Towards an ontology based conceptual model …

… establishing maximum interoperability for interactive and distributed processing of geoscientific information

Rainer Haener

Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences, CeGIT, Potsdam, Germany

Henning Lorenz

Uppsala University, Department of Earth Sciences, Sweden

Sylvain Grellet, Marc Urvois

The French Geological Survey, Orleans, France

Eberhard Kunz

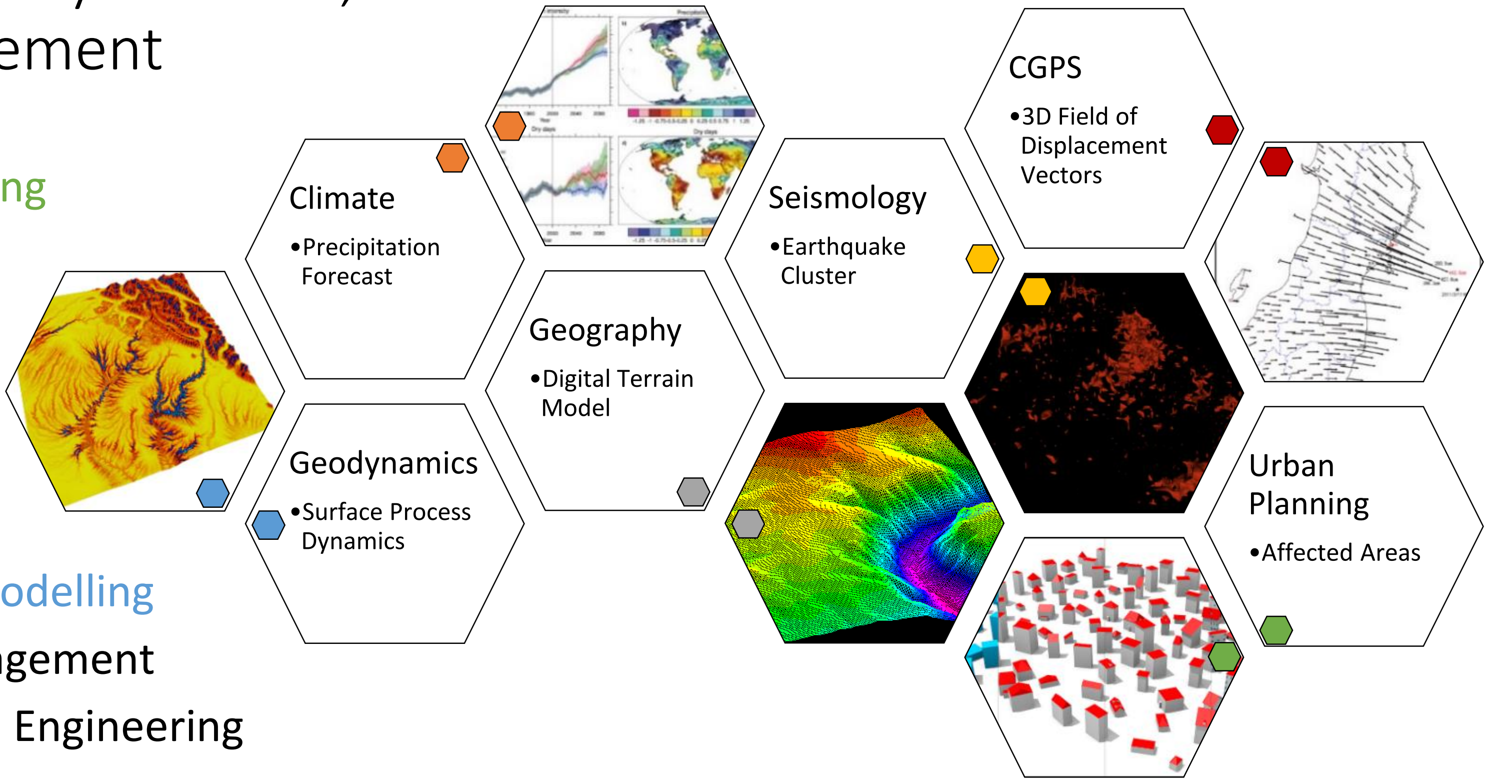
Federal Waterways Engineering and Research Institute, Karlsruhe, Germany

Outline

- Introduction
- Design Principles
  - Composable Infrastructures
  - Dependency Injection
- Formal Approach
  - Semantics
  - Meta-Model
- Implementation
  - Borehole Interoperability
  - Geological Classification
- Conclusion

The “running” Interoperability Scenario, Risk Management

- Urban Planning
- Meteorology
- Climatology
- Geology
- Geography
- Geophysics
- Numerical Modelling
- Facility Management
- Geotechnical Engineering



GFZ
Helmholtz Centre
POTSDAM

HELMHOLTZ
ASSOCIATION

Introduction: Conceptual Model, anticipated

Meta-Model (Pragmatic Interoperability)	OGC:	Sensor Web Enablement (SWE), SWE Common
	ISO:	19139 (Metadata), 19136 (GML), 19148 (Linear Referencing)
	W3C:	RDF/XML, OWL
Semantics (Semantic Interoperability)	Identifiers, Classifications, Characteristics, and Categories	
	Domain Driven Vocabularies	
	Concepts & Ontologies	
Architecture (Dynamic Interoperability)	Composability, Event-Driven & Service-Oriented Architectures, REST-Based	
	Dependency Injection (Standardised Interfaces)	

GFZ
Helmholtz Centre
Potsdam

HELMHOLTZ
ASSOCIATION

The Hollywood Principle, aka Inversion of Control, or Dependency Injection (DI)

Source: [AZ Quotes]

A Landslide, observed (numerically modelled)

```xml
<?xml version="1.0" encoding="UTF-8"?>
<om:OM_Observation gml:id="landslide">
    <om:procedure          xlink:href="ground instability model"/>      <!– Process (Sensor) Identification -->
    <om:observedProperty   xlink:href="landslide"/>                     <!– the Phenomenon observed -->
    <om:featureOfInterest  xlink:href="landslide-geometry"/>            <!– Sampled Feature -->

    <om:result>
        <gml:Surface>                                                   <!– Dependency (Geometry) Injection -->
            <gml:patches>
                <gml:PolygonPatch>
                    <gml:exterior>
                        <gml:LinearRing>
                            <gml:posList>0 1 1 2 3 5 8 13 21 34 55 89 144</gml:posList>
                        </gml:LinearRing>
                    </gml:exterior>
                </gml:PolygonPatch>
            </gml:patches>
        </gml:Surface>
    </om:result>

</om:OM_Observation>
```

An observation container, completely composed of Dependency Injections (e.g., procedure, result), which in turn are subject to a (Semantic) Registry

GFZ Helmholtz Centre Potsdam

Adapted from [Presentation EGU 2018, ESSI2.1]

HELMHOLTZ ASSOCIATION

A Building, observed as well

```
<?xml version="1.0" encoding="UTF-8"?>
<om:OM_Observation gml:id="townhall">
    <om:procedure          xlink:href="lidar"/>                       <!– Process (Sensor) Identification -->
    <om:observedProperty   xlink:href="townhall"/>                    <!– the Phenomenon observed -->
    <om:featureOfInterest  xlink:href="building ground plan"/>        <!– Sampled Feature -->

    <om:result>
        <gml:Surface>                                                 <!– Dependency (Geometry) Injection -->
            <gml:patches>
                <gml:PolygonPatch>
                    <gml:exterior>
                        <gml:LinearRing>
                            <gml:posList>4 8 15 16 23 42</gml:posList>
                        </gml:LinearRing>
                    </gml:exterior>
                </gml:PolygonPatch>
            </gml:patches>
        </gml:Surface>
    </om:result>

</om:OM_Observation>
```

An observation container, completely composed of Dependency Injections (e.g., procedure, result), which in turn are subject to a (Semantic) Registry

GFZ Helmholtz Centre Potsdam

Adapted from [Presentation EGU 2018, ESSI2.1]

HELMHOLTZ ASSOCIATION

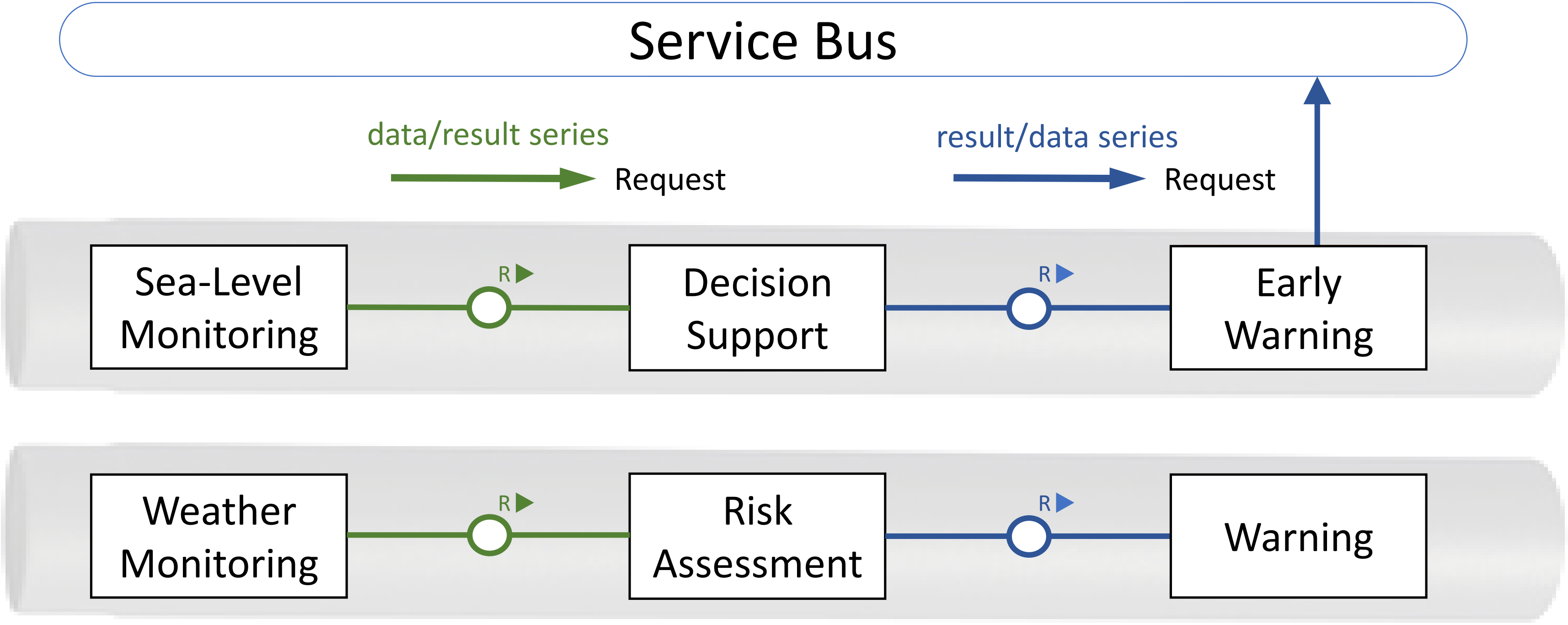
Composability vs. Integration

Favouring Composition over Inheritance (Integration) is essential for establishing Interoperability! “When systems are interoperable, they have the ability to not only share information, but to interpret incoming data and present it as it was received, preserving its original context”, Bobby Roberts

GFZ
Helmholtz Centre
POTSDAM

HELMHOLTZ
ASSOCIATION

Integration: Bound Components (Stove-Pipes)



Dependency Injection, Composability and Composable Infrastructures

Processing-Components (e.g., Risk Assessment) "delegate" publishing (e.g., Warning), either by "calling" (don't call us, we call you) suitable services or - according to a choreography - broadcast messages and let dedicated services autonomously decide according to user and stakeholder demands (e.g., Early Warning):

An Early Warning Focal Point for example, has subscribed to certain crisis topics. It receives notifications directly addressing ("calling" the Focal Point) these topics. The Focal Point DOES NOT need to know any (implementation-) details about the components used.

GFZ Helmholtz Centre Potsdam

HELMHOLTZ ASSOCIATION

Composability: Autonomous Components

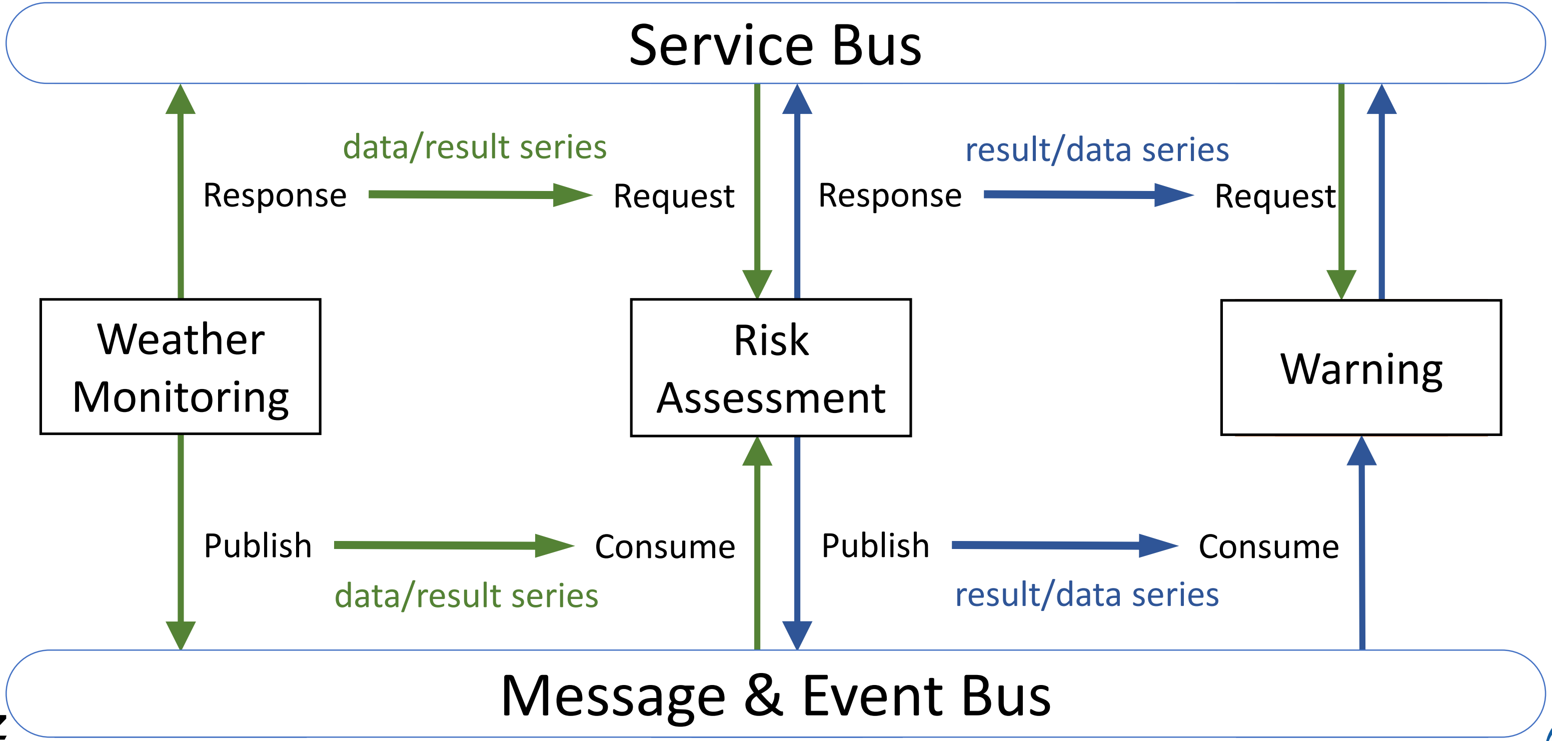


Diagram based on the Fundamental Modelling Concept (FMC) Adapted from [Presentation EGU 2017, ESSI2.2]

Semantics

An indispensable part of a Conceptual Model is detailed semantics, which not only requires terms from Domain-Controlled Vocabularies, but also Ontologies, providing qualified statements about the relationship between data and associated concepts. This is of major importance for evolutionary systems that are able to comprehend and react to state changes.

GFZ Helmholtz Centre Potsdam

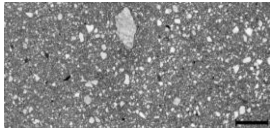
HELMHOLTZ ASSOCIATION

Semantics Use Case, Data Mining

Search for Observations that …		
	contain certain Classifications	Grain Size
		Characteristic Sequences (Tertiary)
	belong to certain Categories	Metamorphic Rocks
		Mineral Deposits
	cover certain Features of Interest	Landslide-Prone Terrain
		Building Sites
	cover certain Concepts	Seismic Velocity reflects Rock-Formation
		Geotechnical Suitability
	are based on certain Processes	Tomography
		3D-Modelling

Semantics Example, Classification (Registry)

```
<?xml version="1.0" encoding="UTF-8"?>

<swe:QuantityRange
        definition= "http://www.gfz-potsdam.de/dictionaries/geology/petrography/grainSize#gU"
        id="gU"

        xmlns:swe="http://www.opengis.net/swe/2.0">
```



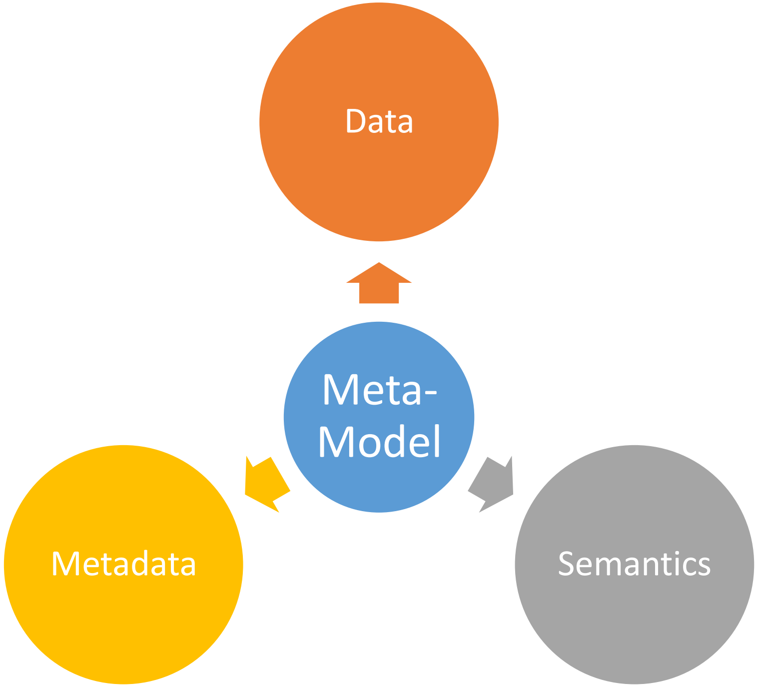
```
        <swe:identifier>http://www.gfz-potsdam.de/dictionaries/geology/petrography/grainSize#gU</swe:identifier>
        <swe:label>Coarse Silt</swe:label>

        <swe:description/>
        <swe:quality/>
        <swe:nilValues/>

        <swe:uom code="mm"/>
        <swe:value>0.02 0.05</swe:value>

</swe:QuantityRange>
```

Comprehensive specification
(classification & characterisation)
of a rock property, completely based on
Domain Controlled Vocabularies.

... Anticipating the utilisation of
SWE Common as part of the Meta-Model

GFZ Helmholtz Centre Potsdam

Helmholtz Association

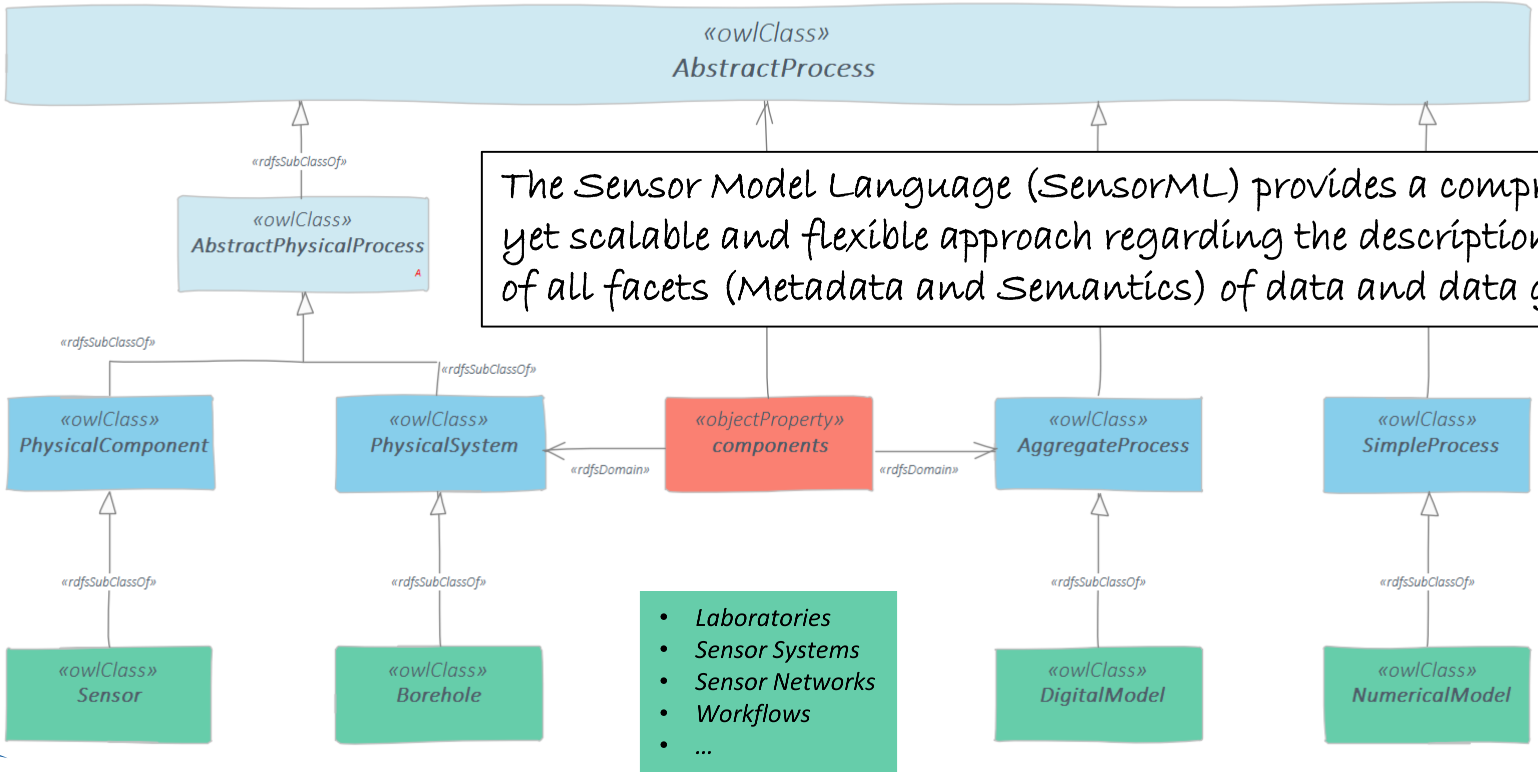


Meta-Model

The goal is the abstraction of all facets of geological information in order to establish a common and uniform model for interactive information processing. A Meta-Model abstracts metadata, data, and semantics as well as their relationships. It uses a formal approach (encoding) in order to establish scalable, re-useable, and globally available information representation.

SensorML: Everything is a Process is a Sensor …

The Sensor Model Language (SensorML) provides a comprehensive, yet scalable and flexible approach regarding the description of all facets (Metadata and Semantics) of data and data generation



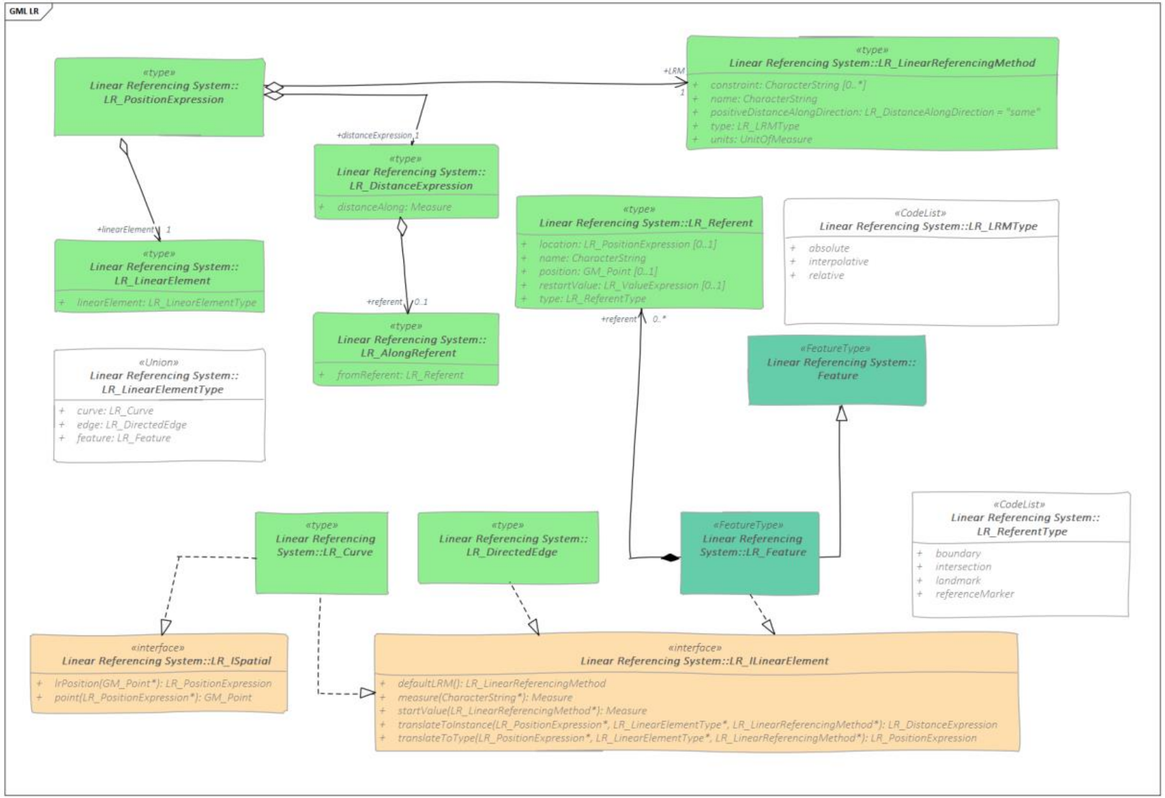
GFZ
Helmholtz Centre
POTSDAM

HELMHOLTZ
ASSOCIATION

Linear Referencing (LR)

“Linear Referencing Systems enable the specification of positions along linear objects” (OGC). Thus, they offer a flexible mechanism to reference both, spatial and temporal observations (so called linear events) to a feature of interest. The according model allows the utilisation of data, based on the pattern Dependency Injection

Linear Referencing: Main Concept



GFZ
Helmholtz Centre
POTSDAM

HELMHOLTZ
ASSOCIATION

Linear Referencing examples, applying DI

```xml
<?xml version="1.0" encoding="UTF-8"?>
<bml:GeoValue
    xmlns:bml="http://www.gfz-potsdam.de/bml/4.0"
    xmlns:gml="http://www.opengis.net/gml/3.2">

    <bml:railwayRoute >                <!– a linear feature, referenced by the event -->
        <gml:Curve>
            <gml:segments/>
        </gml:Curve>
    </bml:railwayRoute >

    <bml:event>
        <bml:GeoVal_Event>
            <bml:location>             <!– 1 meter from start along the railway -->
                <bml:AtLocation>
                    <bml:distanceAlong uom="m">1</bml:distanceAlong>
                </bml:AtLocation>
            </bml:location>

            <bml:value>
                <bml:GeotechnicalParameter>        <!– Dependency Injection -->
                    <bml:extent uom="m">23</bml:extent>
                    <bml:vp uom="m/s">137</bml:vp>
                </bml:GeotechnicalParameter >
            </bml:value>
        </bml:GeoVal_Event>
    </bml:event>

</bml:GeoValue>
```

```xml
<?xml version="1.0" encoding="UTF-8"?>
<bml:Borehole
    xmlns:bml="http://www.gfz-potsdam.de/bml/4.0"
    xmlns:gml="http://www.opengis.net/gml/3.2" >

    <bml:boreholePath >                <!– a linear feature, referenced by the event -->
        <gml:Curve>
            <gml:segments/>
        </gml:Curve>
    </bml:boreholePath >

    <bml:event>
        <bml:BoreholeEvent>
            <bml:location>             <!– 1 meter from start along the borehole path -->
                <bml:AtLocation>
                    <bml:distanceAlong uom="m">1</bml:distanceAlong>
                </bml:AtLocation>
            </bml:location>

            <bml:value>
                <bml:GroundwaterObservation>        <!– Dependency Injection -->
                    <bml:drillingDepth uom="m">815</bml:drillingDepth>
                    <bml:overflow>true</bml:overflow>
                </bml:GroundwaterObservation>
            </bml:value>
        </bml:BoreholeEvent>
    </bml:event>

</bml:Borehole>
```

GFZ Helmholtz Centre Potsdam

HELMHOLTZ ASSOCIATION

Implementation: Borehole Interoperability Experiment (BoreholeIE)

"Geological and geophysical data are crucial for many domains of our society. Any domain that need to probe the underground relies on these data: raw material exploration, hydrology, oil and gas, mining, civil engineering (constructions, transportation), and environmental sciences. In most cases, the according human activities that need to investigate the underground are relying on indirect observations methods like geophysical measurements (e.g., seismic or geoelectrical exploration). However, for direct investigations of the subsurface and for the verification of indirect measurements and models, drilling is essential. Notwithstanding, each geoscientific domain has its own set of methods and interest (resource geologists look at rock properties relevant to mineral occurrences; hydrogeologists at water resources; civil engineers at mechanical properties; and the energy sector at fossil fuel potential) they all use the same basic engineering feature, a borehole …

Involving key implementers and editors of the existing standards, this Interoperability Experiment aims at defining a domain agnostic and comprehensive (umbrella) vocabulary for a general concept of boreholes, which may eventually become its own specification and be properly re-used by various domains when needed."(OGC)

BoreholeIE Example (LR & DI): Referenced Object

```xml
<?xml version="1.0" encoding="UTF-8"?>
<BhML:BoreholeEventCollection xmlns:bml="http://www.gfz-potsdam.de/boreholeml/4.2" xmlns:gml="http://www.opengis.net/g3.2"
xmlns:BhML="https://raw.githubusercontent.com/opengeospatial/boreholeie/master/schemas">

    <BhML:collectionLRS>
        <BhML:BoreholeLRS>
            <BhML:boreholeReferencingMethod/>
            <BhML:linearElement>
                <BhML:BoreholeTrajectory>
                    <BhML:curve>
                        <gml:LineString srsDimension="3" srsName="http://www.opengis.net/gml/srs/epsg.xml#4979">
                            <gml:posList>0 0 0 0 0 10 0 0 20 0 0 42</gml:posList>
                        </gml:LineString>
                    </BhML:curve>
                </BhML:BoreholeTrajectory>
            </BhML:linearElement>
            <BhML:trajectoryReferent/>
        </BhML:BoreholeLRS>
    </BhML:collectionLRS>

    <BhML:collectionMember/>

</BhML:BoreholeEventCollection>
```

Referenced Object (Feature):
Borehole Trajectory (GML Line String)

GFZ Helmholtz Centre POTSDAM

HELMHOLTZ ASSOCIATION

BoreholeIE Example (LR & DI): Referencing Object

```xml
<?xml version="1.0" encoding="UTF-8"?>
<BhML:BoreholeEventCollection>
    <BhML:collectionLRS/>

    <BhML:collectionMember>
        <bml:GroundwaterObservationEvent>
            <BhML:locatedMember>
                <bml:GroundwaterObservation>
                    <bml:drillingDepth uom="http://qudt.org/vocab/unit/M">42</bml:drillingDepth>
                    <bml:overflow>true</bml:overflow>
                </bml:GroundwaterObservation>
            </BhML:locatedMember>
            <BhML:location>
                <BhML:AtLocation>
                    <BhML:atPosition>
                        <BhML:MeasureOrEvent>
                            <BhML:distanceAlong uom="http://qudt.org/vocab/unit/M">23</BhML:distanceAlong>
                        </BhML:MeasureOrEvent>
                    </BhML:atPosition>
                </BhML:AtLocation>
            </BhML:location>
        </bml:GroundwaterObservationEvent>
    </BhML:collectionMember>
</BhML:BoreholeEventCollection>
```

Dependency Injection:
Groundwater Observation (BoreholeML)

GFZ Helmholtz Centre POTSDAM

HELMHOLTZ ASSOCIATION

Implementation: Formal Classification and Support of Expert-Opinions

Further developments aim at correlating borehole logs, geological or geotechnical surveys, and geoscientific models. Since results of surveys are often only available as non-schematised interpretations in text form, interoperability requires formal classifications, which can be derived from machine learning methods applied to the interpretations. As part of a Conceptual Model, such classifications can be used for an automated exchange of standard-conform borehole logs or to support the generation of expert opinions on soil investigations

One encoding example (of many), representing geological information

- Ignoring existing and well-established standards like GML and ISO 19139
  - <latitude>63.401788</latitude>
  - <longitude>13.202924</longitude>
  - <coordinate_system>WGS84</coordinate_system>
  - <elevation>522.04</elevation>
  - <elevation_unit>m</elevation_unit>
- Complete lack of separation of concerns (Data, Metadata & Semantics)
  - <city>Åre</city>
  - <material>Rock</material>
- Content used for structural elements
  - <is_private>0</is_private>
- Non-schematised, non-classified free text
  - <collection_method>Coring>RockCorer</collection_method>
- Redundancies
  - <collection_date_precision>day</collection_date_precision>
- Strange nomenclature
  - <publish_date>2017-3-1</publish_date>

This example illustrates the need to use (machine learning based) formal classification of geoscientific information

GFZ Helmholtz Centre Potsdam

HELMHOLTZ ASSOCIATION

Semantics, DI, & Meta-Model: Lithology Sample

```xml
<?xml version="1.0" encoding="UTF-8"?>

<swe:DataRecord
    definition="lithology sample" id="rcc12"

    xmlns:swe="http://www.opengis.net/swe/2.0"
    xmlns:xlink="http://www.w3.org/1999/xlink">


    <!– Dependency Injections of classification from a Semantic Registry -->
    <swe:field name="stratigraphy"
        xlink:href="http://www.gfz-potsdam.de/dictionaries/stratigraphy#Tertiary"/>

    <swe:field name="grainSize"
        xlink:href="http://www.gfz-potsdam.de/dictionaries/petrography/grainSize#gU"/>

    <swe:field name="weightPercentage"
        xlink:href="http://www.gfz-potsdam.de/dictionaries/classification#veryFew"/>

</swe:DataRecord>
```

Semantics:

1. Semantic Registry (Grain Size)
2. Domain Controlled Vocabularies

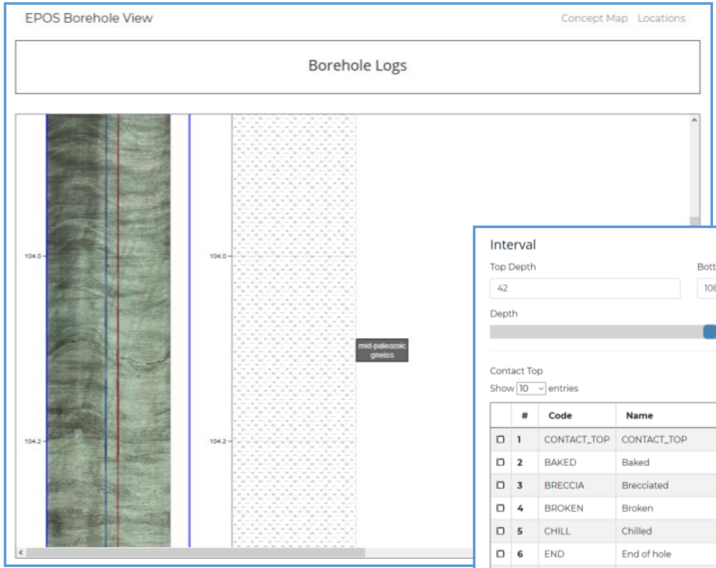
Dependency Injection:

1. Classification (e.g., Tertiary)

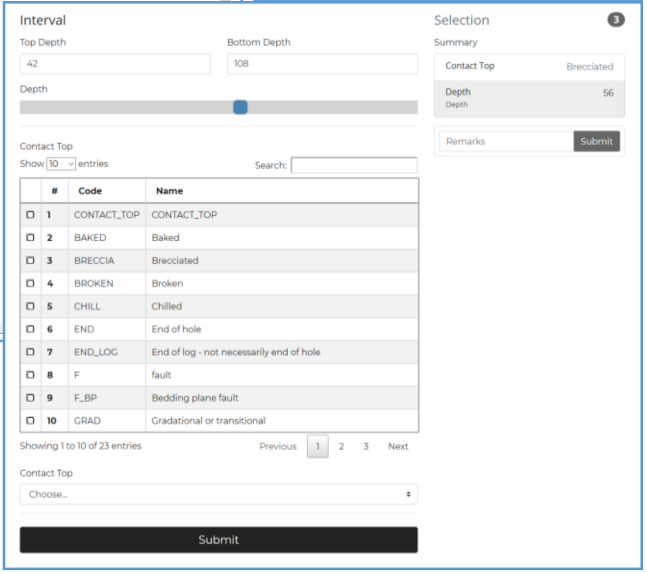
Meta-Model

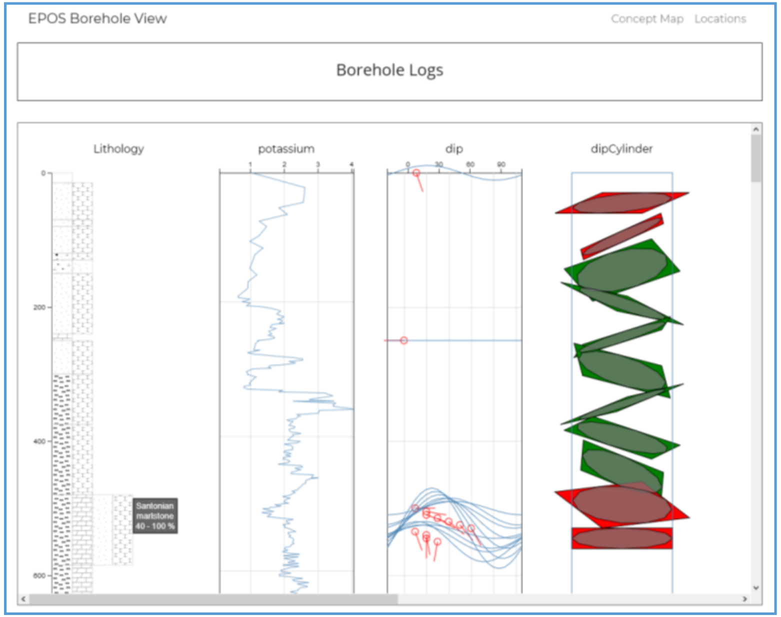
1. SWE Common
2. GML

Interactive and web-based Well-Log Analysis



Interactive assignment of classifications and characteristics





GFZ
Helmholtz Centre POTSDAM

HELMHOLTZ ASSOCIATION

Conclusion: Establishing Systems-of-Systems

Collaboration & Interoperability	Interactive Processing of Information
	Coarse-Grained Functionality Clusters
	Delegation of Responsibilities
Autonomy & Adaptivity	Independent Processing of Events and Messages
	Intelligent Monitoring Strategies
	Robustness and Responsiveness
Evolvement & Evolution	Seamless Integration of new Resources
	Composability and Re-Use of Functionality
	Scalability in terms of Systems and Domains

Adapted from Presentation EGU 2019, ESSI2.1

Thank you for your attention

Rainer Häner (rainer.haener@gfz-potsdam.de)

Helmholtz Centre Potsdam GFZ (German Research Centre for Geosciences)

eScience Centre

https://www.epos-ip.org/tcs/geological-information-and-modeling/data-services/wp15-services-and-architecture