The template for a Semantic SensorThings API with the GloSIS use case

Luís Moreira de Sousa[1], Rául Palma[2], Bogusz Janiak[2], Paul van Genuchten[1]

[1] ISRIC - World Soil Information, Wageningen, The Netherlands - (luis.desousa, paul.vangenuchten)@isric.org
[2] Poznań Supercomputing and Networking Center - PSNC, Poznań, Poland - (rpalma, bjaniak)@man.poznan.pl

Keywords: SensorThings API, Semantic Web, Linked Data, Observations & Measurements, Soil

Abstract

Since 2016 the Open Geospatial Consortium (OGC) has embarked on a new specification paradigm for web access to geo-spatial data. Moving on from the SOAP/XML pattern underlying the many web services it specified earlier, the OGC has since issued novel standards based on ReST API specifications such as OpenAPI and OData. In tandem, data encoding formats considered in the OGC's standards have been greatly expanded. This new direction has greatly aligned the OGC's work with the Semantic Web, the collection of specifications for data on the web based on the Resource Description Framework (RDF) issued by the World Wide Web Consortium (W3C).
In this context, this communication explores the SensorThings API specification as means to serve environmental data in standardised fashion. With concrete examples, it demonstrates how the adoption of the Semantic Web in Spatial Data Infrastructures opens up a fresh approach to the development and deployment of data services, especially those in which domain data is deeply structured. Two use cases are reviewed herewith: the GloSIS web ontology for soil data exchange and the Iliad Ocean Digital Twin.

1. Introduction

Spatial Data Infrastructures (SDI) developed for the exchange of environmental data have heretofore been greatly shaped by the standards issued by the Open Geospatial Consortium (OGC). Based on the Simple Object Access Protocol (SOAP), services like WMS, WFS, WCS or CSW became digital staples for researchers and administrative bodies alike.

In 2017 the Spatial Data on the Web Working Group (SDWWG) of the World Wide Web Consortium (W3C) published a report titled "Spatial Data on the Web Best Practices", questioning the overall philosophy of the OGC's standards based on the ageing SOAP technology (Tandy et al., 2017). The main issues identified by the SDWWG can be summarised as:

- Spatial resources are not identified with URIs.
- Modern API frameworks, e.g. OpenAPI (Miller et al., 2021), are not being used.
- Spatial data are still shared in silos, missing links to other resources.
- SDIs based on OGC web services are difficult to use:
  - content indexing by search engines is not facilitated.
  - catalogue services only provide access to metadata, not the data themselves.
  - not possible to access data trough links (e.g. URLs).
  - data is difficult to understand for non-domain-experts.

To address these issues the SDWWG proposed a five point strategy inspired on the Five Star Scheme (Berners-Lee, 2006):

- **Linkable**: use stable and discoverable global identifiers.
- **Parseable**: use standardised data meta-models such as CSV, XML, RDF, or JSON.
- **Understandable**: use well-known, well-documented, vocabularies/schemas.
- **Linked**: link to other resources whenever possible.
- **Usable**: label data resources with a licence.

The work of the SDWWG triggered vast transformations to the approach taken by the OGC, with a shift to specifications based on OpenAPI and OData. But while convenience of use is the focus, semantics has been left unattended, a Linked Data philosophy has not yet been actively pursued.

However, the modern API approach opens the door to a complete coupling of OGC services with the Semantic Web, especially considering the possibility of adopting JSON-LD as syntax to OGC API response documents. The introduction of a semantic layer to digital environmental data shared through state-of-the-art OGC APIs is becoming a reality, with great benefits to researchers using or sharing data.

This communication lays down a simple SDI set up to serve semantic environmental data through a SensorThings API created with the `glrc` software. It starts by reviewing the relevant standards and software in Section 2. A set of use cases follows in Section 3, exemplifying the approach, starting with two soil data services compliant with the GloSIS web ontology (Section 3.1), to which adds the Iliad Ocean Twin (Section 3.2). The article closes with a few directions for future work (Section 4).

2. Materials and Methods

2.1 Sensor Things API

SensorThings API (STA) is an OGC standard specifying a unified framework to interconnect Internet of Things resources

SensorThings API	SOSA
FeatureOfInterest	FeatureOfInterest
ObservedProperty	ObservedProperty
Sensor	Sensor, Procedure
Datastream	Observation
Observation	Result

Table 1. Correspondence between classes in the SesnorThings API domain model and those specified in SOSA.

over the Web (Liang et al., 2016). SensorThings API aims to address both the semantic, as well as syntactic, interoperability. It follows ReST principles (Fielding and Taylor, 2002), promotes data encoding with JSON, the OASIS OData protocol (Chappell, 2011) and URL conventions.

The SensorThings API is underpinned on a domain model aligned with the ISO/OGC standard Observations & Measurements (O&M) (Cox, 2011), targeted at the interchange of observation data of natural phenomena. O&M puts forth the concept of `Observation` has an action performed on a `Feature of Interest` with the goal of measuring a certain `Property` through a specific `Procedure`. SensorThings API mirrors these concepts with `Observation`, `Thing`, `ObservableProperty` and `Sensor`. This character makes of SensorThings API a vehicle for the interoperability of heterogeneous sources of environmental data. In 2019 the W3C published a Semantic Web counterpart to O&M, the Sensor, Observation, Sample, and Actuator ontology (SOSA) (Janowicz et al., 2019). Table 1 provides a high-level correspondence between concepts in SOSA and the SensorThings API domain model.

2.2 glrc

`grlc` (pronounced "garlic") is a lightweight server that translates SPARQL queries into Linked Data web APIs (Meroño-Peñuela and Hoekstra, 2016) compliant with the OpenAPI specification (Miller et al., 2021). Its purpose is to enable universal access to Linked Data sources through modern web-based mechanisms, dispensing the use of the SPARQL query language. While losing the flexibility and federative capacities of SPARQL, web APIs present developers with an approachable interface that can be used for the automatic generation of source code.

A `glrc` API is constructed from a SPARQL query to which a meta-data section is prepended. This section is declared with a simplified YAML syntax, within a SPARQL comment block, so the query remains valid SPARQL. The meta-data provide basic information for the API set up and most importantly, the SPARQL end-point on which to apply the query. Listing 1 shows an example. A special SPARQL variable formulation is used to map into API parameters. By adding an underscore (_) between the question mark and the variable name, `glrc` is instructed to create a new API parameter. A prefix separated again with an underscore informs `glrc` of the parameter type. The `?band_label` variable in Listing 1 can be expanded to `?_band_label_iri` to create a new API parameter of the type IRI.

Listing 1. A SPARQL query mapping variables to a `glrc` API

```
#+ endpoint: http://dbpedia.org/sparql

PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX dbr: <http://dbpedia.org/resource/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-
    schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-
    syntax-ns#>

SELECT ?band_label {
    ?band rdf:type dbo:Band ;
          dbo:genre dbr:Hard_Rock ;
          rdfs:label ?band_label .
} ORDER BY ?band_label
```

3. Results

3.1 GloSIS

The Global Soil Partnership (GSP) is a network of stakeholders in the soil domain established in 2012 by members of the Food and Agriculture Organisation of the United Nations (FAO). Its broad goals are to raise awareness to the importance of soils and to promote good practices in land and soil management towards a sustainable agriculture.

The GSP acknowledged difficulties in exchanging harmonised soil data as an important obstacle to its goals (Partnership, 2017). Soil data is mostly collected and curated by national or regional institutions, focused on their local context, largely abstract from global concerns. To tackle this problem the GSP set as one of its priorities the development of a soil exchange infrastructure, both leading to a global and federated soil information system and providing the foundation for data harmonisation.

In 2019 the Global Soil Partnership launched an international consultancy to assess the state-of-the-art in soil information exchanges and propose a path towards its operationalisation as the back-bone of the Global Soil Information System (GloSIS). The ISO 28258 domain model for soil quality (ISO, 2013) was selected and augmented with container classes encapsulating the Guidelines for Soil Description issued by the FAO (Jahn et al., 2006), plus numerous code-lists necessary for data exchange (Řezník and Schleidt, 2020). This domain model was later transformed to a Web Ontology, relying on SOSA and other Semantic Web standards such as GeoSPARQL (Consortium, 2011), QUTD (QUDT.org, 2011) and SKOS (Miles and Bechhofer, 2009). The GloSIS web ontology has been successfully demonstrated as a vehicle to exchange soil information as Linked Data (Palma et al., 2024).

A prototype API for the GloSIS ontology demonstrates how the same set of SPARQL queries can be used to query through a ReST API any end-point available over the internet, sharing linked soil data in accordance with the GloSIS ontology. It is a clear step towards the federated and harmonised system envisioned by the GSP.

3.1.1 The WoSIS knowledge graph prototype The World Soil Information Service (WoSIS) is the result of a decadal effort towards an harmonised soil observation dataset at the global scale (Batjes et al., 2020). WoSIS comprises a relational database containing more than 500 000 geo-referenced soil profiles, reporting to 180 different countries. In total, over 18 million observations are recorded on these soil profiles, informing on physio-chemical characteristics such as soil texture, nutrients and metals content, water potential, pedo-genetic description and more. Source data are subject to rigorous quality control and harmonisation, thus producing a globally consistent dataset. For those data whose original licence allows, geo-spatial services such as WFS facilitate access to the general public.

A pilot WoSIS knowledge graph has been maintained at ISRIC employing an ontological alignment with GloSIS. Automated mappings were created, translating from relations in the WoSIS database to GloSIS and SOSA classes. Relevant features of interest are encoded in RDF as individuals of the types `GL_Site`, `GL_Profile` and `GL_Layer`, with measurements making use of the SOSA classes `ObservableProperty`, `Procedure`, `Observation` and `Result`. Primary keys in the relational database are used in the composition of URIs for each individual. Geo-spatial features of interest, such as those of type `GL_Site`, are also expressed as instances of GeoSPARQL classes. The PostGIS function `ST_AsText` is used to obtain the WKT literal matching the GeoSPARQL `hasGeometry` object property of each geo-spatial feature. Similar transformations produce RDF triples for soil layers, soil properties, observations, procedures and results. The resulting knowledge graph, encoding observations from a single WoSIS source dataset, is deployed to the Virtuoso triple store, accessible through a SPARQL endpoint [1] and the Virtuoso Faceted Browser [2]. This knowledge graph is annotated with meta-data predicates from the Dublin Core, VCard and DCat web ontologies.

Listing 2 presents a fragment of the WoSIS knowledge graph, starting with a geo-localised site (`wosis_sit:65321`) and an associated soil profile (`wosis_prf:923095`) indicating a particular soil investigation. In this fragment a single layer is declared for the profile, extending from the surface down to a depth of 34 cm (`wosis_lyr:64318`). An observation of total Nitrogen content is recorded for this layer (`wosis_obs:4357731`), with a numerical result of 3.1[3] (`wosis_res:4357731`).

Listing 2. Fragment of the WoSIS knowledge graph.

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>
    .
@prefix dcterms:         <http://purl.org/dc/
    terms/> .
@prefix qudt: <http://qudt.org/schema/qudt/> .
@prefix sosa:  <http://www.w3.org/ns/sosa/> .
@prefix geo: <http://www.opengis.net/ont/
    geosparql#> .
@prefix iso28258: <http://w3id.org/glosis/model/
    iso28258/2013/> .
@prefix glosis_sp: <http://w3id.org/glosis/model
    /siteplot/> .
@prefix glosis_pr: <http://w3id.org/glosis/model
    /profile/> .
@prefix glosis_lh: <http://w3id.org/glosis/model
    /layerhorizon/> .
@prefix glosis_proc: <http://w3id.org/glosis/
    model/procedure/> .
@prefix wosis_sit: <http://wosis.isric.org/site
    /> .
@prefix wosis_prf: <http://wosis.isric.org/
    profile/> .
@prefix wosis_lyr: <http://wosis.isric.org/layer
    /> .
@prefix wosis_obs: <http://wosis.isric.org/
    observation/> .
@prefix wosis_res: <http://wosis.isric.org/
    result/> .

wosis_sit:65321 a glosis_sp:GL_Site, geo:Point ;
    geo:asWKT "POINT(-80.25 22.82)"^^geo:
        wktLiteral .

wosis_prf:923095 a glosis_pr:GL_Profile, sosa:
    FeatureOfInterest ;
    dcterms:isPartOf wosis_ds:247 ;
    iso28258:Profile.profileSite wosis_sit:65321
        .

wosis_lyr:64318 a glosis_lh:GL_Layer, sosa:
    FeatureOfInterest ;
    iso28258:ProfileElement.elementOfProfile
        wosis_prf:923095 ;
    iso28258:ProfileElement.upperDepth "0"^^xsd:
        float ;
    iso28258:ProfileElement.lowerDepth "34"^^xsd
        :float .

wosis_obs:4357731 a glosis_lh:NitrogenTotal,
    sosa:Observation ;
    sosa:hasFeatureOfInterest wosis_lyr:64318 ;
    sosa:usedProcedure glosis_proc:
        nitrogenTotalProcedure-TotalN_kjeldahl ;
    sosa:hasResult wosis_res:4357731 .

wosis_res:4357731 a glosis_lh:NitrogenTotalValue
    , sosa:Result ;
    qudt:numericValue "3.10"^^xsd:float .
```

3.1.2 WoSIS SensorThings API

An API compliant with the SensorThings specification has been developed for the WoSIS knowledge graph using `glrc`[4]. This demonstrator service is deployed directly from a GitHub repository[5], making full use of the facilities provided by `glrc`, thus dispensing a dedicated infrastructure. As the WoSIS knowledge graph grows and/or as the API densifies, a bespoke deployment may become necessary, however this demonstrator showcases the ease and speed of API development with `glrc`.

Listing 3 presents one of the maps between a SensorThings API request and a SPARQL query. The argument to the request is declared with the `?_id` keyword and transformed to the relevant URI through a `BIND` statement. In this case the query returns all triples associated with the concerned `Observation` instance.

Listing 3. `glrc` code mapping the STA individual Observation request to a SPARQL query on the WoSIS knowledge graph.

```
#+ summary: Returns all predicates and objects
    related to an Observation
#+          instance.
#+ tags:
#+   - sensor-things

PREFIX sosa: <http://www.w3.org/ns/sosa/>
PREFIX wosis_obs: <http://wosis.isric.org/
    observation#>
```

[1] https://virtuoso.isric.org/sparql/
[2] https://virtuoso.isric.org/fct/
[3] The units of this result are declared in the observation class: http://w3id.org/glosis/model/layerhorizon/NitrogenTotal
[4] https://grlc.io/api/ldesousa/wosis-sta
[5] https://github.com/ldesousa/wosis-sta

```
SELECT ?predicate ?object
WHERE {
    ?obs_uri a sosa:Observation ;
             ?predicate ?object .
    BIND (URI(CONCAT("http://wosis.isric.org/
        observation/", ?_id)) AS ?obs_uri)
}
```

3.1.3 LUCAS 2015 Topsoil knowledge graph and APIs

The LUCAS Programme is a series of statistical surveys conducted by the Eurostat within an area frame in order to monitor changes over time in land use and land cover across the EU (Jones et al., 2020). These surveys have been conducted every three years since 2006, consisting of the visual assessment of environmental and structural elements of the landscape in predetermined geo-referenced locations. The latter are set at the intersections of a 2 by 2 km grid covering the territory of the EU, totalling circa 1 million individual locations. In each survey soil material samples are collected for posterior physio-chemical analysis.

In 2015, the LUCAS survey was carried out in all EU-28 Member States, with more than 27 000 locations selected for sampling. In practice just under 24 000 locations were visited, still an enormous figure in soil surveying, posing significant logistic and methodological challenges. Soil material was collect at a depth of $20cm$ following a common procedure. Discounting the samples whose identifier or location were lost, the 2015 LUCAS survey managed to gather consistent soil and agro-environmental data on almost 22 000 locations.

The LUCAS dataset includes the identification code `Point_ID` of the samples and data of physical and chemical properties for each sample. These properties include: Coarse fragments, clay, silt, sand, pH in $CaCl\_2$ and in $H\_2\emptyset$, Electrical Conductivity, Organic carbon, Carbonates, Phosphorus, total nitrogen, and extractable potassium. Additionally, each sample includes the elevation at which the soil sample was taken, land cover class, land use class, and NUTS codes of the sampling location. These data were transformed into a knowledge graph using the GloSIS web ontology, employing the classes `GL_Site` (geo-spatial location), `GL_Profile`, `GL_Layer` plus multiple observation classes such as `GL_PH`. Listing 4 presents selected segments of this knowledge graph.

The LUCAS 2015 knowledge graph was the first GloSIS-compliant dataset to be published on-line, becoming a test-bed for the development and testing of APIs built on `glrc`[6]. A STA end-point was also made available, possibly the first attempt at a systematic coupling between a SOSA knowledge graph and the STA specification[7]. However, it was necessary to add additional features to `glrc` in order to fully comply with the STA specification. Among these are the `limit` and `offset` parameters used for pagination. This work resulted in a `glrc` fork specialised for STA, currently maintained by the Poznań Computing Centre[8].

[6] A first iteration is accessible at: http://w3id.org/glosis/open/LUCAS/topsoildata/.
[7] API end-point: https://sta-sta.apps.paas-dev.psnc.pl/glosis-test/api/v1.0/. Swagger documentation: https://sta-sta.apps.paas-dev.psnc.pl/glosis-test/api/v1.0/docs
[8] https://gitlab.pcss.pl/daisd-public/dpi-pipelines/grlc

Listing 4. LUCAS site #26761786 and associated soil profile, layer and observations

```
<#site_26761786> a g_sp:GL_Site ;
    rdfs:label "LUCAS #26761786" ;
    gsp:hasGeometry <#site_geo_26761786> ;
    gn:parentADM1 nuts:PT1 ;
    gn:parentADM3 nuts:PT150 ;
    gn:parentCountry nuts:PT ;
    gn:parentADM2 nuts:PT15 ;
    iso28258:Site.typicalProfile <#
        profile_26761786> .
<#site_geo_26761786> a gsp:Geometry ;
    gsp:asWKT   "POINT(-8.621613437
        37.336764358)"

<#profile_26761786> a g_pr:GL_Profile ;
    rdfs:label "Profile for #26761786" ;
    iso28258:Profile.element <#layer_26761786> .
<#layer_26761786> a g_lh:GL_Layer ;
    rdfs:label "Layer for #26761786" .

<#lu_26761786> a g_sp:LandUseClass ;
    rdfs:label "Land use for #26761786" ;
    sosa:hasFeatureOfInterest <#site_26761786> ;
    sosa:hasResult <#luvalue_U111> ;
    sosa:observedProperty g_sp:
        landUseClassProperty .

<#phCaCl2_26761786> a g_lh:PH ;
    rdfs:label "pH in CaCl2 for #26761786" ;
    sosa:hasFeatureOfInterest <#layer_26761786>
        ;
    sosa:hasResult <#phCaCl2_value_26761786> ;
    sosa:observedProperty g_cl:
        physioChemicalPropertyCode-pH ;
    sosa:usedProcedure g_pd:pHProcedure-pHCaCl2
        .

<#phCaCl2_value_26761786> a g_lh:PHValue ;
    rdfs:label "pH in CaCl2 value for #26761786"
        ;
    qudt:numericValue "4.30"^^xsd:float ;
    qudt:unit unit:PH .
```

3.2 The Iliad Ocean Twin

Iliad is a research project funded by the EU's Horizon 2020 programme, aiming to develop digital twins of the ocean[9]. The project builds on two decades of investments on the blue economy, aiming to establish a digital twin that is interoperable, data-intensive and cost-effective. It capitalises on the fast expanding wealth of data produced by multiple advanced infrastructures, such as cloud computing, high-performance computing (HPC), Internet of Things, Big Data, social networking, etc. Iliad aims at an ambitious, inclusive, augmented engaging paradigm, addressing all Earth Data challenges.

The Iliad project has comprised the development of a web ontology for ocean data named Ocean Information Model[10]. It is meant to become a common vocabulary for the project, providing the basis for semantic interoperability across smart ocean

[9] https://ocean-twin.eu/ grant agreement No 101037643.
[10] https://github.com/ILIAD-ocean-twin/OIM

solutions. Like GloSIS, OIM makes use of established Semantic Web standards such as SKOS, QUDT or DCAT. It too applies SOSA to define features of interest (such as the ocean's depth layers) and associated properties. Albeit still in development, this web ontology has already conveyed the genesis of an extensive and elaborate data service compliant with the SensorThings API[11].

The Iliad API is a considerably more complex construct, including some relevant customisations of `glrc`. This API adds a pagination feature, preventing a high number of requested individuals from overwhelming the service. For requests that may result in a denser or deeper set of triples a transformation is applied to create a knowledge graph, applied in the `transform` section in the query mapping heather. The mapping returning the full list of features of interest is a good example of these advanced techniques[12].

4. Future Work

4.1 OMS and SOSA v2

The SensorThings API domain model was further developed into a new version of the original Observations & Measures specification. Among the evolutions is a renewed focus on sampling, resulting in a new model name: Observations, Measurements and Samples (OMS) (Rinne et al., 2023). This new domain model was recently approved as a standard by both the OGC and ISO. In the wake of this development the SDWWG of the W3C started working on an update to SOSA. This update will feature a specific component on Sampling, alongside the Actuation and Observation components. In what is expected to become version 2 of the web ontology, SOSA will also include a new module explicitly aligning with concepts in OMS.

Whereas SOSA version 2 might imply updates to compliant ontologies, knowledge graphs and APIs, its explicit alignment with OMS paves the way for an even more streamlined approach to semantic data reporting observations on natural and environmental phenomena. Among the results from this alignment is the automated generation of APIs from (or for) SOSA-compliant knowledge graphs.

4.2 Automated API generation

A knowledge graph made of SOSA classes instances can provide for the automatic generation of APIs compliant with the SensorThings specification. Consider for instance the SPARQL query in Listing 3. It is a simple formulation that might require further differentiation in certain contexts, however it applies successfully to any knownledge graph compliant with SOSA and including instances of the `Observation` class. The same reasoning applies to classes such as `ObservableProperty`, `Procedure`, `Sampling`, `Result`, etc. The concept of a mechanism generating such queries with minimal to no human intervention becomes evident.

Such an automated mechanism is currently being developed within the EU funded DEMETER project. For the time being known as the STA Generation Service[13], it applies a set of assumptions on the internal structure of a knowledge graph so that, through inspection, it can generate an API based on the STA fork of `glrc`. Once the knowledge graph complies with the given assumptions, the user can merely provide the location of a SPARQL end-point to obtain a fully functional SensorThings API[14]. More recently the STA Generation Service was also applied to the LUCAS knowledge graph referenced in Section 3.1.3, producing yet another standardised API for those data [15].

Whereas still in active development, the work conducted within the DEMETER project, and continued in various ongoing projects like ILIAD and AD4GD, clearly points the direction for standards-compliant APIs of environmental data. A semantic approach to this kind of data, especially based on the SOSA web ontology, is transforming work on this field from a services-centred task to an ontology-oriented undertaking.

[11] https://github.com/ILIAD-ocean-twin/JF-API
[12] https://github.com/ILIAD-ocean-twin/JF-API/blob/main/FeaturesOfInterest.rq
[13] https://docs.psnc.pl/display/DEM/STA+generation+service
[14] https://sensor-things-api-sensor-things-api.apps.dcw1.paas.psnc.pl/
[15] https://sta-sta.apps.paas-dev.psnc.pl/glosis-test/api/v1.0/docs

References

Batjes, N. H., Ribeiro, E., Van Oostrum, A., 2020. Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019). *Earth System Science Data*, 12(1), 299–320.

Berners-Lee, T., 2006. Linked data.

Chappell, D., 2011. Introducing odata. *Data Access for the Web, The Cloud, Mobile Devices, and More*, 1–24.

Consortium, O. G., 2011. OGC GeoSPARQL - A Geographic Query Language for RDF Data. Technical report, OGC - Open Geospatial Consortium.

Cox, S., 2011. Observations and measurements-xml implementation. version 2.0. Technical report, Open Geospatial Consortium.

Fielding, R. T., Taylor, R. N., 2002. Principled design of the modern web architecture. *ACM Transactions on Internet Technology (TOIT)*, 2(2), 115–150.

ISO, 2013. Soil quality – Digital exchange of soil-related data. Standard, International Organization for Standardization, Geneva, CH.

Jahn, R., Blume, H., Asio, V., Spaargaren, O., Schad, P., 2006. *Guidelines for soil description*. FAO.

Janowicz, K., Haller, A., Cox, S. J., Le Phuoc, D., Lefrançois, M., 2019. SOSA: A lightweight ontology for sensors, observations, samples, and actuators. *Journal of Web Semantics*, 56, 1–10.

Jones, A., Fernandez-Ugalde, O., Scarpa, S., 2020. Lucas 2015 topsoil survey. presentation of dataset and results, eur 30332 en, publications office of the european union.

Liang, S., Huang, C.-Y., Khalafbeigi, T., 2016. OGC SensorThings API Part 1: Sensing, Version 1.0.

Meroño-Peñuela, A., Hoekstra, R., 2016. grlc Makes GitHub Taste Like Linked Data APIs. *The Semantic Web: ESWC 2016 Satellite Events, Heraklion, Crete, Greece, May 29 – June 2, 2016*, Springer, 342–353.

Miles, A., Bechhofer, S., 2009. Skos simple knowledge organization system reference. Technical report, W3C - World Wide Web Consortium.

Miller, D., Whitlock, J., Gardiner, M., Ralphson, M., Ratovsky, R., Sarid, U., 2021. The openapi specification (version 3.1.0).

Palma, R., Janiak, B., de Sousa, L. M., Schleidt, K., Řezník, T., van Egmond, F., Leenaars, J., Moshou, D., Mouazen, A., Wilson, P., Medyckyj-Scott, D., Ritchie, A., Yigini, Y., Vargas, R., 2024. Glosis: The global soil information system web ontology.

Partnership, G. S., 2017. Plan of action for pillar five of the global soil partnership. Technical report, GSP - Global Soil Partnership.

QUDT.org, 2011. Quantities, units, dimensions and types (qudt). Technical report, QUDT.org.

Rinne, I., Schleidt, K., van den Brink, L., Grellet, S., Portele, C., van der Schaaf, H., Beaufils, M., Bressan, H., Cox, S., Feliachi, A., Huisman, R., Ritchie, A., Sőrés, L., Klausen, J., 2023. OGC Abstract Specification Topic 20: Observations, measurements and samples. http://www.opengis.net/doc/as/om/3.0.

Tandy, J., van den Brink, L., Barnaghi, P., 2017. Spatial data on the web best practices.

Řezník, T., Schleidt, K., 2020. Data Model Development for the Global Soil Information System (GloSIS). Technical report, GSP - Global Soil Partnership.