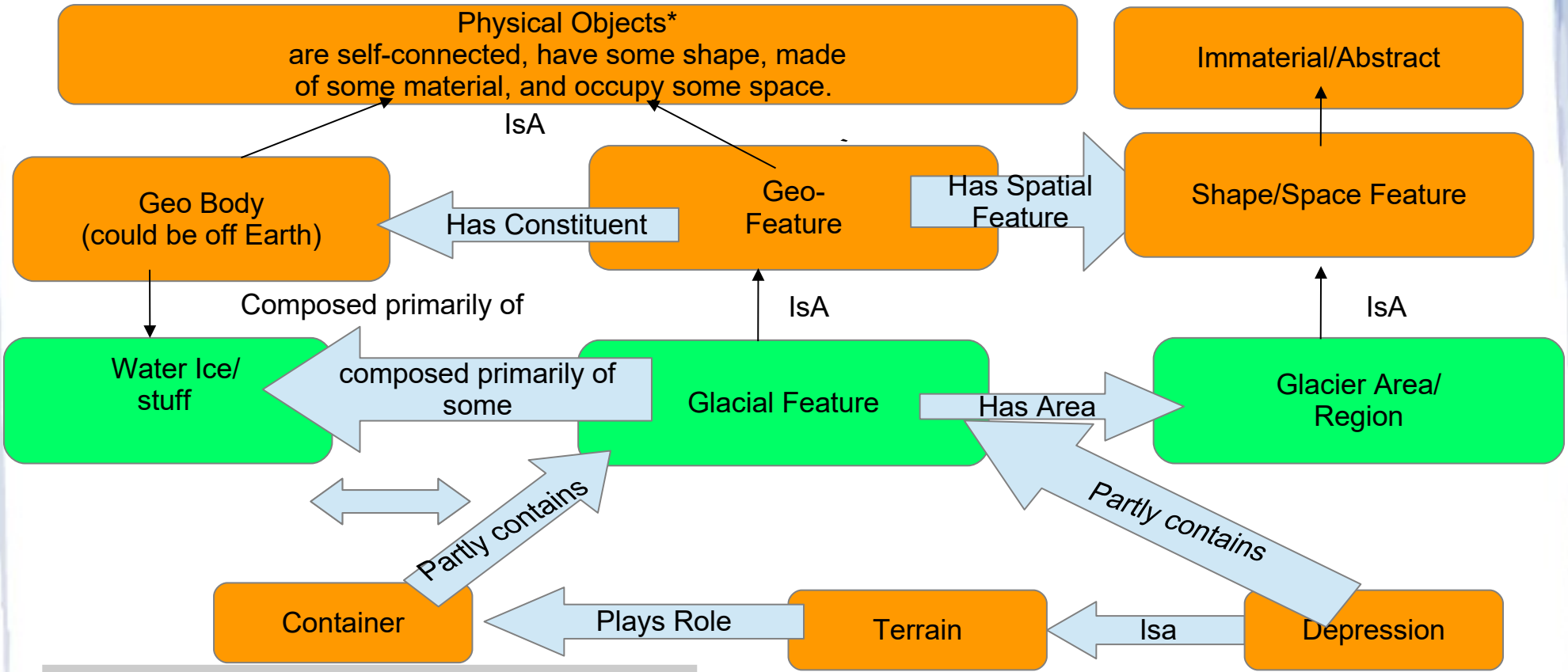
*We will revisit the **age** example at the end of this talk.*

Envisioned, Pattern-based Arcitecture (Horizontal)



Patterns act as **fallback** level that ensures **minimal interoperability** while preserving **heterogeneity** (i.e., local, repository-specific ontologies can differ).

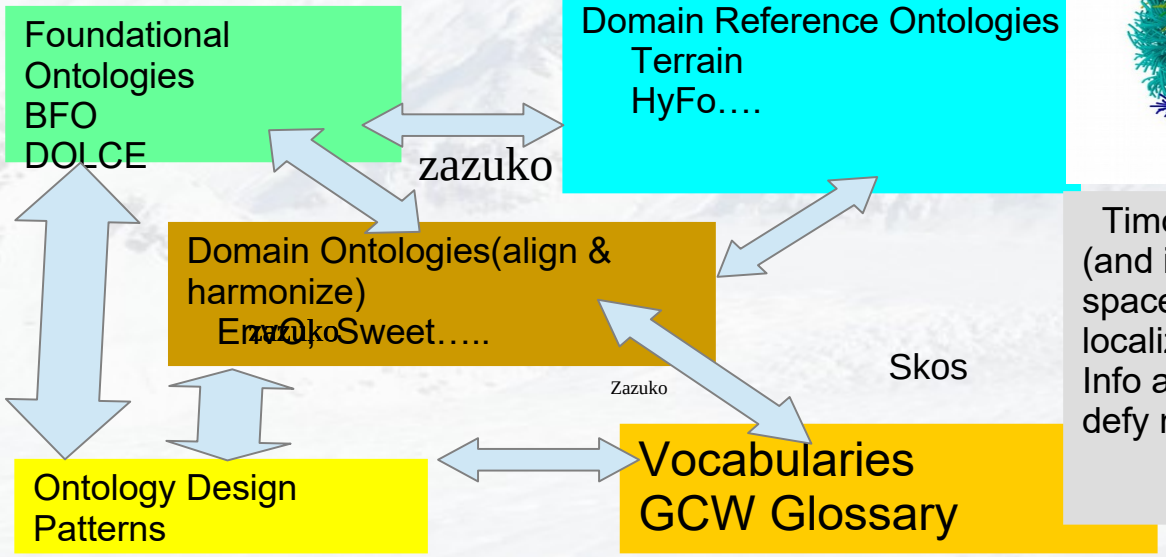
Foundational Distinctions: Spatial regions & physical bodies are distinct entities maybe a problems for an “Integrated” model like EnvO relying on BFO. Some domains they lack easily used upper domain reference model and/or ODPs

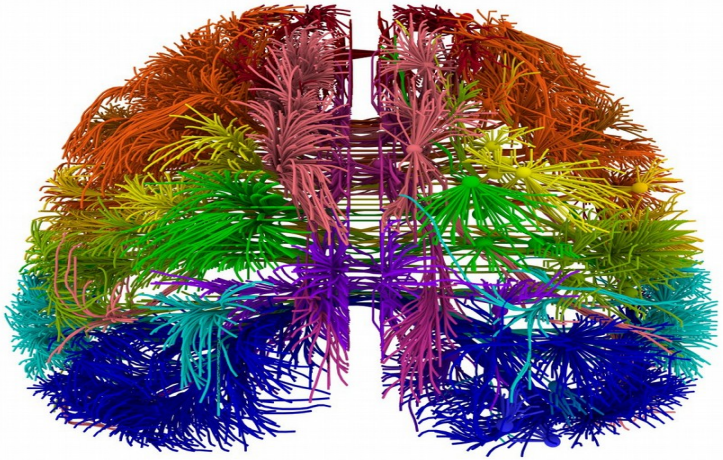


* Ontology of Solid Physical Objects:
colore.oor.net/opo/opo.clif

Challenge of Large Systems of Knowledge - “HyperModel Objects”

We need all these varied disparate model entities making a model hyperobject system





Timothy Morton uses the term to explain objects (and info) so massively distributed in time and space (but related) as to transcend their localization.
Info about these objects is a problem that seems to defy not only our control but also our understanding.

The next steps include the following:

- Distribute notes from the meeting and provide access to slides
- Start of monthly calls the third Wednesday of the month at 2 pm EST. The first meeting to be Jan. 22.
- Using ESIP slack channel for discussion
- Panel session on semantic harmonization at US2TS in Raleigh March 9-11, 2020.
- Session at Summer ESIP meeting July 14-17 (Vermont).
- Possible virtual “Vocamp” on topic of domain harmonization..