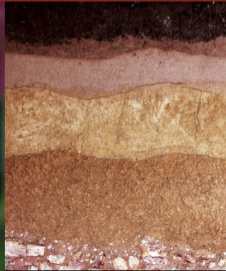


GloSIS

An operational Soil Ontology with the Semantic Web

L. M. de Sousa ¹ R. Palma ² B. Janiak ² J. Leenaars ¹ F. van Egmond ¹

May 26, 2024





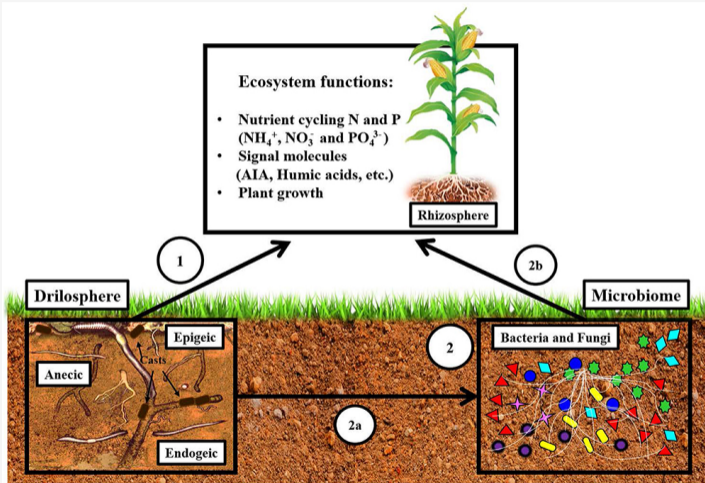
Acknowledgments





Soil: its nature and importance

What is Soil?



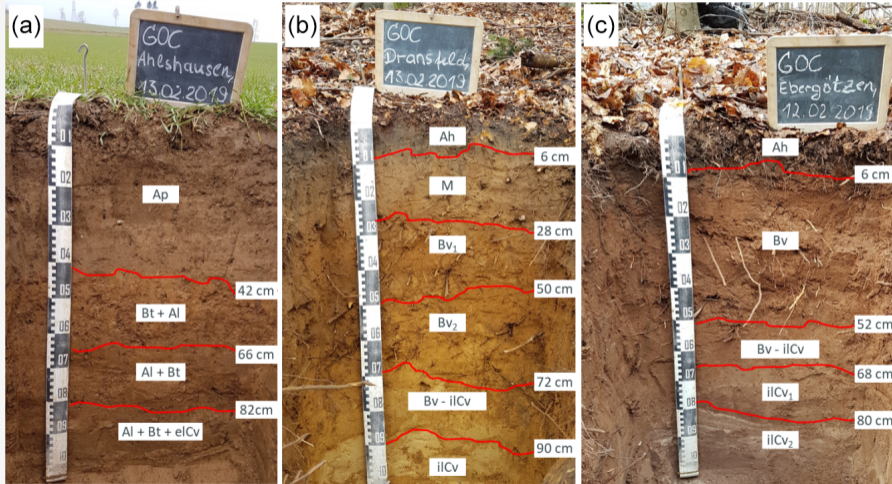
Sustainable development



Soil survey

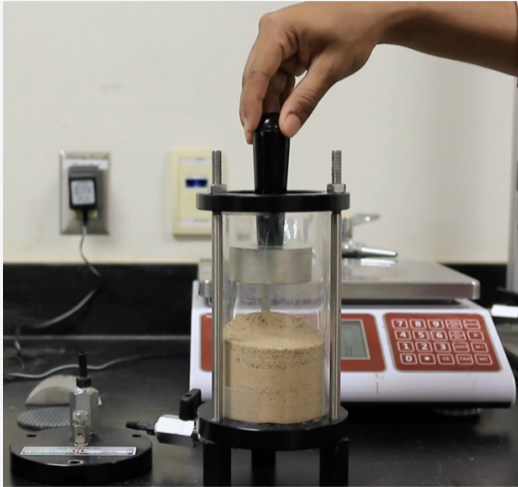


Soil description





Physio-chemical analysis





GloSIS and the GSP



The Global Soil Partnership

Promoted by the FAO

- Network of stakeholders in the Soil Domain.
- Established in 2012.
- Broad goals:
 - raise awareness to the importance of soils;
 - help attaining sustainable agriculture;
 - promote good practices in soil management.
- Majority of the world's soil information institutions involved.
 - Gathered around the International Network of Soil Information Institutions (INSII).



GLOBAL SOIL
PARTNERSHIP



The Global Soil Partnership

2017: Five Pillars of the GSP

- Pillar 1 – **Soil management** – promote the sustainable management of soil resources.
- Pillar 2 – **Awareness raising** – encourage investment, technical cooperation, policy, education and awareness.
- Pillar 3 – **Research** – focused on identified gaps, priorities and synergies with related actions.
- Pillar 4 – **Information and data** – enhance the quantity, quality and availability of soil data and Information.
- Pillar 5 – **Harmonisation** – targeting methods, measurements and sustainable management indicators.



Pillar 5 – Harmonisation

Action Plan for Pillar 5

- Difficulties identified:
 - data collected and curated by national or regional institutions;
 - focused on local context;
 - lack of heterogeneity for global use;
 - data transfer across regions/countries prone to errors.
- Among the key priorities in Pillar 5:
 - development of a **soil information exchange** mechanism.



Pillar 5 – Harmonisation

Soil Information Exchange

[...] a conceptual soil feature information model provid[ing] the framework for harmonisation such that the efficient exchange and collation of globally consistent data and information can occur.

- Data exchange as an **essential** harmonisation component.
- Facilitating data **integration, analysis** and **interpretation**.
- A **foundation** to all other GSP Pillars.



International Consultancy



Requirements

Understanding stakeholders needs

- **Requirements** inventory.
 - Meetings and interviews with GSP stakeholders.
- **Re-use** existing assets as much as possible:
 - soil domain ontologies;
 - soil data exchange mechanisms.
- Assess **implementation** possibilities.



Ontologies and data models assessed

Wealth of soil ontologies

- **ANZSoilML** [Simons et al. 2013]
- **INSPIRE** [INSPIRE Thematic Working Group Soil 2013]
- **ISO-28258** [ISO 2013]
- **OGC Soil IE** [OGC 2016]
- **WoSIS** [Batjes et al. 2020]
- **SOTER** [Oldeman & Engelen 1993]



Implementation

From model to data exchange mechanism

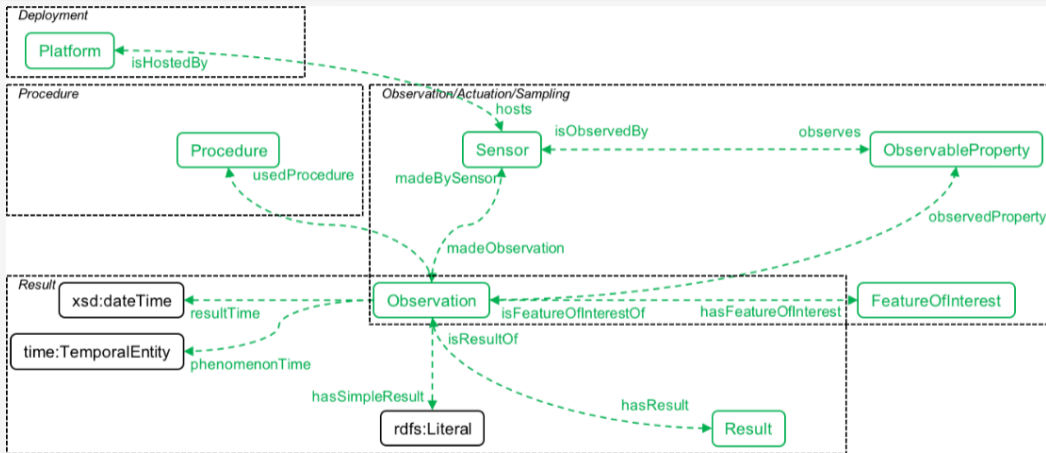
- **XML:** Observations and Measurements [Cox 2011]
- **Semantic Web:** Sensor, Observation, Sample, and Actuator ontology (SOSA) [Janowicz et al. 2019]



From domain model to the web ontology



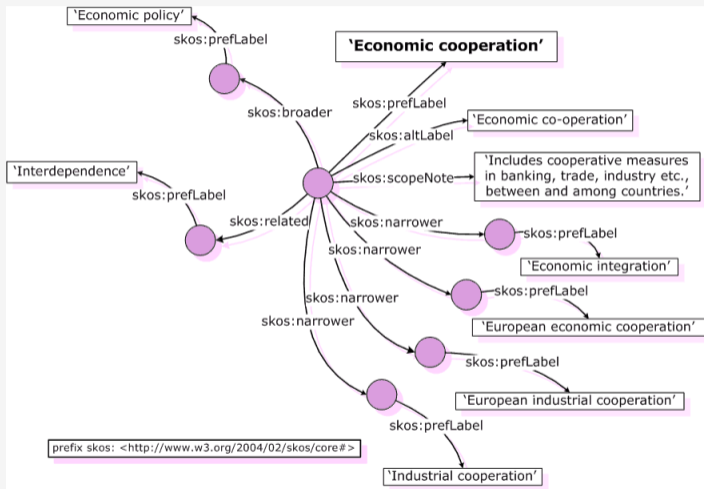
SOSA: semantic web counterpart to O&M



<http://www.w3.org/ns/sosa/>



SKOS: Simple Knowledge Organisation System





Physio-chemical properties and procedures

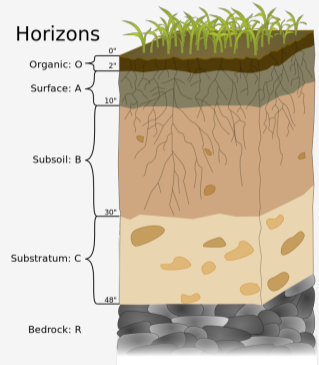
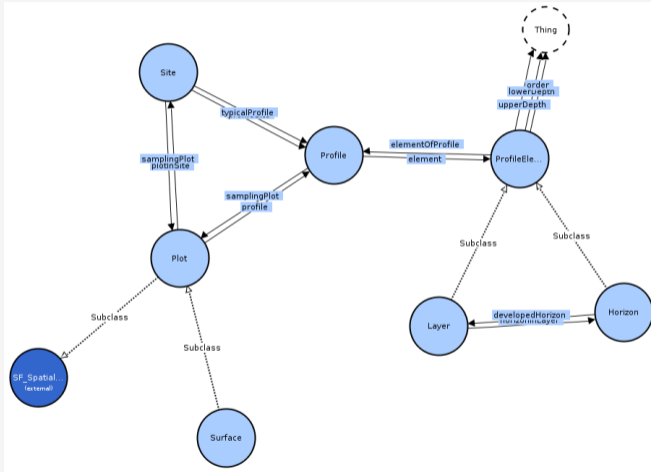
Making use of existing soil information systems

- Main sources:
 - **AfSP**: Africa Soil Profiles
 - **WoSIS**: World Soil Information Service
- Human-friendly descriptions.
- Bibliographic references.
 - Making use of URIs as much as possible.



Overview

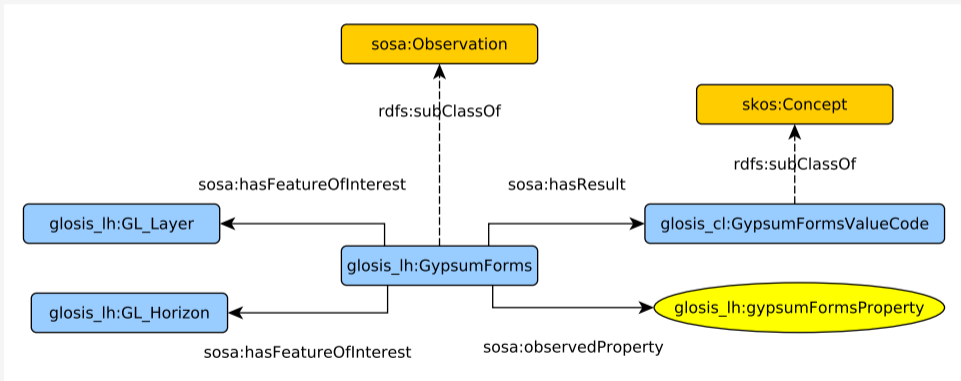
Features of Interest



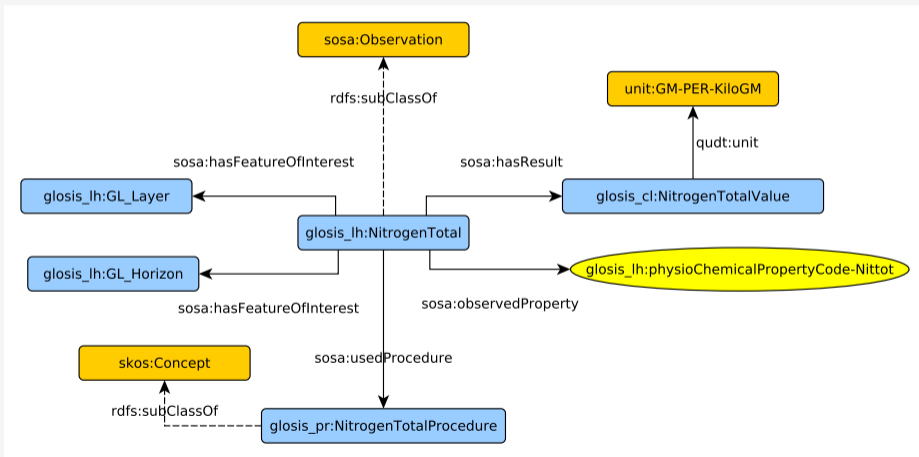
Source: [Wikimedia Commons](#)



Observations - Descriptive properties



Observations - Physio-chemical properties



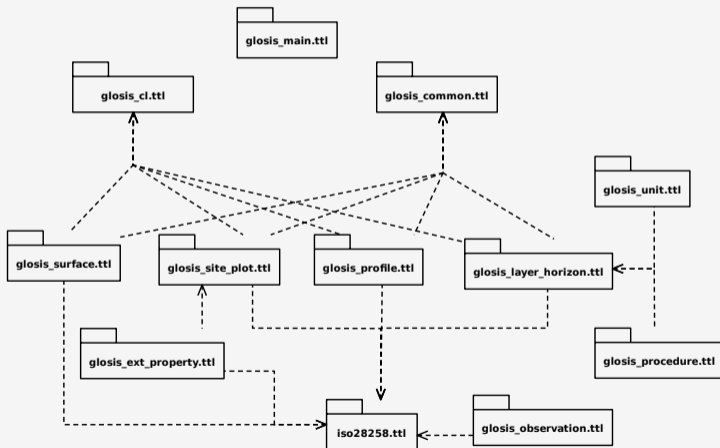


Code-lists

Main assets set up with SKOS

- Code-lists for **Descriptive** properties (FAO Guidelines of Soil Description)
 - 40 Site/Plot
 - 90 Layer/Profile
 - 5 Surface
- 80 **Physio-chemical** properties (AfSP and WoSIS)
- 230 **Procedures** (AfSP and WoSIS).

Modules





Benefits of the Semantic Web



Transparency and ease of access

Open by design

- Human-friendly code-list index(es).
- HTML content negotiation:
 - provide rich HTML to humans;
 - accessible RDF to machines.
- A galaxy away from UML diagrams;
 - often requiring proprietary tools.

[Concept](#) | [BulkDensityWholeSoilProcedure](#) | [html](#) [turtle](#) [rdf+xml](#) [ld+json](#) [n3](#) [n-triples](#) | [Alternates view](#)

BlkDensW_we-cl-fc

URI: http://w3id.org/glosis/model/procedure/bulkDensityWholeSoilProcedure-BlkDensW_we-cl-fc

[In concept scheme](#)

[Code list for bulkDensityWholeSoil analysis procedures - codelist scheme](#)

Definition

Whole earth. Clod samples (natural clods), at field capacity (0.33 bar, 33 kPa, 330 cm, pF 2.5), not corrected for coarse fragments if any

[Top concept of](#)

- [Code list for bulkDensityWholeSoil analysis procedures - codelist scheme](#)

notation

BlkDensW_we-cl-fc

scope note

[isric_report_2014_01.pdf](#)

scope note

Leenaars J.G.B., A.J.M. van Oostrum and M. Ruyter Gonzalez, 2014. Africa Soil Profiles Database, Version 1.2. A compilation of georeferenced and standardised legacy soil profile data for Sub-Saharan Africa (with dataset). ISRIC Report 2014/01. Africa Soil Information Service (AFSIS) project and ISRIC - World Soil Information, Wageningen, the Netherlands. See Annex 4.



The ontology can be queried

```
SQL> DB.DBA.TTLP_MT (file_to_string_output ('/database/gloasis_cl.ttl'), '',  
                    'http://w3id.org/gloasis/model/codelists#');
```

```
Done. -- 74 msec.
```




The ontology can be queried

```
SQL> DB.DBA.TTLP_MT (file_to_string_output ('/database/gloasis_cl.ttl'), '',
                    'http://w3id.org/gloasis/model/codelists#');
```

```
Done. -- 74 msec.
```

```
SQL> sparql
PREFIX gloasis_cl: <http://w3id.org/gloasis/model/codelists#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?code ?label
WHERE {
    ?code skos:topConceptOf gloasis_cl:landformComplexValueCode ;
        skos:prefLabel ?label .
};
```



The ontology can be queried

LONG VARCHAR

http://w3id.org/glosis/model/codelists#landformComplexValueCode-CU
http://w3id.org/glosis/model/codelists#landformComplexValueCode-DO
http://w3id.org/glosis/model/codelists#landformComplexValueCode-DU
http://w3id.org/glosis/model/codelists#landformComplexValueCode-IM
http://w3id.org/glosis/model/codelists#landformComplexValueCode-IN
http://w3id.org/glosis/model/codelists#landformComplexValueCode-KA
http://w3id.org/glosis/model/codelists#landformComplexValueCode-RI
http://w3id.org/glosis/model/codelists#landformComplexValueCode-TE
http://w3id.org/glosis/model/codelists#landformComplexValueCode-WE

LONG VARCHAR

Cuesta-shaped
Dome-shaped
Dune-shaped
With intermontane plains (occupy
Inselberg covered (occupying > 1
Strong karst
Ridged
Terraced
With wetlands (occupying > 15%)

9 Rows. -- 20 msec.



The ontology can be queried

```
SQL> sparql
PREFIX glosis_cl: <http://w3id.org/glosis/model/codelists#>
SELECT ?predicate ?object
WHERE {
  glosis_cl:landformComplexValueCode ?predicate ?object
};
```

LONG VARCHAR

LONG VARCHAR

```
http://www.w3.org/1999/02/22-rdf-syntax-ns#type
http://www.w3.org/2000/01/rdf-schema#label
http://www.w3.org/2000/01/rdf-schema#seeAlso
http://www.w3.org/2004/02/skos/core#definition
http://www.w3.org/2004/02/skos/core#note
http://www.w3.org/2004/02/skos/core#prefLabel
```

```
http://www.w3.org/2004/02/skos/core#ConceptScheme
Code list for LandformComplexValue - codelist schem
http://w3id.org/glosis/model/codelists#LandformComp
table 5
This code list provides the LandformComplexValue.
Code list for LandformComplexValue - codelist schem
```

6 Rows. -- 1 msec.



Federated queries

No more silos!

```
PREFIX sosa: <http://www.w3.org/ns/sosa/>
PREFIX qudt: <http://qudt.org/schema/qudt/>
PREFIX glosis_lh: <http://w3id.org/glosis/model/layerhorizon/>

SELECT ?obs ?value
      (REPLACE(STR(glosis_lh:NitrogenTotal),
              "^.*#([\^#]*)$", "$1") as ?property)
WHERE
{
  SERVICE <https://virtuoso.isric.org/sparql/> {
    ?obs a glosis_lh:NitrogenTotal;
        sosa:hasResult ?res ;
        sosa:hasFeatureOfInterest ?lay .
    ?res qudt:numericValue ?value .
    FILTER (?value > 2) .
  }
}
```





Extending code-lists

Example: a local procedure for texture fractions defined at 5, 50 and 2000 μm .



Extending code-lists

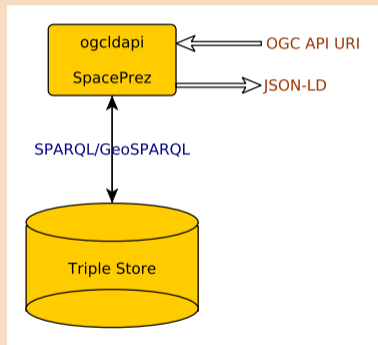
Example: a local procedure for texture fractions defined at 5, 50 and 2000 μm .

```
my_soil:textureProcedure-SaSiCl_5-50-2000u a skos:Concept, glosis_pr:TextureProcedure;  
  skos:topConceptOf glosis_pr:textureProcedure;  
  skos:prefLabel "SaSiCl_5-50-2000u"@en ;  
  skos:notation "SaSiCl_5-50-2000u" ;  
  skos:definition "Sand, silt, clay fractions as used in my country (5-50-2000um)" ;  
  skos:inScheme glosis_pr:textureProcedure .
```

Straightforward adoption of data provision standards

Adoption of OpenAPI by the OGC is a game changer

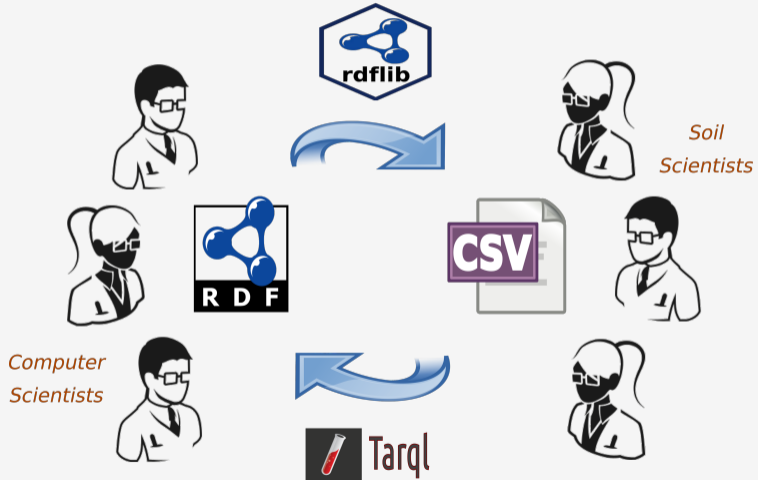
- Features referenced with URIs:
 - <https://my-soil.org/collections/07/items/101>
- JSON is one of the response formats:
 - RDF viable with JSON-LD (semantic context);
- Response document structure not prescribed:
 - going beyond flat tables and rigid hierarchies;
 - data approached as a network (linked).



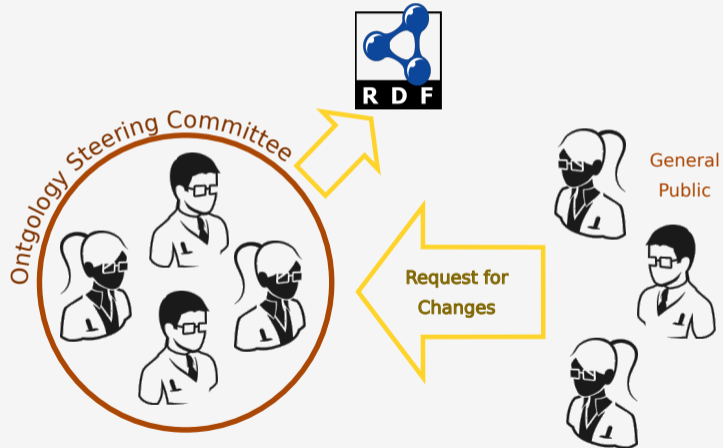
<https://github.com/surroundaust ralia/Prez>



Human engagement



Governance





Links



Join us!

Where to find GloSIS

Code forge

<https://github.com/gloasis-ld/gloasis>

Documentation

<https://gloasis-ld.github.io/gloasis>

GloSIS in peer review

<https://www.semantic-web-journal.net/content/gloasis-global-soil-information-system-web-ontology-1>

<https://doi.org/10.48550/arXiv.2403.16778>



Resources

Code-lists

Hosted by the OGC names server

<https://defs.opengis.net/vocprez/vocab/>

Hosted at ISRIC

<https://vocab.isric.org/>



Resources

Example Knowledge Graphs

GloSIS web ontology and WoSIS preview

<https://virtuoso.isric.org/sparql>

Global soil respiration database (SRDB) and LUCAS

<https://www.foodie-cloud.org/sparql>



Faceted knowledge graph browser

<https://virtuoso.isric.org/fct/>

About: [Code list for physioChemicalProperty - codelist scheme](#) [Goto](#) [Sponge](#) [NotDistinct](#)

[Permalink](#)

An Entity of Type : [skos:ConceptScheme](#), within Data Space : [virtuoso.isric.org](#) associated with source [document\(s\)](#)

Type: Command:

Attributes	Values
rdfs:type	skos:ConceptScheme
rdfs:label	Code list for physioChemicalProperty - codelist scheme (en)
rdfs:seeAlso	Code list for PhysioChemicalProperty - codelist class
skos:note	This code list provides analysis procedures for physioChemical. (en)
skos:pref_label	Code list for physioChemicalProperty - codelist scheme (en)
skos:definition	ISRIC Report 2019/01: Tier 1 and Tier 2 data in the context of the federated Global Soil Information System, Appendix 3
is rdfs:seeAlso of	Code list for PhysioChemicalProperty - codelist class
is skos:topConceptOf of	Acidity - exchangeable Aluminium (Al+++) - exchangeable Available water capacity - volumetric (EC to WP) Base saturation - calculated Boron (B) - total Calcium (Ca++) - exchangeable Calcium (Ca++) - extractable Calcium (Ca++) - total Carbon (C) - organic Carbon (C) - total smatte
is skos:inScheme of	Acidity - exchangeable Aluminium (Al+++) - exchangeable Available water capacity - volumetric (EC to WP) Base saturation - calculated Boron (B) - total Calcium (Ca++) - exchangeable Calcium (Ca++) - extractable Calcium (Ca++) - total Carbon (C) - organic Carbon (C) - total smatte

About: [Code list for biologicalFeaturesValue - codelist scheme](#) [Goto](#) [Sponge](#) [NotDistinct](#)

[Permalink](#)

An Entity of Type : [skos:ConceptScheme](#), within Data Space : [virtuoso.isric.org](#) associated with source [document\(s\)](#)

Type: Command:

Attributes	Values
rdfs:type	skos:ConceptScheme
rdfs:label	Code list for biologicalFeaturesValue - codelist scheme (en)
rdfs:seeAlso	Code list for BiologicalFeaturesValue - codelist class
skos:note	This code list provides analysis procedures for biologicalFeatures. (en)
skos:pref_label	Code list for biologicalFeaturesValue - codelist scheme (en)
skos:definition	Guidelines for Soil Description issued by the FAO: table B2
is rdfs:seeAlso of	Code list for BiologicalFeaturesValue - codelist class
is skos:topConceptOf of	Artefacts Burrows (unspecified) Infilled large burrows Open large burrows Charcoal Earthworm channels Other insect activity Pedotubules Termite or ant channels and nests
is skos:inScheme of	Artefacts Burrows (unspecified) Infilled large burrows Open large burrows Charcoal Earthworm channels Other insect activity Pedotubules Termite or ant channels and nests

Diving further

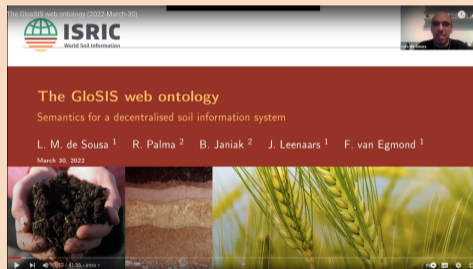


Video material on GloSIS and related

On-line address to ESIP Soil on GloSIS

EJP Soil training: Concepts on Ontology

EJP Soil training: Introduction to GloSIS



Thank you!

luis.desousa@isric.org

[@luis_de_sousa@mastodon.social](https://mastodon.social/@luis_de_sousa)

www.isric.org

