**GeoSciML SWG**

**Minutes of Dublin F2F and teleconference, 23 June 2016**

**Attendees:**In person - Eric Boisvert (GSC), Alistair Ritchie (Landcare NZ), Bruce Simons (CSIRO), Marcus Sen (BGS), Harvey Thorleifson (Minnesota Uni), Francois Robida (BRGM), Mikael Beaufils (BRGM), Sylvain Grellet (BRGM), Christelle Loiselet (BRGM) , *some others…?*Teleconference - Ollie Raymond (GA), Tim Duffy (BGS)

**Conceptual Model**

GSML-Lite has to be described as a encoding of the same conceptual model, or as a transformation of GeoSciML. The spec had to be fix to reflect either or both. It now does both by introducing NADM as a conceptual model and offers property mapping between GSML and Lite. It has been concluded that Lite is a transformation.

GeoMorphologicalView had issues because it has properties that did not exists (landformClass and landforClass\_uri) , but after further discussions, we realised it should be the type (natural vs anthropogenic), so it is a transformation.. But to improve linkage, we need to rename the property name (action from Carlo)

**Terms and definition: complete ?**

Should avoid duplication between normative statement and text around it ***\*DECISION: Yes\****

**Vocabularies as HTTP URI**

Is this restricted to XML representation only (as opposed to GeoJSon for example)? *[Ollie - Would it be a problem if we recommended that Json should use uri’s for vocabulary concept identifiers too?]****\*DECISION: Recommend that all encoding implementations use HTTP URI.\****

Impact: Querying on GeoSciML will depend of the format

Recommend that we move Linked Open Data into its own Requirement class that might or might not be implemented [*Ollie - yes*] ***\*DECISION: Yes\****

Portrayal

* Identifier : recommendation URI [*Ollie - yes*] ***\*DECISION: SHALL be formatted as URI\****
* Vocabulary (\_uri) : mandatory (or nil) , should resolve (eg: HTML) [*Ollie - yes*] ***\*DECISION: These should be optional attributes, not mandatory - Recommendation for community or practice to profile if they want them to be mandatory. \* OLLIE - 12 Aug 2016 - need to fix UML and UML scope notes to reflect this decision!!!***
* Specification\_uri : optional , should resolve [*Ollie - yes. see decision at* [*http://external.opengeospatial.org/twiki\_public/GeoSciMLswg/MeetingMinutes*](http://external.opengeospatial.org/twiki_public/GeoSciMLswg/MeetingMinutes)*. The UML cardinality and schema currently says not mandatory, but we need to change the scope notes that still describe “specification\_uri” as mandatory. ie, GeologicUnitView, ContactView, GeomorphologicUnitView, Shear..View*] ***\*DECISION: Optional - need to correct scope notes. \****

**8.2.8 Identifier**

Should it be a requirement that URI point to an XML representation (or a XML be one of the mandatory representation) ? WFS has this requirement (GML is a mandatory format), but some groups try to have this rule removed. ESRI REST essentially works in JSON, Future OGC RESTFul is on the drawing board.

See above [*Ollie - no for “identifier”. But yes for vocabulary concepts]*

**Text strings in GSML Lite**p. 35 , first paragraph:

*. In robust services the free-text fields will contain well-structured summaries of data in a format suitable for reading by the intended users*

This comes from the scope note and I don’t understand what “robust services” and “well-structured” are.

Rephrase and turn into recommendation. [Ollie - “Best practice is that free text fields will, where possible, contain well-structured summaries of data in a format suitable for reading by the intended users. For instance, an agreed common format like comma-delimited values should be adopted by user communities.”] ***~~\*DECISION: Defer to later\*~~ DONE (used Ollie’s sentence)***

**GSML Lite**

General discussion about gsml-lite identifiers :

Often the identifier of the MappedFeature

Is it unique (a primary key) for the dataset ? yes *[Ollie - yes]*

Could be a URI, but does have to. *[Ollie - yes]*

Ie: URI not mandatory -> any structure

Versus

URI mandatory (as a syntax), but might or might not resolve. [Ollie - prefer this. Personally, I think a vocabulary concept URI “SHALL resolve” - there are enough vocabularies out there now to expect that - but maybe we should use “SHOULD resolve” just because we are trying to make the user entry point for GSML-Lite as low as possible?]

Larger discussion, Why URI mandatory then. *[Ollie - Suggest URI format is only mandatory for vocabulary terms and specification\_uri, not for identifier?]*

**Specification\_uri**

Even Larger discussion, why specification\_uri for provider who don’t have complex GSML ? Meant to be a global identifier for the GeologicFeature *[Ollie - specification\_uri could resolve to a HTML page in cases where a complex XML GeologicFeature is not available. For example:* [*http://dbforms.ga.gov.au/pls/www/geodx.strat\_units.sch\_full?wher=stratno=25947*](http://dbforms.ga.gov.au/pls/www/geodx.strat_units.sch_full?wher=stratno=25947)*]*

***\* DECISION: Specification\_URI not mandatory \****

***\*DECISION: Specification\_URI should, but does not have to, resolve to a complex GSML feature. That is too much to expect for low-end data providers. \****

**Nillable**

We still have several clauses that mention that a nilReason should be provided , but it’s not consistent. Not sure it’s GML SF-0 valid and if this is doable in, say, MapServer*. [Ollie - nilReason can be done in a simple URL using OGC nil URL’s - http://www.opengis.net/def/nil/OGC/0/missing . Shouldn’t be a problem for GSML-Lite mandatory vocabulary concept URI’s]* **\*DECISION: Do Ollie’s comment\***

We just don’t understand this clause (8.4.3.1 Null values in GML SF-0) - **NOTE that section numbers are changed because we added a LOD ReqClass (8.3 -> 8.4)**

**8.3.3.12 positionalAccuracy**

In GSML Lite, it’s a string.. Confirm it’s a text description of the positional accuracy and not just a number.

Ok, it’s a quantity in a string representation*.*

*[Ollie - examples given in the scope notes could be better. For example, from the new ERML-Lite UML: “ Text description of the accuracy of the feature location. (eg, accurate, approximate, diagrammatic, indefinite, unknown, 5 metres, 1 kilometre)”]*

**8.3.9.9 observedValue DONE**

*. Although this field is implemented as a character string to allow reporting various type of values, the value may be numeric (e.g.; 235 degrees, 50 ppm) or textual (e.g.; red). Unit of measures shall be reported in observedValueUom*

The original note mentioned only string to “conform with simple feature requirements,” – that’s not true. Other literal types are valid. But it also said “it can be encoded as a swe:Record.” Must be Ollie. I don’t think it’s SF-0. Ok.. fixed. It was just a comment. [Ollie - note that UoM of numeric values are delivered in the observedValueUom property, so the comment on observedValue should not include UoM in the examples. Change comment to: *“This field is implemented as a character string to allow reporting various type of values, the value may be numeric (e.g.; 235) or textual (e.g.; red). Units of measure shall be reported in observedValueUom”.*

**8.4.1.1.1 observationMethod *DONE***

*. For a borehole, the GeologicFeature observation method specifies how the geologic properties were determined (e.g., visual observation, or standard corporate logging procedure (described in detail somewhere else)).*

*(…)*

*Scoped name because intention is asserted by author of the data instance.*

*Just delete [Ollie - yes, delete it]*

Eric - I don’t understand (from scope notes)

**8.4.1.1.6 relatedFeature : *DECISION: NO ACTION***

*. There is always a single source and a single target for a given FeatureRelation (which is abstract in GeoSciML Basic).*

This seems to imply that a new AbstractFeatureRelation must be created (one of its subtype anyway) for any new relation between 2 features. But AbstractFeatureRelation is a “Type”, which has an identity, and therefore can be pointed to. Since relatedFeature can do a xlink:href, it technically means it can point to an existing relation and therefore break the “single source and single target” rule. What’s wrong here , the rule or the encoding ?

*will allow a “by reference” value using a pointer (for example xlink:href) to an external instance.*

That an interesting side effect. Basic cannot instanciate a AbstractFeatureRelation, but can xlink:href to one.

At the end of the day you can’t prevent people putting garbage in their datasets. ☺ ☺

**p. 82, top** **DONE**

*• the specific bounded occurrence, such as an outcrop or map polygon*

*• the association with a Geologic Feature (legend item) provides specification of all the other descriptors*

*• the association with a Sampling Feature provides the context and dimensionality*

Eric - I don’t understand

Need to rework from the scope notes. ***DECISION - delete these dot points. Incomprehensible****.*

***8.4.1.3.2 rank******DONE***

*The property rank:RankTerm shall contain a term that classifies the geologic unit in a generalization hierarchy from most local/smallest volume to most regional. Scoped name because classification is asserted, not based on observational data.*

I don’t understand the last sentence – delete if nobody remembers [Ollie - delete. Unnecessary words.]

***8.4.1.4.1 role******DECISION: NO ACTION***

*The role:GeologicUnitHierarchyRoleTerm property shall provide a term describing the nature of the parts, e.g. facies, stratigraphic, interbeds, geographic, eastern facies.*

Some roles, such as stratigraphic hierarchy, and probably most of them, might require extra rules, such as hierachyLink shall not be cyclic. (This falls to the category of : to what extent do we impose rules that are otherwise common sense ?)

Probably overkill [Ollie - I agree]

***8.4.1.6.2 Purpose DONE***

*The purpose:DescriptionPurpose property shall provide a specification of the intended purpose/level of abstraction for the given EarthMaterial. The intent is the same a GeologicFeature’s purpose (see 7.4.1.1.2) and it shares the same vocabulary (instance, typicalNorm, definingNorm).*

There is a small “semantic” risk of reusing the same type. If a new purpose emerge in either EarthMaterial or GeologicUnit, it automatically become available for the other class (because they use the same “DescriptionPurpose”)

EM list actually has a different term (Values: Instance, TypicalNorm, ***IdentifyingNorm***.)

Is it a typo in scope notes of EM, or should they be two differents vocabs ? ***DECISION - Typo, I think. use definingNorm***

Should be bring back the “known absence” discussion ? No. ;-)

8.4.3.2 Contact ***DONE***

*Bedding measured as discrete surfaces in the case that those are the feature of interest (e.ulg. individual cross set surfaces for paleocurrent analysis) should be represented here.*

Not sure what this is for – anyone remember what this means or delete *[Ollie - I think I get what is trying to be said here… “Contacts may include discrete bedding surfaces that separate small scale geologic units (eg, surfaces between individual cross-bedding sets in paleocurrent analysis”].* ***\*DECISION: Delete comment. Not needed\****

#### Fold *DONE*

A fold is formed by one or more systematically curved layers, surfaces, or lines in a rock body. Fold denotes a structure formed by the deformation of a GeologicStructure to form

We are pretty sure it should be “GeologicUnits” **\*DECISION: Steve explained 🡪 geologic structure\***

**8.4.5.1.2 collectionType *DONE***

The collectionType:CollectionTypeTerm property shall be a term from a controlled vocabulary describing the type of collection.

We need a better description. Basically describes the content of the collection eg: geologicalMap, boreholes. *[Ollie - eg, geologic map, borehole log, 3D model]*

**8.4.6.2 GSML\_PlanarOrientation** ***DONE*** – removed overkill requirements

A planar orientation is composed of two values; the azimuth (a compass point) and a dip (the angle from the horizontal). Polarity of the plane indicates whether the planar orientation is associated with a directed feature that is overturned, upright, vertical, etc. **There are several conventions to encode a planar orientation and this specification does not impose one but provides a convention property to report it. It must be noted that allowance of different convention makes manipulation of the data more difficult.**

Should we have a recommendation that community stick to a single convention (eg: OneGeology uses this convention)

I suggested some constrains in the document as requirement for numerical values (90 degrees max, etc),

Overkill ?? probably [*Ollie - maybe reword the bold bit just to say: “****It is recommended that user communities adopt a single measurement convention****.” We’ll never get the Americans to agree to change from strike/dip to dip/dipDirection ;-P , so there’s no point recommending which convention*.]

***\*DECISION: Keep the comments to community recommendation, not specification. Remove requirement.\****

**p. 113** ***NO ACTION***

The map also has a cross section through the same Ttv unit (Figure 38) showing an example of a non-map mapping frame.

There should be a convention that the SRS of the geometries on the cross-section shall be a reference to the plane that makes the cross-section.

Don’t mention it ? Or. There are no known best practice on how to deal with C-S CRS *[Ollie - yikes. Can of worms. Leave it out for now.]*

**8.5.2.4.3 shape**

A good example that a property name alone is not a good identifier. Most of the “shape” I’ve seen in model are actually the GM\_Object. BTW, if I understand correctly, the new 19109 will forbid duplicate property names in an application schema (in the same namespace). Is this important enough to rename ? [*Ollie - can rename to particleShape. Shouldn’t cause a problem. It’s a minor change in a non-core part of the model.]*

***\* DECISION: change to “particleShape” \* ?? was this the decision??***

**8.5.2.4.3 size**

The property size (SWE::QuantityRange) shall report the size that specifies particle grainsize. Values may be reported using absolute measurements (e.g.: range, mean, median, mode, maximum) ~~or as descriptive terms from a schema appropriate to the type of Compound Material (e.g.: the Udden-Wentworth sheme for clastic sedimentary rocks - silt, sand, gravel; volcaniclastic rocks - ash, lapilli, bomb; crystalline rocks - fine, medium, coarse, cryptocrystalline).~~

This description does not fit with a swe:QuantityRange

Change encoding or description *[Ollie - description is left over from GSML v2. Need to change the description to leave out the “descriptive terms” bit.]*

***\* DECISION: Change old scope notes to reflect SWE::QuantityRange \****

**8.5.2.5.1 role** ***NO ACTION***

The role: ConstituentPartRoleTerm property shall contain a term from a controlled vocabulary that describes the role a ConstituentPart plays in a CompoundMaterial aggregation. The same EarthMaterial may occur as different ConstituentParts playing different roles within one CompoundMaterial. For example, feldspar may be present as groundmass (“groundmass” is a ConstituentPart::role) and as phenocrysts (“phenocryst” is another ConstituentPart::role) within a single igneous rock.

Both ConstituentPart.role ? Is this intentional ? ok. Looks correct. *[Ollie - OK]*

/req/gsml4-extension/contact-boundary ***DONE***

A contact SHALL have 2 and only 2 two instances of GeologicFeatureRelation which roles are boundaries and targets are GeologicUnit. Still struggle to make a sentence that makes sense here. Need someone who masters English to write this one [Ollie - Rethinking this one. For instance, one contact may form the top of a group, formation, and member GeologicUnit. So potentially there could be more than 2 GeologicUnits either side of a Contact.]

***\*DECISION: Take it out\****

**Figure 67: *NO ACTION***



Would’t it be more informative to have ShearDisplacementStructure in this diagram (to show that SDSDescription adds description to SDS*) [Ollie - maybe. But context diagrams are, AFAIK, are meant to just show the immediate associations to a class, not transitive ones. The summary diagrams show the broader context.]*

**8.6.1.1.1 primaryGuidingCriterion** ***NO ACTION***

The property primaryGuidingCriterion:Primitive::CharacterString shall contain a description of the primary criterion used to establish this stratigraphic point.

Circular definition ? This one is from Simon [Ollie - I’ve seen worse. Leave it] \*DECISION: Leave it\*

**8.6.2.2 TimeOrdinalEra** ***NO ACTION***

TimeOrdinalEra is a period of time between two boundaries. **The association of an era with a stratotype is optional. In the GSSP approach recommended by ICS for the Global** Geologic Timescale, Unit Stratotypes are not used. Rather, the association of an Era with geologic units and sections is indirect, via the association of an era with Boundaries, which are in turn tied to Stratotype Points, which occur within host Stratotype Sections. TimeOrdinalEra can be composed or other era and organized into an arbitrarily deep nested tree.

In bold, this is the very same description than GeochronologicEra. I suspect this description is for GeochronologicEra because it references stratotype.

**8.6.2.2.3 start** ***DONE***

The start property shall be an association to a TimeOrdinalEraBoundary that defines the start of the era.

Is there a rule that says that start must be younger than end ? [Ollie - note in UML on the ‘start’ association says “Older time boundary of an era”] ***\*DECISION: Add requirement\****

**BOREHOLE**

I missed the fact that BoreholeInterval has both distance and Geometry, so the geometry SRS discussion is still an issue

A BoreholeInterval is a special kind of Mapped Feature whose shape is 1-D interval and uses the SRS of the containing borehole. The "mappedIntervalBegin" and "mappedIntervalEnd" elements are included here as a measure to overcome problems with the delivery and queryability of 1D GML shapes via the "shape" property.

***\*DECISION: Amend description to use distance\****

Does this mean that Log interval are absolute coordinate (elevation) ? Not what map interval says (relative) *[Ollie - we separated out the begin and end values because FES couldn’t query a GML number pair.]*

Reword “special kind”, because “mixin”.

It does not have explicit SRS. Must rephrase as “distance”. [Ollie - OK]

**Flag HTTP URI:** ***DONE ?***

<http://www.ietf.org/rfc/rfc2616> - only an issue for gml:identifier.

Is this still how it should be done. This codeSpace == the content is a URI

GeoSciML has adopted a practice of flaggin HTTP URI using codeSpace ...

\*DECISION: gml:name codespace is used for naming authority, not for method of interpreting the identifier. Get rid of the rfc requirement bit.\*

**9.2.1 XML document validation** ***DEFER***

An XML instance shall validate to both the XSD and schematron rules provided by this specification for each of the XML requirements classes.

As we painfully experienced, the location of the schema is almost as important as its content. So the schema location could almost be defined as “normative”. The other way to deal with this is to have Oasis Catalog with the schema that allow 1) schemas to be distributed with the spect and 2) ensure that the externalities are pulled from the right place. THESE (XSD+Oasis catalog) would be the normative XSD

**Decide if we publish on schema.opengis.net or schema.geosciml.org ?**

***\*DECISION: Ollie to get confirmation from Scott Simmons.\**** General discussion at the meeting was that publishing the GeoSciML v4.1 canonical schemas at schemas.opengis.net will be done and that while there will be some pain for users in a transition period between schemas.geosciml.org to schemas.opengis.net, it can be managed.

**Xsd:any**

Right now, the schema does not state if the user properties are in the same namespace or another namespace.

As is (we don’t impose anything, best practice, to avoid future conflicts, should use ##other)

If we force another namespace , it should be

<xs:any namespace="##other" processContents="lax" minOccurs="0" maxOccurs="unbounded"/>

Otherwise, if we want to force the same namespace, it should be

<xs:any processContents="lax" minOccurs="0" maxOccurs="unbounded"

namespace="##targetNamespace"/>

*[Ollie - do we really have to specify this? Here’s an example of a* [*gsmlp:BoreholeView*](http://schemas.geoscience.gov.au/geosciml) *schema extended for use with Geoserver. In practice, you have to rewrite the schema like this if you want to add extra attributes, because Geoserver can’t deliver “any:lax” by pointing to another schema.]*

***\*DECISION: Recommend different namespace\* DONE***  *(Ollie would like to see a Geoserver implementation of this before we finalise this decision.)*

**Discussion about links between GeoSciML and GeoSciML-Lite**

***\* DECISION:*** *Maintain description in spec documentation that Lite is a different mapping of the same conceptual model, but can also be derived by a transformation of the GeoSciML logical model. \**

**Geomorphology - LandformClass *(Carlo needs to review this) DONE***

***\* DECISION:*** *In order to maintain better link to the GeoSciML-Basic Geomorphology model, change Landform Class to GeomorphologicFeatureType and GeomorphologicFeatureType\_uri. Vocabulary is only two terms: natural geomorphologic feature, and anthropogenic geomorphologic feature. \**

**Publishing and voting timeframe for spec documentation**

From Scott Simmons to Ollie (23 June 2016): *…..(Ollie to insert text here when his email server gets back up)*