Open Geospatial Consortium

Publication Date: 2016-07-26

Approval Date: 2016-06-23

Posted Date: 2016-05-05

Reference number of this document: OGC 16-088r1

Reference URL for this document: http://www.opengis.net/doc/PER/OGC-Soil-IE

Category: Public Engineering Report

Editor: Alistair Ritchie

OGC Soil Data Interoperability Experiment

Warning

This document is not an OGC Standard. This document presents a discussion of technology issues considered in an initiative of the OGC Interoperability Program. This document does not represent an official position of the OGC. It is subject to change without notice and may not be referred to as an OGC Standard. However, the discussions in this document could very well lead to the definition of an OGC Standard.

Document type:	OGC® Engineering Report
Document subtype:	NA
Document stage:	Approved for public release
Document language:	English

License Agreement

Contents

Page

Figures

Page

Tables

Page

Abstract

This engineering report describes the results of the Soil Data Interoperability Experiment (the IE) conducted under the auspices of the OGC Agriculture Domain Working Group in 2015. Soil data exchange and analysis is compromised by the lack of a widely agreed international standard for the exchange of data describing soils and the sampling and analytical activities relating to them. Previous modeling activities in Europe and Australasia have not yielded models that satisfy many of the data needs of global soil scientists, data custodians and users. This IE evaluated existing models and proposed a common core model, including a GML/XML schema, which was tested through the deployment of OGC web services and demonstration clients. IE time constraints and limited participant resources precluded extensive modeling activities. However, the resulting model should form the core of a more comprehensive model to be developed by a future OGC Soil Data Standards Working Group.

Business Value

Modern digital mapping and modeling techniques are of huge value in better understanding our imperfectly sampled soil resource. However they depend on dispersed, inconsistent and difficult to access digital data. Modern, harmonized and interoperable information systems are required to quickly and efficiently integrate these data into a consistent set of soil data for analysis and decision making.

Keywords

ogcdocs, soil, wfs, wcs, wps, sos, skos, rdf, observations, sampling, soil data exchange

OGC Soil Data Interoperability Experiment

1 Introduction

1.1 Scope

This Open Geospatial Consortium (OGC®) engineering report describes the use cases, modeling, implementation, results and findings of the Soil Interoperability Experiment. This OGC® document is applicable to the sharing of soil data. The document reviews previous work, defines the conceptual, logical and physical models created by those involved in the interoperability experiment and provides an overview of the demonstration tools.

The International Union of Soil Sciences (IUSS) Working Group on Soil Information Standards (WGSIS) is working to consolidate the various existing soil information models and reconcile them into a single language for the exchange of globally consistent soil information. The WGSIS chose to progress this work through an OGC Interoperability Experiment.

1.2 Document contributor contact points

All questions regarding this document should be directed to the editor or the contributors:

Name	Organization
Alistair Ritchie	Landcare Research New Zealand Ltd
David Medyckyj-Scott	Landcare Research New Zealand Ltd
Pierre Roudier	Landcare Research New Zealand Ltd
Bruce Simons	CSIRO, Australia
Peter Wilson	CSIRO, Australia
Jorge Mendes de Jesus	ISRIC — World Soil Information, The Netherlands
Peter Dahlhaus	Federation University of Australia
Andrew MacLeod	Federation University of Australia
Giovanni L'Abate	Agribiology and Pedology Research Centre, Italy
Jessica Lucido	US Geological Survey (USGS)
Joshua Lieberman	Harvard University, USA
Paul Finnell	USDA-Natural Resources Conservation Service
Dave Hoover	USDA-Natural Resources Conservation Service
Brent Watson	Horizons Regional Council, New Zealand

1.3 Revision history

Date	Release	Editor	Primary clauses modified	Description
2016-04-01	0.1	Alistair Ritchie	All	First draft
2016-04-04	0.2	Bruce Simons	All	First read through
2016-04-06	0.3	Alistair Ritchie	All	Worked through proposed changes, augmented model comparison tables.
2016-04-07	0.4	David Medyckyj-Scott	All	First read through with proposed edits
2016-04-09	0.5	Alistair Ritchie	All	Worked through proposed changes, additional content where requested.
2016-04-12	0.6	Peter Wilson	All	First read through with proposed edits
2016-04-13	0.7	Bruce Simons	All	Minor edits.
2016-05-04	0.8	Alistair Ritchie	All	Updates from community review.
2016-05-05	0.9	David Medyckyj-Scott	All	Moved initial drafts into the OGC template
2016-05-22	0.10	Alistair Ritchie	All	Final draft.
2016-07-20	1.0	Scott Simmons	All	Edit for publication

1.4 Future work

This model forms the basis of a normative specification to be defined by a future OGC Soil Data Standards Working Group.

1.5 Forward

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

2 References

The following documents are referenced in this document. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the latest edition of the normative document referred to applies.

ISO: ISO 19101:2002 – Geographic information – Reference model. (2002).

ISO: ISO/TS 19101-2:2008 – Geographic information – Reference model – Part 2: Imagery (2008)

ISO: ISO/TS 19103:2005 – Conceptual Schema Language (2005)

ISO: ISO 19109:2005 – Geographic information – Rules for application schema (2005)

ISO: ISO 19123:2005 – Geographic information – Schema for coverage geometry and functions (2005)

ISO: ISO 19135-1:2015 – Geographic information – Procedures for item registration – Part 1: Fundamentals (2015)

ISO: ISO 19136:2007 – Geographic information – Geography Markup Language (GML) (2007)

ISO: ISO 19143:2010 – Geographic information – Filter encoding (2010)

ISO: ISO 19156:2011 – Geographic information – Observations and measurements (aka OGC 10-004r3) (2011)

OGC: OGC 05-007r7 – OpenGIS Web Processing Service (2007)

OGC: OGC 06-009r6 – Sensor Observation Service (2007)

OGC: OGC 06-121r3 – OGC Web Services Common Specification (2007)

OGC: OGC 08-094r1 – OGC SWE Common 2.0 (2011)

OGC: OGC 09-025r1 – OpenGIS Web Feature Service 2.0 Interface Standard (aka ISO/DIS 19142) (2010)

OGC: OGC 09-110r4 – OGC WCS 2.0 Interface Standard- Core: Corrigendum (2012)

OGC: OGC 12-006 – OGC Sensor Observation Service Interface Standard (2012)

OGC: OGC 15-043r3 – Timeseries Profile of Observations and Measurements (2016)

Auscope PID: Persistent Identifier Service, https://www.seegrid.csiro.au/wiki/Siss/PIDService

UKGovLD: Registry Core, https://github.com/UKGovLD/registry-core/wiki

W3C: WEBARCH – Architecture of the World Wide Web, https://www.w3.org/TR/webarch/

W3C: LDP – Linked Data Platform, https://www.w3.org/TR/ldp/

W3C: RDF – Resource Description Framework (RDF), https://www.w3.org/RDF/

W3C: SKOS – Simple Knowledge Organization System (SKOS), https://www.w3.org/TR/skos-reference/

3 Terms and definitions

For the purposes of this report, the definitions specified in Clause 4 of the OWS Common Implementation Standard [OGC 06-121r3] shall apply. In addition, the following terms and definitions apply.

3.1
application schema
Conceptual schema for data required by one or more applications.

[ISO 19101:2002, definition 4.2]

3.2
coverage
Feature that acts as a function to return values from its range for any direct position within its spatial, temporal or spatiotemporal domain.

[ISO 19123:2005, definition 4.17]

3.3
data type
Specification of a value domain with operations allowed on values in this domain.

[ISO/TS 19103:2005, definition 4.1.5]

EXAMPLE Integer, Real, Boolean, String, Date (conversion of a date into a series of codes).

NOTE Data types include primitive predefined types and user-definable types. All instances of a data type lack identity.

3.4
domain feature
Feature of a type defined within a particular application domain.

NOTE: This may be contrasted with observations and sampling features, which are features of types defined for cross-domain purposes.

[ISO 19156:2011, definition 4.4]

3.5

element <XML>

Basic information item of an XML document containing child elements, attributes and character data.

NOTE: From the XML Information Set — each XML document contains one or more elements, the boundaries of which are either delimited by start-tags and end-tags, or, for empty elements, by an empty-element tag. Each element has a type, identified by name, sometimes called its 'generic identifier' (GI), and may have a set of attribute specifications. Each attribute specification has a name and a value.

[ISO 19136:2007, definition 4.1.23]

3.6

feature

Abstraction of a real-world phenomena.

[ISO 19101:2002, definition 4.11]

3.7

GML application schema

Application schema written in XML Schema in accordance with the rules specified in ISO 19136:2007.

[ISO 19136:2007, definition 4.1.34]

3.8

GML document

XML document with a root element that is one of the elements AbstractFeature, Dictionary or TopoComplex, specified in the GML schema or any element of a substitution group of any of these elements.

[ISO 19136:2007, definition 4.1.35]

3.9

GML profile

Subset of the GML schema.

[ISO 19136:2007, definition 4.1.36]

3.10

GML schema

Schema components in the XML namespace ―http://www.opengis.net/gml/3.2‖ as specified in ISO 19136:2007.

[ISO 19136:2007, definition 4.1.37]

3.11
Linked Data
The use of HTTP for accessing, updating, creating and deleting resources from servers that expose their resources according to these rules of Linked Data: use URIs as names for things; use HTTP URIs so that people can look up those names; when someone looks up a URI, provide useful information, using the standards (RDF, SPARQL); include links to other URIs, so that they can discover more things.

[W3C LDP]

3.12
measurement
Set of operations having the objective of determining the value of a quantity.

[ISO/TS 19101-2:2008, definition 4.20]

3.13
observation
Act of observing a property.

NOTE: The goal of an observation may be to measure or otherwise determine the value of a property.

[ISO 19156:2011 definition 4.11]

3.14
observation procedure
Method, algorithm or instrument, or system which may be used in making an observation.

[ISO19156:2011, definition 4.12]

3.15
observation result
Estimate of the value of a property determined through a known procedure.

[ISO 19156:2011, definition 4.14]

3.16
property
Facet or attribute of an object referenced by a name.

[ISO 19143:2010, definition 4.21]

EXAMPLE: Abby's car has the color red, where "color red" is a property of the car instance.

3.17
register
Set of files containing identifiers assigned to items with descriptions of the associated items.

[ISO 19135-1:2015, definition 4.1.9]

3.18
registry
Information system on which a register is maintained.

[ISO 19135-1:2015, definition 4.1.13]

3.19
resource
Items of interest on the World Wide Web that are identified by global identifiers called Uniform Resource Identifiers.

[W3C WEBARCH]

3.20
sampled feature
The real-world domain feature of interest, such as a groundwater body, aquifer, river, lake, or sea, which is observed.

[ISO 19156:2011]

3.21
sampling feature
Feature, such as a station, transect, section or specimen, which is involved in making observations of a domain feature.

NOTE: A sampling feature is purely an artefact of the observational strategy, and has no significance independent of the observational campaign.

[ISO 19156:2011, definition 4.17]

3.22
schema <XML Schema>
XML document containing a collection of schema component definitions and declarations within the same target namespace.

Example Schema components of W3C XML Schema are types, elements, attributes, groups, etc.

NOTE: The W3C XML Schema provides an XML interchange format for schema information. A single schema document provides descriptions of components associated with a single XML namespace, but several documents may describe components in the same schema, i.e. the same target namespace.

[ISO 19136:2007, definition 4.1.54]

3.23
sensor
Type of observation procedure that provides the estimated value of an observed property at its output.

Note: A sensor uses a combination of physical, chemical or biological means in order to estimate the underlying observed property. At the end of the measuring chain electronic devices often produce signals to be processed.

[OGC 08-094r1, definition 4.5]

3.24
timeseries
Sequence of data values which are ordered in time.

[OGC 15-043r3, definition 4.9]

4 Conventions

4.1 Abbreviated terms

ANZSoilML	Australia and New Zealand Soil Mark-up Language
API	Application Program Interface
CSIRO	Commonwealth Scientific and Industrial Research Organisation
SOTER	Soil and Terrain Database
GeoSciML	GeoScience Mark-up Language
GML	OGC Geography Mark-up Language
GWML2	Groundwater Mark-up Language, version 2.0
INSPIRE	Infrastructure for Spatial Information in the European Community
ISO	International Organization for Standardization

ISRIC	ISRIC – World Soil Information
OGC	Open Geospatial Consortium
O&M	OGC Observations and Measurements Abstract Specification
OMXML	OGC Observations and Measurements XML Implementation
RDF	Resource Description Framework
SISS	AuScope Spatial Information Services Stack
SKOS	Simple Knowledge Information System
SoilIEML	Soil Interoperability Experiment Mark-up Language
SoilML	Soil Mark-up Language (ISO 28258:2013)
SOS	Sensor Observation Service
THREDDS	Thematic Real-Time Environmental Distributed Data Services
TSML	Timeseries Mark-up Language
UML	Unified Modeling Language
UTC	Coordinated Universal Time
URI	Universal Resource Identifier
URL	Universal Resource Locator
WaterML2	Water Mark-up Language, version 2.
WCS	Web Coverage Service
WFS	Web Feature Service
WGSIS	International Union of Soil Sciences (IUSS) Working Group on Soil Information Standards
WMS	Web Map Service
WPS	Web Processing Service
XML	Extensible Markup Language
XSD	W3C XML Schema Definition Language

4.2 UML notation

The diagrams that appear in this standard are presented using the Unified Modeling Language (UML) static structure diagram, as described in Subclause 5.2 of [OGC 06-121r3].

5 Soil Interoperability Experiment Overview

The promotion of sustainable soil and land management is critical to ensuring productive food systems, improved rural livelihoods and a sustainable healthy environment. A number of large environmental and production challenges have been identified as ongoing threats to the sustainability of soil resources across the world. These threats are complex, difficult to resolve, and are inextricably linked across a number of environmental and land management domains. The need for an improved understanding of soil distribution, function and state is important to support science and policy development and to improve productivity in a sustainable manner. This understanding needs to be underpinned by soil data and information that can be organized, aggregated and made accessible in a consistent, granular and consumable form.

Modern digital mapping and modeling techniques are of huge value in providing a more quantitative estimate of the variation of properties across the imperfectly sampled soil resource. However, they depend on dispersed, inconsistent and difficult to access digital data. It is increasingly clear that modern, harmonized and interoperable information systems are required to integrate these data into a consistent set of soil data. Various initiatives have started work on this by defining soil information models. Examples include: the European INSPIRE, e-SOTER and ISO 28258:2013 [6] (ISO SoilML) standards; the international GlobalSoilMap consortium, and the Australian and New Zealand ANZSoilML projects. The UN FAO Global Soil Partnership recognizes the value of all of these projects, particularly for Pillars 4 (Global Soil Information System) and 5 (Harmonization). Yet despite these endeavors, soil scientists still have a situation where they must reconcile multiple systems that often attempt to do the same thing.

The International Union of Soil Sciences (IUSS) Working Group on Soil Information Standards (WGSIS) is working to consolidate these information models and reconcile them into a single language for the exchange of globally consistent soil information. The WGSIS aimed to progress the work by running an Open Geospatial Consortium Interoperability Experiment.

This experiment had the goal of defining a simplified, yet implementable, soil information model (referred to as SoilIEML for the purposes of this document) by consolidating core concepts and features from existing standards, and testing the result (through working implementations) against an agreed set of use cases for the exchange and analysis of soil data.

Participants were not expected to change their information systems; instead they provided web services using a common exchange language and protocols used to communicate with all components in the system. Those undertaking the experiment gained valuable insights that can be used to improve existing information systems or fast-track the development of new systems.

Harmonizing inconsistent data is a time-consuming process – some estimates state that 80% of a scientist's time can be occupied getting data into a state where it can be analyzed ('data wrangling'). This process must be repeated with each project. This initiative aims to reverse that by allowing data acquisition to be automated as much as possible, allowing more time to be spent understanding and managing our soils.

6 Objectives of the Soil Interoperability Experiment

The objectives of the Soil Interoperability Experiment were:

— Further develop and test a GML compatible information model for soil features, based on existing international initiatives;

— Seek participation by a number of soil agencies from across the world to establish data services using the defined model;

— Demonstrate delivery of standardized soil data services from multiple and disparate sources, and the use of these services within a number of user focused applications; and

— Prepare an OGC/IUSS engineering report with intent to develop it into a data specification subsequent to the IE.

The Technical Approach for this Interoperability Experiment followed existing principles for developing information models as exemplified by GeoSciML, GWML2 and WaterML2.

Testing involved:

— Evaluation of the GML-UML against GML principles;

— Validation of the XSD;

— Evaluation of the functionality of web services;

— Evaluation of instance documents against use-cases; and

— Evaluation of data exchange between participants.

7 Use Cases

Four use-cases were developed for the IE to constrain the requirements of the information model and test implementations. The use cases addressed scenarios commonly encountered by providers and users of soil data. The use cases are described in detail in Annex A.

Use cases 1 to 3 involve the delivery of data describing soil features using web services while use case 4 involves the use of algorithms or other processing methods to generate new data. A fifth use case addressing soil type mapping was proposed but not developed due to a lack of resources.

7.1 Use Case 1: Soil Data Integration

This use case involves using the schema as a canonical structure into which heterogeneous soil data formats are transformed and published. An example application is the exchange of comprehensive soil description, sampling and analytical data captured at a site.

7.2 Use Case 2: Soil Sensor Data

This use case involves identifying the location and properties of sensors monitoring dynamic soil properties, and the provision of the measurements made at those sites, to inform interested parties about local soil conditions. An example application is the provision of soil moisture monitoring data to control variable rate precision irrigation.

7.3 Use Case 3: Soil Property Modeling and Predictions

This use case involves the provision of high resolution estimates of functional soil properties generated using digital soil mapping techniques to support the development and implementation of policies for the management of the soil resources. An example application is the publication of the world-wide soil property predictions made by the GlobalSoilMap project (http://www.globalsoilmap.net/).

7.4 Use Case 4: Pedo-transfer Functions

This use case involves the delivery of observed and interpreted soil properties (by soil type and/or by spatial distribution) in a standard format that allows the use of pedo-transfer functions - algorithms that calculate additional interpreted soil properties. An example application is the provision of data to a processing service that predicts the values of relevant functional properties of a soil to inform farm management practices.

In a fully interoperable system there would be interdependencies between the use cases.

1. Use case 4 would use data published by services implemented to provide data according to use cases 1, 2 and 3.

2. Use case 3 would include data from services supporting use cases 1, 2 and 4 (along with other environmental datasets) when generating the predictions.

8 Comparison of Existing Soil Information Models

Various national and international initiatives have worked on information models to support the exchange of soil data. This experiment intended to reconcile core concepts and features of these models into a single coherent, fully attributed, implementable, albeit provisional, standard. The five models that were reviewed in this IE were:

1. Australia and New Zealand Soil Mark-up Language (ANZSoilML);

2. e-SOTER Soil and Terrain Mark-up Language (SoTerML);

3. INSPIRE D2.8.III.3 Data Specification on Soil (INSPIRE Soil);

4. ISO 28258:2013 Soil quality – Digital exchange of soil-related data (ISO SoilML); and

5. IUSS/ISO ‘Wageningen Proposal’ (a variation of 4 with reference to 1, 2 and 3).

This section provides an overview of the models, comparing their breadth of content and modeling technique. A summary of each model can be found in Annex B.

8.1 Basis for comparison

The models were compared according to their scope, the modeling techniques and patterns used in their definition, whether they are readily available and implemented, and the context in which they have been used (e.g. production or prototype). Only documents that could reasonably be considered normative and were of unambiguous origin were used for the comparison (UML models and XML Schema Documents).

Scope refers to the breadth of information captured by each model. As each model has a different level of abstraction and conceptual base an external basis for comparison was necessary. The FAO Guidelines for Soil Description [3], with additional guidance from the USDA Field Book for Describing and Sampling Soils [11], was adapted to define and group the key dimensions of a soil. The comparison is only intended to be indicative of scope and does not imply strict conformance to the FAO and USDA guidelines.

The criteria for comparison can be grouped into eight categories:

1. Site registration: identity, location, timing and other metadata about sampling sites;
2. Soil formation: environmental and human factors influencing the formation of the soil;
3. Soil description: physical, chemical and organic character of a soil;
4. Sampling: collection of physical samples;
5. Observation: field or laboratory measurements of soil properties;
6. Classification: categorization of soil and horizons according to formal taxonomies;
7. Vocabularies: systems for managing terms and their definitions; and
8. Mapping: mapping the distribution of soils according to their type.

Modeling approach refers to how the model was defined. Models may be:

— comprehensive, attempting to cover as many dimensions of the soil as possible and favoring hard-typed properties (where properties are explicitly defined and bound to classes). Additional properties may be specified using a soft-typing mechanism (values presented as property-value-pairs where the property refers to a dynamic register of properties);

— targeted, hard-typing a selected set of essential properties while relying on soft-typing for a significant set of properties; or

— framework, a model that simply provides a framework of classes. Soft-typed properties are used almost exclusively.

Accessibility refers to the availability of the model (UML models, XML schema, specification documents and other artefacts) in terms of access constraints or charges for access.

Implementation readiness attempts to capture whether a model may be implemented (e.g. deployed as web services) with comparative ease (high), some difficulty but minimal technical impediment (medium) or with significant effort and/or technical impediments (low).

Implementation shows whether a standard has been implemented, either as a prototype in a production environment.

8.2 Comparison of models

Table 1 is a comparison of the modeling approach and scope of each information model and Table 2 summarizes matters of implementation. No existing model was accepted by the IE participants as the primary basis for a soil information model, due to:

— Lack of agreement on the concepts in each of the models;

— Incompleteness due to a reliance on soft-typing.

The soil model developed as part of this IE (SoilIEML), although not part of the analysis of existing models, is included in Table 1 for completeness.

Of the models, ANZSoilML was the most comprehensive and explicit model. However, it was also the most difficult to understand as it relied on a single model for conceptual and physical purposes. In contrast, SoTerML and INSPIRE Soil were less-comprehensive as they took a targeted modeling approach, explicitly defining some properties, but relying on soft-typing for most soil properties. ISO SoilML and the modified ISO/IUSS Wageningen proposal are the least comprehensive models, instead providing a framework of the highest level concepts, without detailing their associated properties.

Soft-typing, allowing generic property-value pairs rather than explicitly specifying property-data type values, has value in physical models used to serialize the data, but at the abstract level simply defers the problem of model definition and therefore encourages independent proliferation of models and implementations. Given substantial effort can be required to agree on even simple property lists, the value of these framework and targeted models is limited for many communities as it hampers fast and consistent adoption. A comprehensive and explicit model is required to reduce the need for this work and to support consistent implementations. This may be accompanied by a more flexible implementation that is nonetheless constrained by the conceptual model.

Explicit conceptual models can be published in normative feature catalogs or data type registers but we could find no obvious authoritative examples for the INSPIRE or SoTerML standards. Searches did uncover a Semantic Web European Soil Thesaurus (https://secure.umweltbundesamt.at/soil/) but its status is unclear. Meanwhile, ISO SoilML explicitly defers the problem, expecting '*… that every data provider extends SoilML classes with properties which are suitable for his or her data model, or simply fill the attributes for which (s)he has data.*' [7].

Soft-typing also causes problems when the property values are complex and require structured data types. Also, many properties describe material components of a soil and need to be related to one another to fully describe the structure, genesis or function of a soil. Without clear guidance as to the organization of these complex types interoperability breaks down. As such, soft-typed models are only reliable for the provision of simple property-value statements about a soil, and even then necessitating the use of well-defined property registers.

The ANZSoilML, INSPIRE Soil and eSoTerML models are implementation ready (see Annex B for details), with GML schema for the logical models available at no cost. Production ANZSoilML services have been deployed in both Australia and New Zealand. INSPIRE Soil and ISO SoilML have been implemented for the European Union GS Soil project (http://www.eurogeosurveys.org/projects/gssoil/). However, the ISO SoilML cannot be implemented as is. There is no normative schema or register of soil properties and no mechanism for providing the property values. Properties are expected to be assigned to classes in other application and XML schema that extend ISO 28258:2013 – for example, the GS Soil project created a 'provider specific' extension of the schema for portal services [1].

Table 1 — Comparison of existing soil information models

Green cells (F) denote aspects of a soil description covered by classes and formally defined properties; yellow cells (S) denote aspects of a soil description covered by classes and soft-typed properties; white cells (X) denote aspects of a soil description that are not covered, or handled by a soft-typing mechanism.

Model / Soil Information	ANZSoilML	eSoTer	INSPIRE Soil	ISO SoilML	IUSS/ISO SoilML	OGC Soil Data IE
Site Registration	F	S	F	F	F	F
Identity	F	S	F	F	F	F
Location	F	S	F	F	F	F
Observation date/time	F	S	F	x	x	x
Author	S	S	x	x	x	x
Soil Formation	F	x	x	x	x	x
Climate/Weather	F	x	x	x	x	x
Landform	F	S	x	x	x	x
Land use	F	S	x	x	x	x
Vegetation	F	S	x	x	x	x
Human	F	x	x	x	x	x
Parent Material	F	S	F	x	x	x
Soil Description	F	S	x	x	x	S
Surface character	F	S	x	x	x	S
Horizon boundary	F	S	F	F	F	F
Primary Constituents	F	S	S	x	x	S
Soil colour (matrix)	F	S	S	x	x	S
Mottling	F	S	S	x	x	S
Redox	S	S	S	x	x	S
Carbonates	F	S	S	x	x	S
Gypsum	S	S	S	x	x	S
Soluble Salts	S	S	S	x	x	F
Field pH	F	S	S	x	x	F

Model Soil Information	ANZSoilML	eSoTer	INSPIRE Soil	ISO SoilML	IUSS/ISO SoilML	OGC Soil Data IE
Odour	S	S	S	x	x	S
Organic Matter	S	S	S	x	x	S
Soil Texture	F	S	S	x	x	F
Soil Structure	F	S	S	x	x	S
Consistence	F	S	S	x	x	S
Soil-water status	F	S	F	x	x	F
Bulk density	F	S	S	x	x	F
Porosity	F	S	S	x	x	S
Concentrations	F	S	S	x	x	S
Roots	F	S	F	x	x	S
Contamination	S	S	F	x	x	S
Bedrock	F	S	S	x	x	S
Soil Sampling	F	S	F	F	F	F
Location	F	S	F	F	F	F
Method	F	S	S	F	F	S
Archive	F	x	x	F	F	F
Observations	F	S	S	F	F	F
Method	F	S	S	F	F	S
Result	F	S	F	F	F	F
Soil Classification	F	F	F	F	F	F
Soil	F	F	F	F	F	F
Horizon	F	F	F	x	x	F
Vocabularies	F (SKOS)	x	F (Codelist)	x	x	F (SKOS)
Soil Mapping	F	F	F	F	F	F

Table 2 — Comparison of model accessibility and implementation

Model	Accessibility	Implementation
ANZSoilML	Public; free of charge	Production services
eSoTer	Public; free of charge	Unknown
INSPIRE Soil	Public; free of charge	GS Soil Project
ISO SoilML	Private; charge for access	GS Soil Project
IUSS/ISO SoilML	No model	No

9 Soil Interoperability Experiment Information Models

9.1 Introduction

As there was no suitable starting model selected from existing candidates, a compromise model was developed using classes and patterns from the existing models wherever possible. The initial focus for modeling was on supporting the IE use cases but the IE Participants were mindful of allowing extension to support soil mapping use cases.

With only a 6-month time frame available for the Soil IE, and with the need to be able to map multiple data providers' data to the model, implement services and develop demonstration clients, minimal soil data modeling was carried out. To constrain the model's scope, the selection of properties to be assigned to soil classes was based on the specifications of the GlobalSoilMap consortium [4], with the intention to allow the model to expand in a modular fashion.

Three models were produced by the IE.

1. A UML conceptual model defining the classes and properties needed to comprehensively describe a soil resource according to the needs of the IE use cases. This model is technology independent.

2. A UML logical model converting the conceptual model into an ISO19109:2005 compliant application schema and formally importing models from other OGC specifications: ISO 19156:2011 – Observations and Measurements (O&M) and OGC 15-043r3 – the Timeseries Profile of Observations and Measurements.

3. An XML physical model defining the encoding of data using Geography Mark-up Language according to the ISO19136:2007.

9.1.1 Modeling Principles

To constrain modeling activity and help ensure that the model defined was not disruptive (conforming closely to existing OGC standards) or parochial (recognizing that soil scientists must integrate data from many environmental domains), the group defined a set of guiding principles, including:

- Open publication of the results (UML, XSD) once approved through the OGC process;

- Open development of the model, subject to appropriate agreements with the OGC;

- Re-use of existing models wherever possible (for example Observations and Measurements, Timeseries Profile of Observations and Measurements, GeoSciML); and

— If an existing model does not meet project needs, then work with the relevant community to address issues before resorting to an extension or branch of the model.

The motivation for the principles was twofold:

— To ensure consistency of data types across domains - soil is a function of geology, climate, topography, and biology, and pedologists must aggregate these data during analysis; and

— To ease process of deployment – a client developed for geology or climate observations encoded as O&M can be used for soil observations without modification (other than dealing with domain specific values).

There was a simplification of models from the conceptual to logical to physical viewpoints when classes explicitly defined at a higher level could be realized without loss of information by a single class. This is an equivalent practice to O&M's virtual typing strategy.

9.1.2 Treatment of Mandatory Properties

In most cases cardinality constraints on properties and associations were removed in the logical and physical model. This allowed the definition of a strict conceptual model that was very clear about the properties and associations needed to comprehensively and meaningfully describe a resource, but permitted a terse digital encoding of data when necessary.

Conceptually, a distinction is made between properties that are mandatory (without them the data has no meaning) and mandatory but voidable (they should be provided but may not be for some reason). In the translation to the logical and physical models, mandatory properties must remain so, but voidable properties become optional. It is up to the community deploying the services to define implementation profiles that define which properties must be present in that context.

Providers can then deploy services that can lie on an appropriate continuum of use case support from comprehensive encodings (where all data are provided and reasons are provided if data are missing) to terse encodings (where only the bare minimum of data required for a specific application are transmitted). Applications may also switch between the two, requesting a terse encoding to display property values, but using a comprehensive data service where appropriate, such as to explore meta-data about how, when and to what quality a value was generated, for dataset exchange or when sending data on to web processing services. Profiles identifying which aspects of the logical model services are expected to provide will require further community agreement.

9.1.3 Observations as a Data-type

Throughout the model the Observation class is used as the data-type for estimated property values, allowing them to be attributed with meta-data about the value's provenance and quality. This is important as a soil property may be measured in several ways, to differing levels of quality, and both affect how values are interpreted or processed.

Using O&M to explicitly decouple procedures, properties and results keeps the number of properties defined in the model manageable. Otherwise, the common practice of defining a property as a combination of the phenomenon observed, how it was measured, for example pH_NaCl and pH_KCl (both describe soil pH, but using different procedures), causes significant increase in properties defined. This not only introduces an inconsistency with O&M, but also reduces interoperability and creates a governance overhead, requiring the core information model to be updated for newly developed processes.

9.2 Conceptual Model

The conceptual model is designed as a technology neutral representation of the semantics of the soil domain. It is defined using UML and describes four aspects of the domain:

1. the description and classification of soil;

2. the soil profile;

3. sampling and field/laboratory observations; and

4. sensor-based monitoring of dynamic soil properties.

In accordance with the use cases defined for this Interoperability Experiment it excludes:

— soil mapping; and

— landscape, land-use, climate and vegetation data (these are recognized as important, but best modelled in collaboration with experts from those domains).

A set of utility classes was defined in this model to provide placeholders for complex data types and classes and are described in Annex C.

A soil (Figure 1) is a body of unconsolidated mineral and organic material developed through soil forming processes near the surface of the earth. It is composed of horizons defined according to their physical, chemical or biological characteristics. Horizons are given a designation according to their character and position and a soil can then be classified according to one of a number of soil taxonomies addressing a soil's form or function.

A description of a soil comprising a vertical section of horizons is called a soil profile (Figure 2). Soil profiles usually describe a sequence at a sampling location (for example, an excavation) but this definition may be extended to include aggregate descriptions (such as a soil profile class) of the expected variation of soil properties over a mapped area.

Parts of a soil may be extracted and analyzed, either immediately in the field, or by subsequent analysis of a physical sample in a laboratory (Figure 3). It is essential to be able to record the method of sampling and analysis, and the quality or uncertainty of the resulting measurement. The results of these measurements may be presented and used in their own right or be incorporated into a soil description.

In contrast to inherent soil properties that change very little, certain soil properties change in response to external forces over measurable (human-scale) periods of time. These dynamic properties are essential to understanding soil, and broader environmental

Figure 1 — An exposed section of soil in farmland.

This soil is classified, according to the New Zealand Soil Classification, as a 'Podzol Soil'. (Image copyright of Landcare Research. All rights reserved. Permission from Manaaki Whenua: Landcare Research New Zealand Limited must be obtained before the re-use of this image.)

conditions and health, particularly where change is a function of human activity (for example land use or management practices). Change over time may be monitored through a variety of techniques from episodic visits to established monitoring sites or from in-situ sensors measuring one or more properties of a soil (Figure 4). Although not limited to these, the dynamic properties include:

— soil moisture content;

— soil temperature;

— bulk density; and

— organic carbon content.

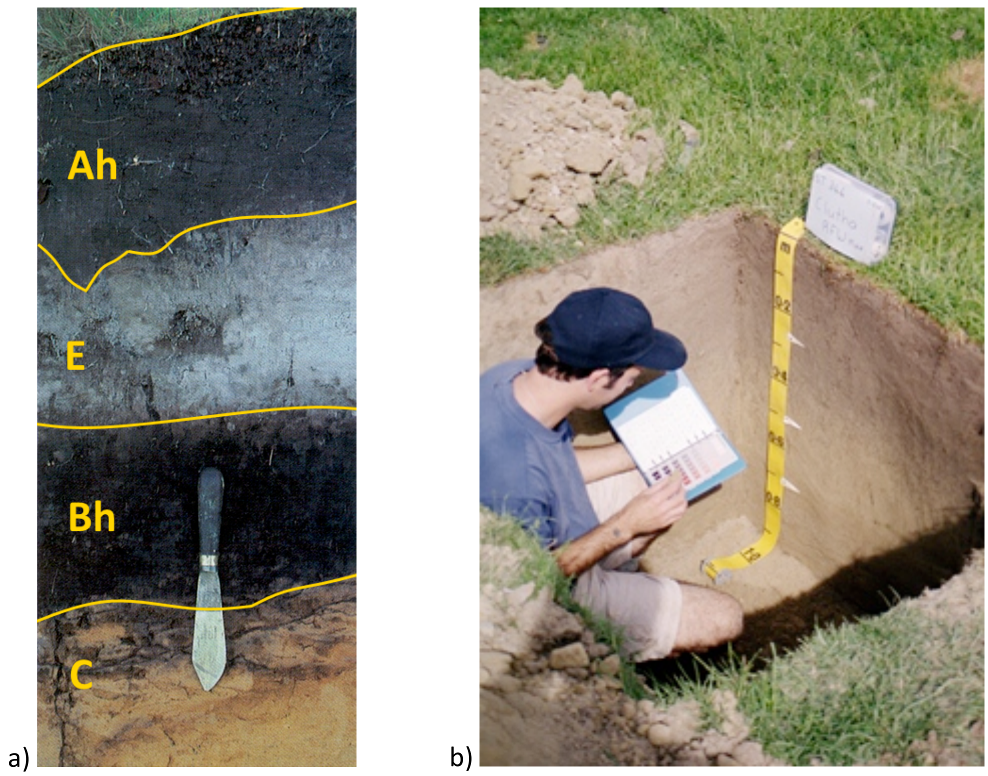


Figure 2 — Soil profiles and sampling.

(a) a soil profile in a Podzol Soil, broken into four horizons, each of which have been designated according to the New Zealand Soil Description Handbook; (b) an example of a pit dug to expose a profile. (Images copyright of Landcare Research. All rights reserved. Permission from Manaaki Whenua: Landcare Research New Zealand Limited must be obtained before the re-use of this image.)

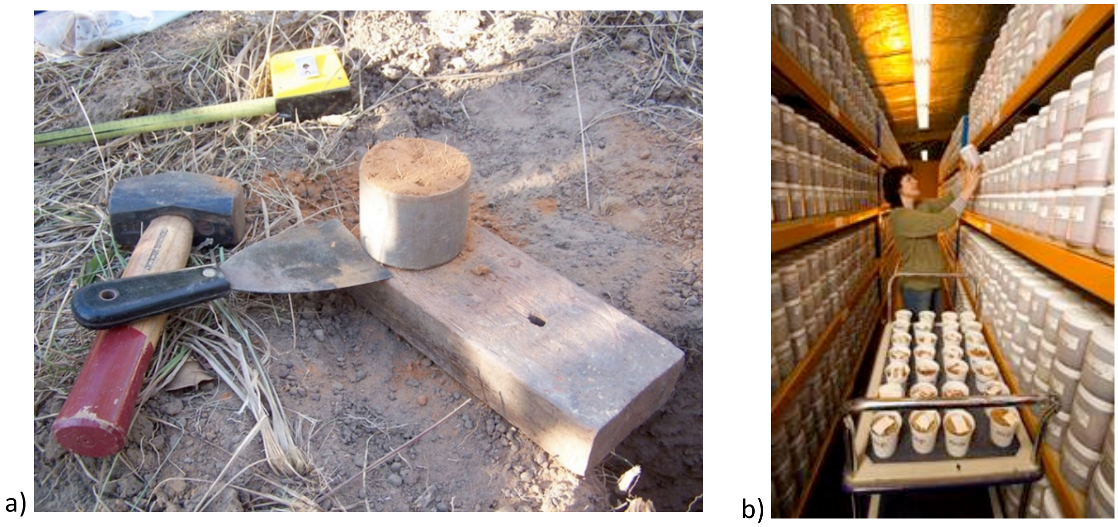


Figure 3 — Soil specimens.

(a) A ring sampler used to extract a volume of soil for bulk density measurements; (b) soil specimens archived for long term storage at CSIRO's National Soil Archive in Canberra, Australia. (Photo (a): T. Mendel. Images copyright of CSIRO.)

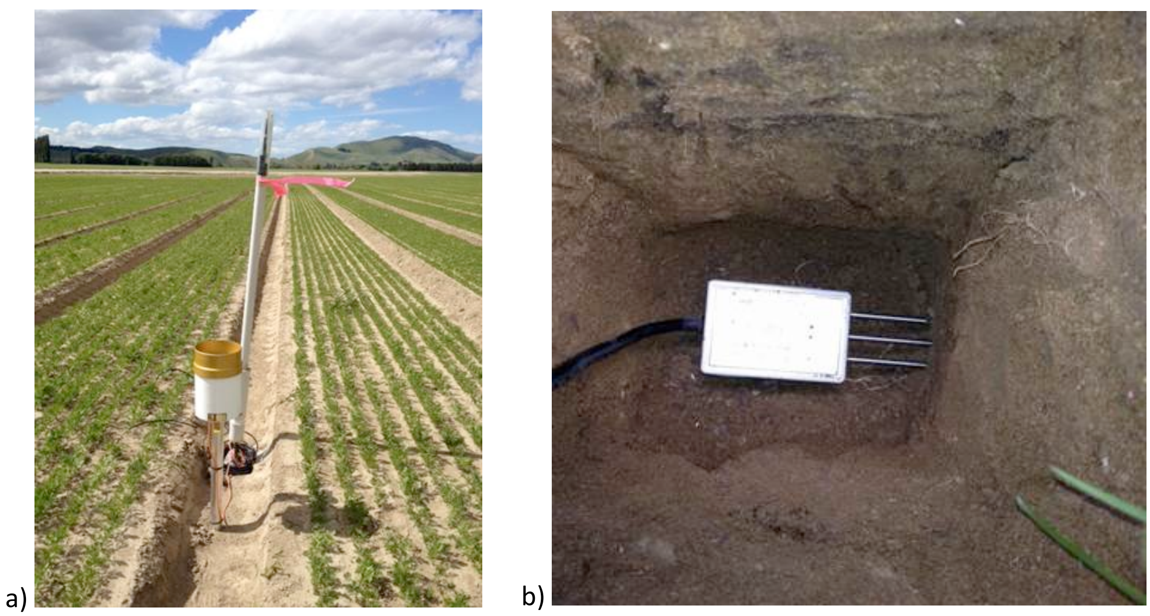


Figure 4 — Soil moisture monitoring site.

(a) At the surface with rain gauge and communications hardware; and (b) the buried moisture soil moisture probe. (Photos: Jagath Ekanayake, Landcare Research.Images copyright of Landcare Research. All rights reserved. Permission from Manaaki Whenua: Landcare Research New Zealand Limited must be obtained before the re-use of this image.)

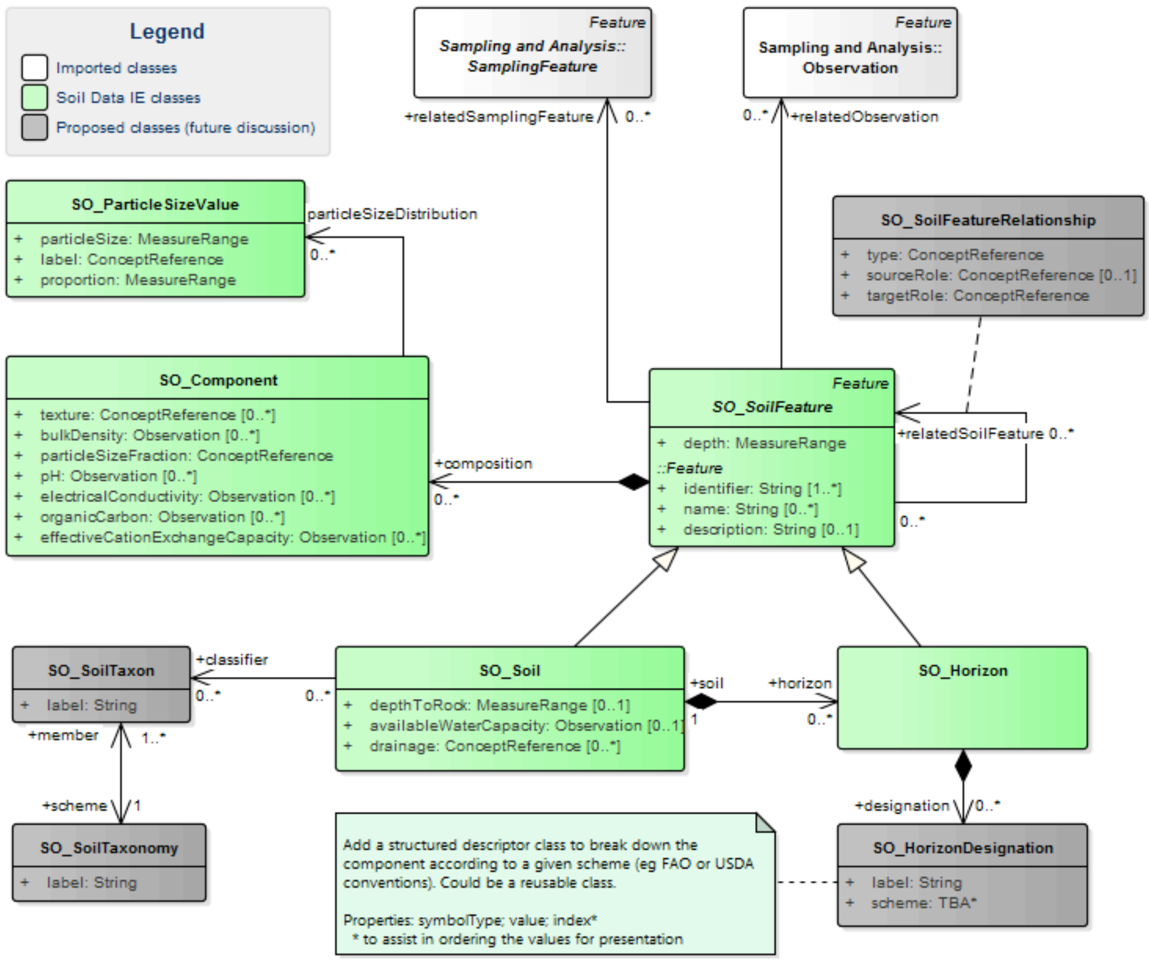


Figure 5 — Class model for the description and classification of soil.

9.2.1 Soil Description

Soil features (Figure 5) are designed to describe and classify a volume of soil and its horizons. The soil may be described as a whole (SO_Soil) or according to horizons (SO_Horizon). Both are sub-classes of the abstract class SO_SoilFeature, which provides classes to capture composition according to soil size fractions (SO_Component) and, although not tested in this IE, the relationships between soils and horizons (SO_SoilFeatureRelationship).

9.2.1.1 SO_SoilFeature

The abstract SO_SoilFeature class is a conceptual feature that exists coherently in the world. It corresponds with a soil body, comprising a set of similar soil profiles with a mappable distribution in some soil landscape, or a soil profile or its constituent parts. The implemented SO_SoilFeature instance acts as the description package. ANZSoilML, following patterns established in GeoSciML, classifies the SO_SoilFeature description according to its purpose as an ‘instance’ (capturing an observation), ‘typical norm’ (capturing a description of the expected range of property values in a soil body), or ‘defining norm’ (capturing the necessary criteria to be classified as a particular class of soil). In contrast, INSPIRE uses separate classes to distinguish observed descriptions from derived descriptions. Neither option was tested during the IE.

Name	**Definition**	**Data Type**	**Multiplicity**
depth	The upper (nearest to Earth's surface) and lower boundary of the described soil or horizon.	MeasureRange	One
composition	The physical or chemical composition of a soil.	SO_Component	Zero or one
relatedSoilFeature	Relates a soil feature to another soil feature.	SO_SoilFeature Relationship	Zero to many
relatedObservation	Any observation made on the soil.	Observation	Zero to many
relatedSamplingFeature	Sampling features related to the soil. May be where the soil was observed or where one or more samples were taken from the soil.	SamplingFeature	Zero to many

9.2.1.2 SO_Soil

The Soil Science Glossary [12] defines a soil as “*(i) The unconsolidated mineral or organic material on the immediate surface of the Earth that serves as a natural medium for the growth of land plants. (ii) The unconsolidated mineral or organic matter on the surface of the Earth that has been subjected to and shows effects of genetic and environmental factors of: climate (including water and temperature effects), and macro- and microorganisms, conditioned by relief, acting on parent material over a period of time. A product-soil differs from the material from which it is derived in many physical, chemical, biological, and morphological properties and characteristics.*”

Name	**Definition**	**Data Type**	**Multiplicity**
depthToRock	The depth to rock. Also known as the R horizon depth or total profile depth.	MeasureRange	Zero or one
availableWaterCapacity	Available water capacity, measured or computed for a set of specified depth increments using a pedo-transfer function that references the values estimated for organic carbon, sand, silt, clay and bulk density.	Observation	Zero or one
drainage	The local soil wetness conditions. For example, 'very poorly drained', 'poorly drained' or 'well-drained’.	Concept	Zero to many
classifier	Reference to a classification of a soil according to a formal soil taxonomy.	SO_SoilTaxon	Zero to many
horizon	Related soil horizons.	SO_Horizon	Zero to many

Although these represent only a limited set of soil properties commonly recorded, they were considered important for the IE use cases providing a variety of data types to test encodings.

9.2.1.3 SO_Horizon

The Soil Science Glossary defines a soil horizon as a "*layer of soil or soil material approximately parallel to the land surface and differing from adjacent genetically related layers in physical, chemical, and biological properties or characteristics such as color, structure, texture, consistency, kinds and number of organisms present, degree of acidity or alkalinity, etc.*"

Name	**Definition**	**Data Type**	**Multiplicity**
soil	The soil that contains the horizon.	SO_Soil	One
designation	Reference to a classification of a horizon according to a formal system.	SO_HorizonDesignation	Zero to many

9.2.1.4 SO_Component

A data type that describes the particle size distribution, and physical and chemical properties of the various soil size fractions (such as 'fine soil material', 'whole soil', 'coarse fragments'), either within a soil horizon or an undifferentiated soil. These properties were taken from the GlobalSoilMap consortium's requirements [4].

Name	**Definition**	**Data Type**	**Multiplicity**
texture	Classification of soil texture according to a defined scheme (e.g. 'sand' or 'clay loam').	Concept	Zero to many
bulkDensity	Bulk density in mass per unit volume of the described size fraction.	Observation	Zero to many
particleSizeFraction	The size fraction of the soil being described (e.g. 'fine soil material', 'whole soil', ''coarse fragments').	Concept	One
pH	The acidity or alkalinity of the described size fraction.	Observation	Zero to many
electricalConductivity	The ability of a soil to transmit an electric current.	Observation	Zero to many
organicCarbon	Carbon present as organic matter in the described size fraction.	Observation	Zero to many
effectiveCationExchange Capacity	The total capacity of the described size fraction to hold exchangeable cations.	Observation	Zero to many
particleSizeDistribution	The proportion of defined particle sizes (e.g. 'sand', 'silt' and 'clay') within the described size fraction.	SO_ParticleSize Value	Zero to many

9.2.1.5 SO_ParticleSizeValue

A data type to describe the proportions of particle size classes in the soil. As the definition of a particle size may vary according to the classification system, this class explicitly states the size range.

Name	Definition	Data Type	Multiplicity
particleSize	The lower and upper limits of the particle size appropriate for the classification system used to provide term used in label.	MeasureRange	One
label	A label to describe the particle size (e.g. 'sand', 'fine sand').	Concept	Zero or one
proportion	The proportion that the soil contains of this particle size.	MeasureRange	One

9.2.1.6 SO_SoilFeatureRelationship (proposed)

A SO_SoilFeatureRelationship class was proposed but not tested. This captures relationships between soils and/or horizons. These include spatial and temporal associations, or aggregation of soils into collections like map complexes. Participants in relationships play roles (e.g. ‘overlies’, ‘truncates’, ‘member’). The design is based on the GeoSciML GeologicFeatureRelationship feature type and can be extended to provide additional information, for example, a boundary relationship between two horizons could capture the distinctness of the boundary.

Name	Definition	Data Type	Multiplicity
type	The type of relationship	Concept	One
sourceRole	The role played by the source soil feature or object.	Concept	Zero or one
targetRole	The role played by the target soil feature or object.	Concept	One

9.2.1.7 SO_SoilTaxonomy

A system to classify soils according to diagnostic criteria based on soil morphology, genesis, function or other characteristics. Examples include the IUSS World Reference Base [8], the USDA Soil Taxonomy [13] or the Australian Soil Classification [5].

Name	Definition	Data Type	Multiplicity
label	Name of the soil classification system.	String	One
member	Reference to the member soil taxa.	SO_SoilTaxon	One to many

9.2.1.8 SO_SoilTaxon

A single class in a soil taxonomy.

Name	Definition	Data Type	Multiplicity
label	Name of the taxon (e.g. 'Allophanic Soil').	String	One
scheme	Reference to the soil taxonomy that defined the taxon.	SO_SoilTaxonomy	One

9.2.1.9 SO_HorizonDesignation

A term to categorize the soil horizon according to certain physical criteria. Horizons may be defined as genetic, morphological, functional or diagnostic horizons (to support differentiation of soils into soil taxa).

Name	Definition	Data Type	Multiplicity
label	A soil horizon designation.	String	One
scheme	The scheme from which the designation value is taken.	To be determined	One

9.2.2 Soil Sampling and Observations

The sampling and observation model (Figure 6) has been designed to make extensive use of the OGC Observations and Measurements model [ISO 19156:2011]. The sampledFeature or featureOfInterest will be a SO_Soil or SO_Horizon as defined in this model. Domain specific classes have been added for informative purposes to show how concepts describing the sampling strategy, many of them not unique to the soil domain, map on to the O&M model. The O&M abstract specification provides a full description of the model.

9.2.2.1 SamplingFeature

O&M defines a sampling feature (SamplingFeature) as a '*feature which is involved in making observations concerning a domain feature. EXAMPLE Station, transect, section or specimen.*' They may be spatially constrained (SpatialSamplingFeature) when '*observations are made to estimate properties of a geospatial feature, in particular where the value of a property varies within the scope of the feature...*'. Specimens are '*a physical sample, obtained for observation(s) normally carried out ex-situ, sometimes in a laboratory.*'

Four informative sub-types of a SpatialSamplingFeature were defined to show how it encompasses common artefacts of soil sampling strategies: the soil survey (SO_Survey), site (SO_Site), plot (SO_Plot) and a simple marker locating a sampling event (SO_Station). An additional subtype (SO_Layer) was added for circumstances where an arbitrary layer of the soil with some lateral extent is sampled. An explicit distinction is made between these sampling layers and a layer within the soil, which is defined by some environmental process, or other character of the soil, but is not a soil horizon. This later concept has not been modelled but could be defined as per the MaterialLayer class in the ISO/IUSS 'Wageningen Proposal'.

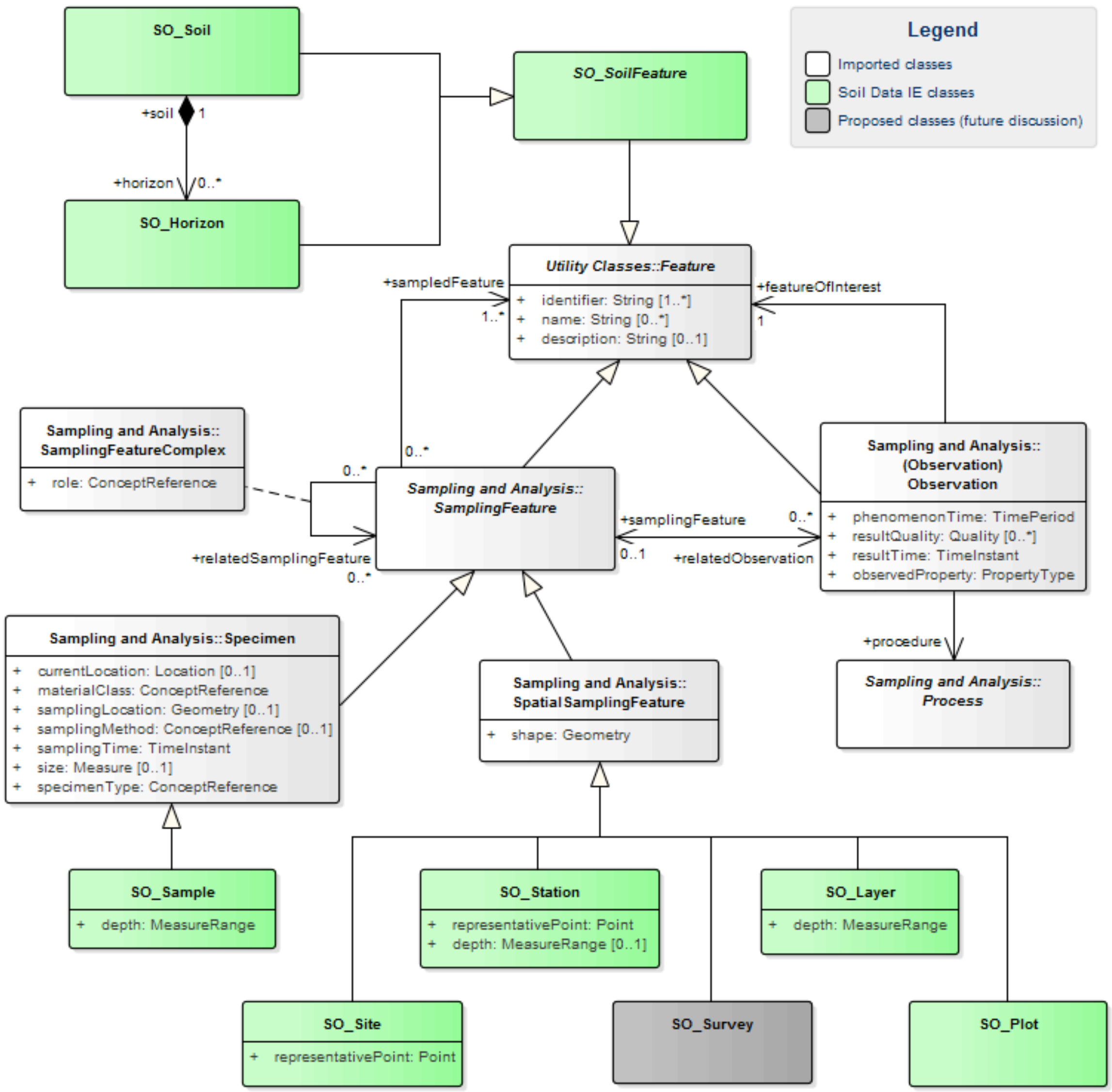


Figure 6 — Soil sampling and observation class model

The SO_Sample class highlights that the common use of the term ‘sample’ within the soil community is equivalent to a Specimen in ISO19156:2011.

Two informative properties were defined for soil sampling features to reinforce their importance to a description: depth (the location of the sampling feature relative to the top of the soil); and representativePoint (a point marker used to locate a sampling feature, which itself may have a greater or more complicated extent – captured by the shape property). These properties are not unique to the soil domain.

9.2.2.2 Observation

O&M defines an Observation as '*an act associated with a discrete time instant or period through which a number, term or other symbol is assigned to a phenomenon. It involves application of a specified procedure, such as a sensor, instrument, algorithm or process chain. The procedure may be applied in-situ, remotely, or ex-situ with respect to the sampling location. The result of an observation is an estimate of the value of a property of some feature.*'

In the model the procedure value is an abstract class (Process) but data describing a documented methodology or algorithm can be provided by reference to terms (for example using SKOS Concepts) in a process register, or the Timeseries Profile of Observations and Measurements has defined a concrete ObservationProcess class that can be used to describe a piece of equipment.

9.2.3 Soil Profile

The soil profile (SO_SoilProfile, Figure 7) is a fundamental concept in soil science, describing the character of a soil as a vertical sequence of soil horizons. The soil profile is also commonly used as a proxy for the soil itself, as well as a method of sampling the soil. Separating these various uses of the term 'soil profile' proved contentious during the modeling work conducted in the IE.

Soil profiles are primarily used to describe a soil at a single location - this reflects the difficulty in describing a soil over larger extents due to the significant variation in the nature of a soil over small distances. Despite this, it is useful to provide descriptions of soil for a larger volume that predict expected values for a property in that space. These are presented as profiles and given names like 'soil class profiles' or 'derived profiles'. Soil map units can represent the occurrence of one or more of these interpreted profiles.

Creating separate classes to describe a soil based on the process (observation or interpretation) used to generate the description and the extent of the sampling is problematic. It can lead to convoluted modeling, where the same sets of properties are repeated across the 'observed' and 'derived' classes. Also, not all soil descriptions are profile-based, focusing on recording only a small subset of properties and often ignoring the internal structure of the soil. Using a profile to capture these simpler descriptions is misleading, overstating the intent of the descriptive activity. Finally, the hybrid sampling and description aspect of a profile doesn't reconcile well with Observations and Measurements, which makes a strong distinction between the sampling feature and the domain feature against which the observations are made.

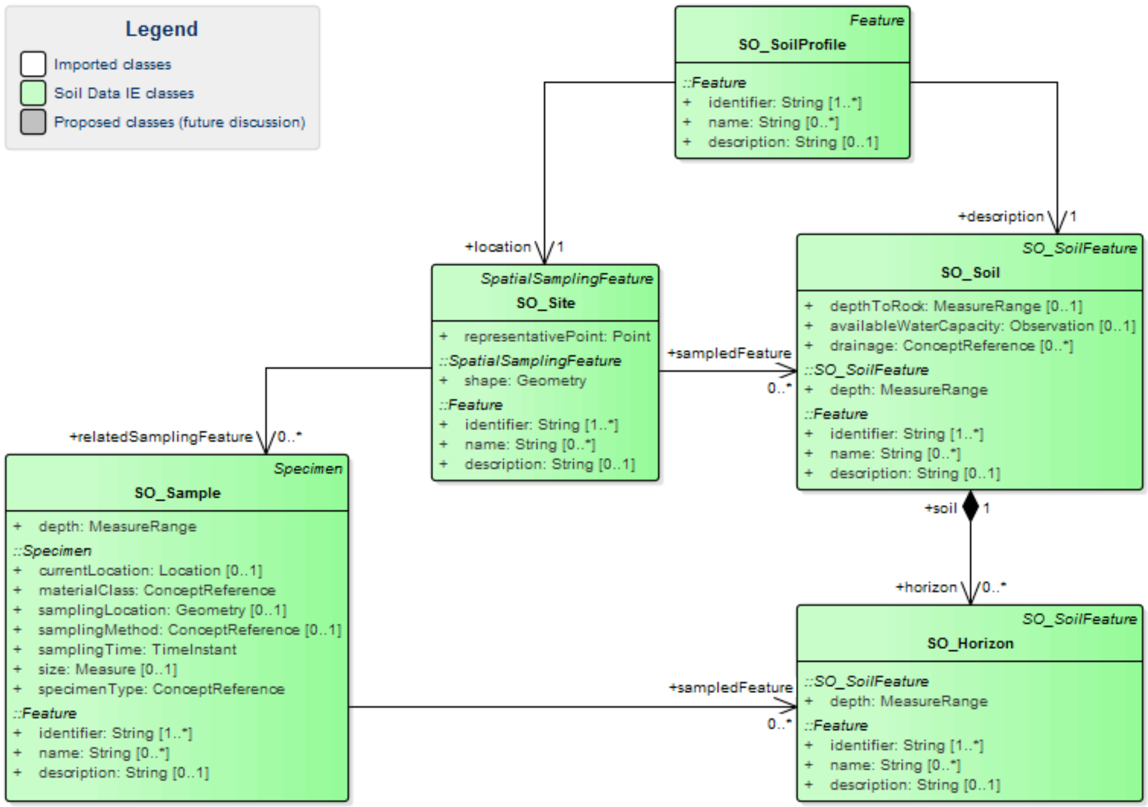


Figure 7 — Soil profile class model.

To align with O&M and simplify modeling, a single SO_Soil class was created that could hold descriptions for observed profiles, derived profiles, or simple soil descriptions. Sampling data are captured by the O&M aligned SamplingFeature classes. These can be aggregated into an identifiable profile using a SO_SoilProfile class. If required, different sub-classes of SO_SoilProfile could be defined (for example to explicitly model INSPIRE Observed and Derived Soil Profiles) with minimal impact on the rest of the model.

9.2.4 Soil Monitoring

This interoperability experiment tested the applicability of OGC 15-043r3 (the Timeseries Profile of Observations and Measurements) for delivery of time-varying data for dynamic soil properties. The conceptual model reflects this (Figure 8) and OGC15-043r3 fully describes the model.

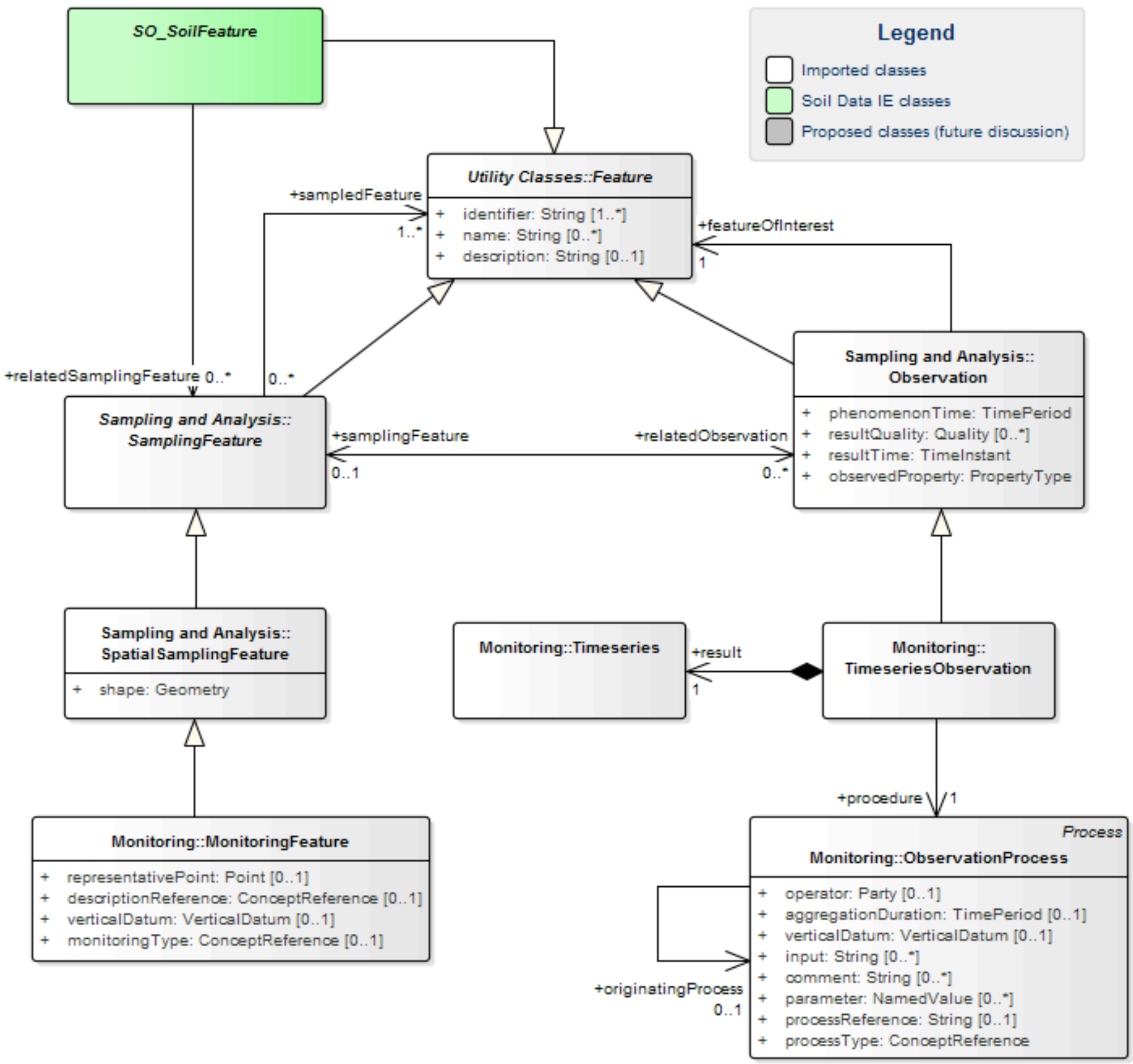


Figure 8 — Monitoring data class model

The Timeseries profile extends O&M to provide a cross-domain model for the delivery of measurements of environmental phenomena in a time series. Aside from providing classes that describe the sensors (ObservationProcess) and the platforms they are mounted on (MonitoringFeature), the model defines a TimeseriesObservation whose result is a set of time value pairs presented in a time-value-pair or domain-range form.

9.3 Logical Model

The logical model is a technology specific document that creates an ISO19109:2005 compliant UML application schema using the semantics of the Conceptual Model (Figure 9). The model consists of a single package containing classes created to describe a soil and its horizons. Classes that describe sampling and monitoring activities are imported, without extension, from O&M (Figure 10) and the Timeseries Profile of O&M (Figure 11).

There is a one-to-one relationship between the scope and semantics of the conceptual and the logical model classes created to describe the soil domain. As such, definitions are not repeated from the Conceptual Model section.

The only significant change to the conceptual model is the removal of soil sampling classes and their replacement with the appropriate O&M super-class (Table 3).

The Logical Model also makes references to Concept and ConceptScheme classes defined by the W3C Simple Knowledge Organisation System (SKOS) for controlled vocabularies of terms (all properties with the data type ScopedName) and soil classification (SO_SoilTaxonomy, SO_SoilTaxon and SO_HorizonDesignation).

Table 3 — Mapping of conceptual soil sampling classes to O&M Sampling Features

Conceptual Model	Observations and Measurements
SO_Sample	SF_Specimen
SO_Site	SF_SpatialSamplingFeature
SO_Station	SF_SpatialSamplingFeature
SO_Survey	SF_SpatialSamplingFeature
SO_Layer	SF_SpatialSamplingFeature
SO_Plot	SF_SpatialSamplingFeature

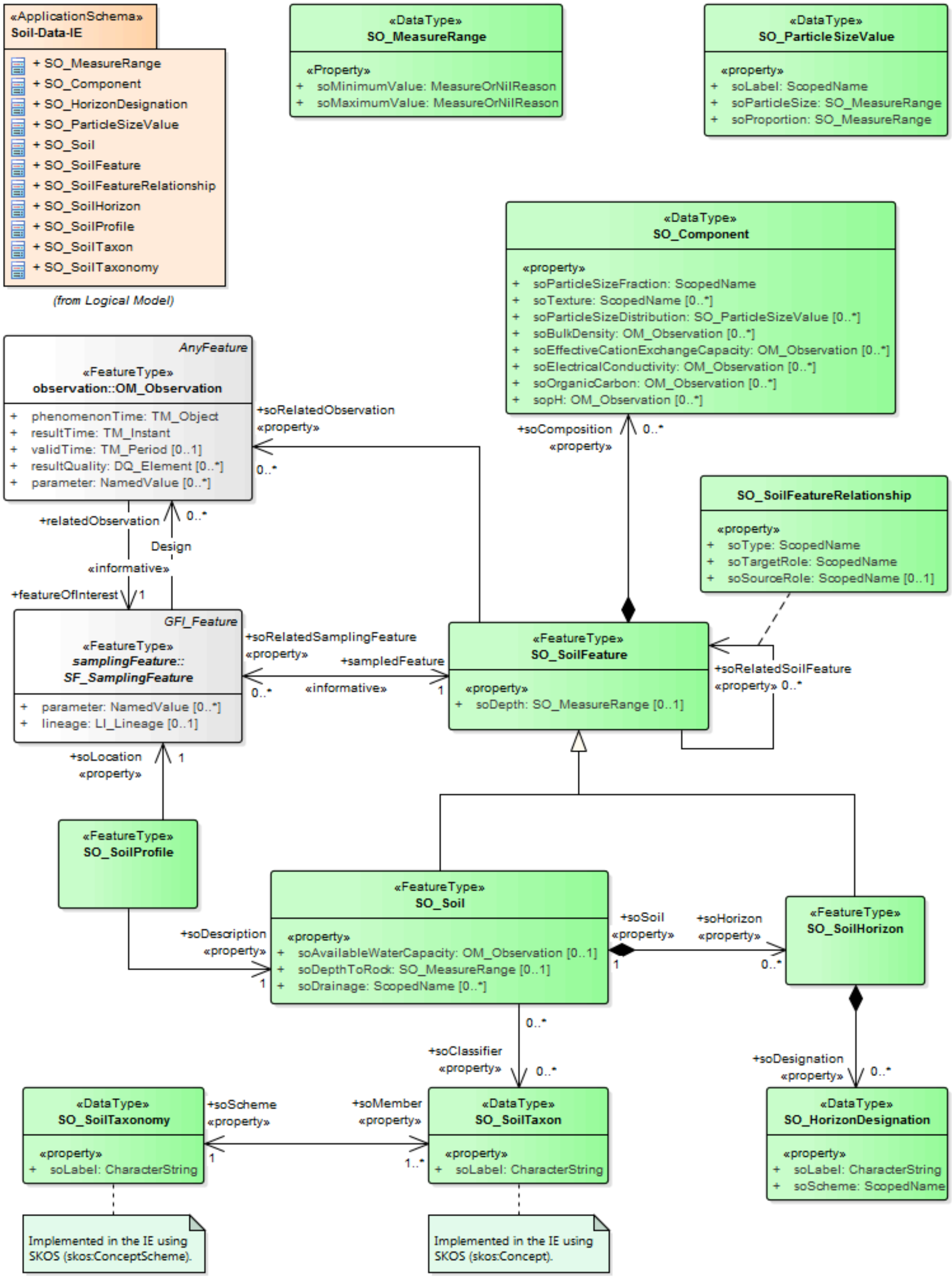


Figure 9 — The Soil Data Interoperability Experiment Logical Model

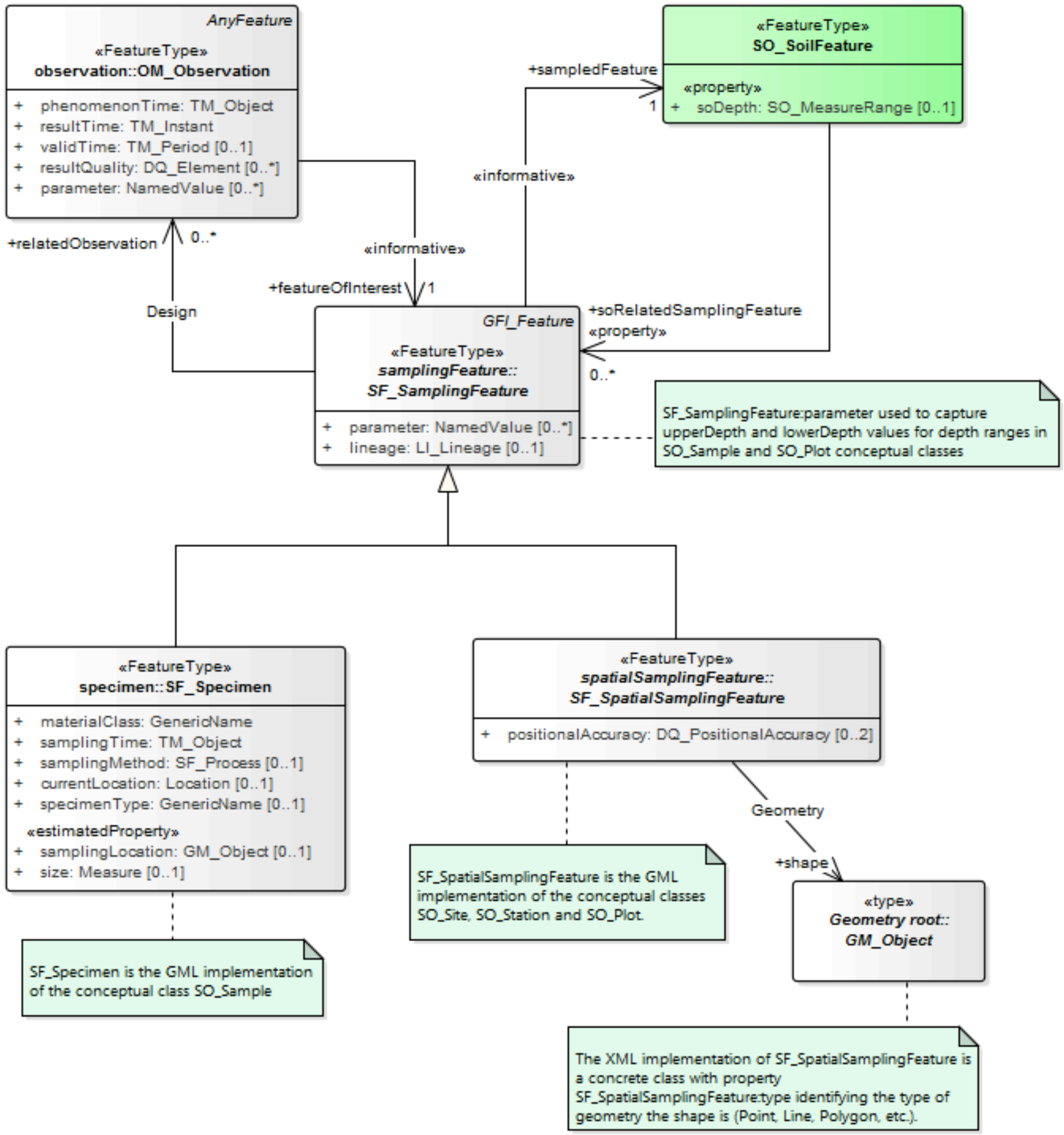


Figure 10 — Logical Model classes from Observations and Measurements

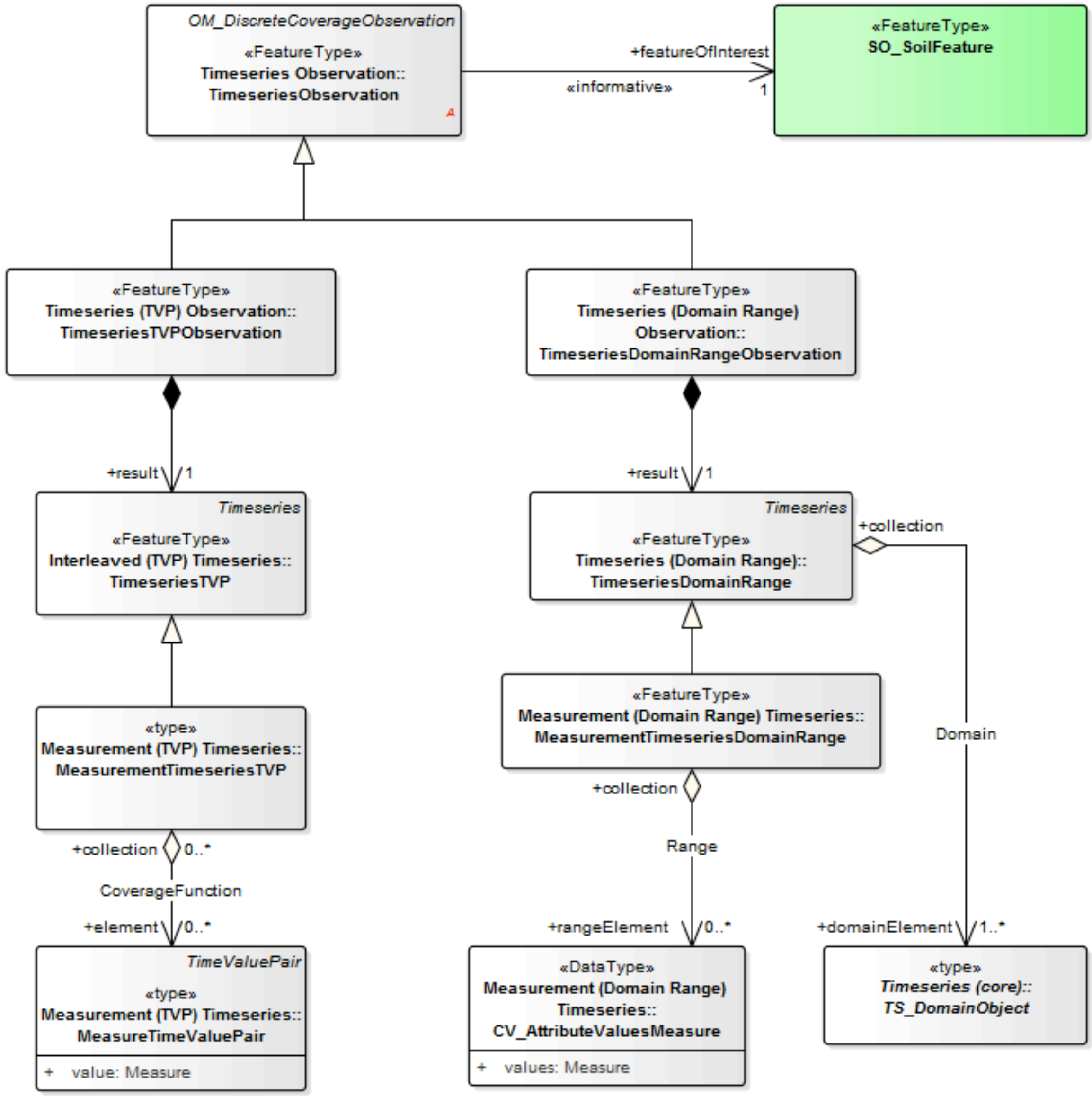


Figure 11 — Logical Model classes from the Timeseries Profile of Observations and Measurements

9.4 Physical Model

The physical model (Figure 12) defines how soil description features and their properties shall be encoded using GML according to the rules specified in ISO19136:2007. The resulting schema has been provisionally titled SoilIEML and the XML schema document is published at:

https://github.com/opengeospatial/SoilDataIE/schema/soil-data-ie.xsd

The XSD was generated manually from the logical model and matches its semantics exactly.

Table 4 lists the namespaces declared or imported by the schema.

The SKOS information model for controlled content cannot be incorporated into the GML schema. References to SKOS Concepts are implemented as HTTP URIs held in XLINK attribute values.

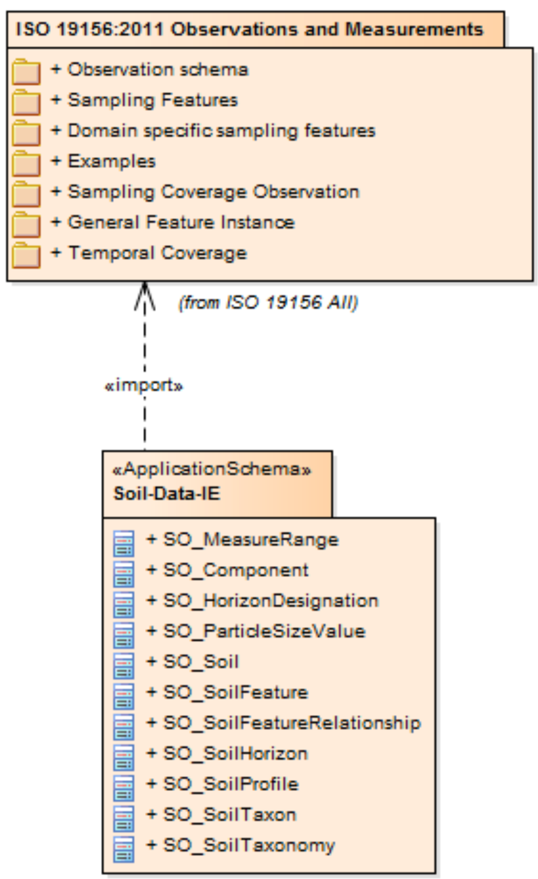


Figure 12 — SoilIEML package diagram

Table 4 — Soil Data IE Physical Model namespaces

Prefix	Namespace	Type
soil	http://www.opengis.net/soil-data-ie/0.1	Target
gml	http://www.opengis.net/gml/3.2	Import
om	http://www.opengis.net/om/2.0	Import
sam	http://www.opengis.net/sampling/2.0	Import
sams	http://www.opengis.net/samplingSpatial/2.0	Import
spec	http://www.opengis.net/samplingSpecimen/2.0	Import

10 Soil Data Interoperability Experiment Architecture

The Soil Data IE architecture (Figure 13) defines the components used to implement a prototype distributed system that exchanged data according to the Soil Data IE information model. Three suites of components were deployed.

— Web services conforming to OGC specifications (WFS, WCS, WPS and SOS) for the delivery and processing of site registration, and soil formation, description, sampling, observation data.

— Web services conforming to World Wide Web Consortium (W3C) Linked Data Platform for the resolution feature URIs and the delivery of soil classification and vocabulary definitions.

— Bespoke tools to present or transform data.

The design is a service oriented architecture where components support the discovery of features and the construction of representations of the features. Complimenting this is a lightweight resource oriented architecture where features are given URIs to uniquely identify them and provide access to digital representations. This allowed the configuration of features as relatively lightweight objects that used URI links to identify and resolve to related features. The client can access related data as it needs, rather than deal with a single large object with all related data returned inline. It also allowed the use of SKOS to provide data for terms and their definitions. This was useful but introduces an inconsistency into the system, as the nature of the data encoding patterns and protocols being used changed to those of the Semantic Web.

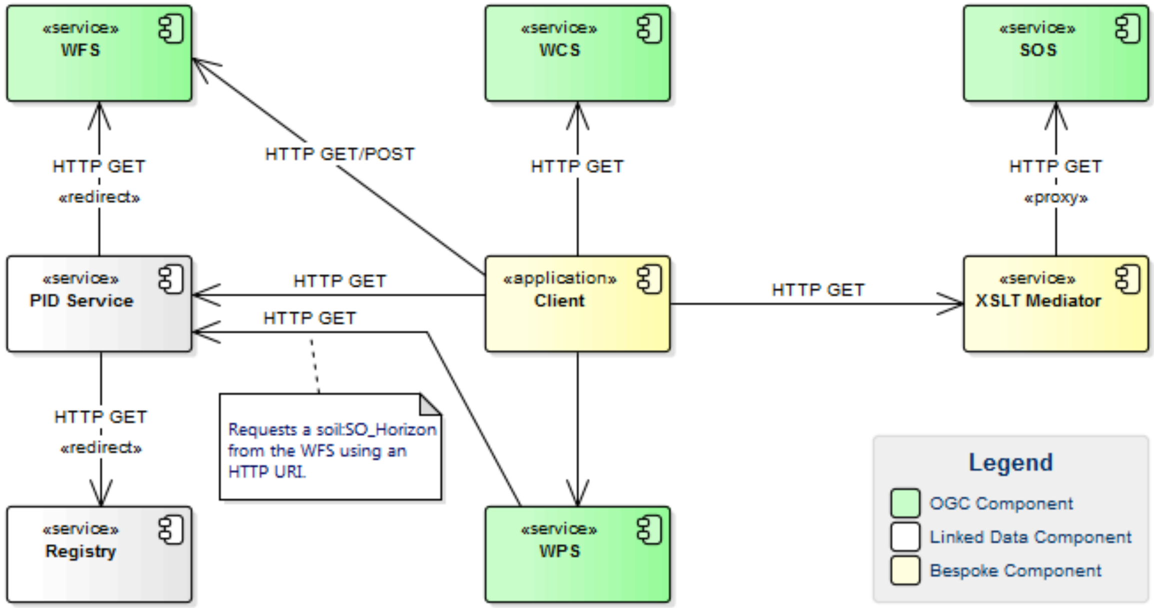


Figure 13 — Overview of the soil data IE component architecture.

The direction of associations is from the component that initiates the interaction between components.

10.1 Components

10.1.1 OGC Web Services

Four OGC-compliant service interfaces were used to deliver the soil data of interest: Sensor Observation Service (SOS); Web Coverage Service (WCS); Web Feature Service (WFS); and Web Processing Service (WPS). The Use Case 1 WFS instances were new and were bound directly to the databases managed by participants. In contrast, the SOS instances used for Use Case 2 were 'virtual services' that used a simple XSLT-based mediator to translate data from existing SOS instances into TimeseriesML. The experiment was able to make use of existing WCS instances that were providing the soil property predictions covered by Use Case 3. Results from WFS queries were forwarded to WPS to generate additional soil properties.

Sensor Observation Service (SOS)	
Version	1.0 [06-009r6] and 2.0 [OGC 12-006].
Implementation	ncSOS-THREDDS Server (USGS); Hilltop Server SOS (Horizons Regional Council; Landcare Research)
Use Case	2
Role	Provision dynamic soil property (soil moisture and temperature) observations.
Comments	SOS instances provided by participants conformed to different versions and provided different content (O&M 1.0, THREDDS, and WaterML 2.0, Hilltop). They were therefore proxied by a bespoke mediator service (see XSLT Mediator below) that translated their responses into TimeseriesML 1.0.

Web Coverage Service (WCS)	
Version	2.0 [OGC 09-110r4]
Implementation	Geoserver (CSIRO, ISRIC)
Use Case	3
Role	Provision of coverages describing variation in soil properties.

Web Feature Service (WFS)	
Version	2.0 [OGC 09-025r1]
Implementation	Geoserver with app-schema plugin (CSIRO, ISRIC); Snowflake Go Publisher WFS (Landcare Research)
Use Cases	1 and 4
Role	Provision of feature describing site registration, and soil formation, description, sampling and observations.

Web Processing Service (WPS)	
Version	1.0.0 [OGC 05-007r7]
Implementation	PyWPS 4.0 (ISRIC)
Use Case	4
Role	Provision of processes encapsulating pedo-transfer functions to calculate soil wilting point, field capacity and bulk density for soil horizon data provided by Use Case 1 WFS instances.

10.1.2 Linked Data Services

Every identifiable feature, classifier and term used by the IE was given an HTTP URI as an identifier and means of linking data from different data types or sources. A very simple interpretation of the Linked Data Platform principles was used to design these URIs and the Linked Data web services were used to dereference them (PID Service) and also provide definitions of properties and terms (Linked Data Registry).

Persistent Identifier Service (PID Service)	
Version	1.1.137 [Auscope PID]
Implementation	AuScope SISS PIDService (Landcare Research)
Use Cases	1, 2
Role	Manages the resolution of URIs identifying GML Features (soil and sampling data) and SKOS Concepts (term and property definitions). It uses REGEX to parse the URI and map it onto a request to an appropriate service, redirecting it to a WFS GetFeature request or an item in the Linked Data Registry.

Linked Data Registry	
Version	1.1.0 [UKGovLD]
Implementation	Epimorphics Linked Data Registry (Core) (CSIRO)
Use Cases	1, 2
Role	Manages the SKOS Concepts used to capture definition of terms and property types used by the experiment.

10.1.3 Bespoke Services and Applications

Several bespoke applications were written to demonstrate the implementation of the IE services; this includes a mediator servlet that used XSLT to create appropriate GML when source services could not. The details of demonstration client applications are included in subsequent sections.

XSLT Mediator	
Version	1.0
Implementation	Bespoke java servlet (Landcare Research)
Use Cases	2
Role	A proxy service used to translate O&M 1.0 and WaterML 2.0 from participating sensor services into TimeseriesML 1.0.

11 Soil Data Interoperability Experiment Demonstration

11.1 Demonstration 1 – Use cases 1 and 4

This demonstration deployed clients and services that implemented both Use Case 1 (Soil Data Integration) and, because the necessary data processing used data provided for Use Case 1, Use Case 4 (Pedo-transfer Functions). Table 5 summarizes the participants and their contribution and Figure 14 shows the components of the IE architecture that were used.

Participants deployed:

- Web Feature Services that provided soil description and sampling and lab analysis data for soil sites;
- a Web Processing Service that took soil horizon descriptions (served by WFS) and generated predictions of soil wilting point, field capacity and bulk density for each horizon; and
- a Register of soil vocabularies and property definitions.

A Persistent Identifier Service managed the resolution of the HTTP URIs allocated to the resources (GML Features and SKOS Concepts) provided by these services.

Two clients were created to demonstrate the use case. An overview of the user interaction, and the services and feature types used to facilitate this, is included in Annex D.

Landcare Research deployed a simple client to demonstrate how linked data principles could be used to access and explore representations of soil description and sampling data at a site (Figure 15). It also included links into the soil definition register to show how a client could be augmented with definitions of terms through calls to SKOS Concepts.

Federation University Australia deployed a more comprehensive client to integrate the complete set of Use Case 1 soil description and observation data from CSIRO, ISRIC and Landcare Research (Figure 16). The application allowed the user to explore the data and provided visualizations of soil particle size distribution data. It also incorporated pedo-transfer functions published by a WPS at ISRIC, allowing the user, at the press of a button, to augment the local data with predictions of soil wilting point, field capacity and bulk density.

These clients were used to demonstrate Use Cases 1 and 2 at the OGC Technical Committee meeting at Sydney in December, 2015. A video of the demonstration is available via the OGC YouTube channel:

https://www.youtube.com/watch?v=oR-c7Viu19k.

Table 5 — Demonstration 1 participants and components

Participant	**Country**	**Component**	**Implementation**
CSIRO	AU	Registry WFS	Epimorphics Linked Data Registry Geoserver (with app-schema plugin)
Federation University Australia	AU	Client	Bespoke (Open Layers 3 and Bootstrap)
ISRIC	NL	WFS WPS	Geoserver (with app-schema plugin) pyWPS 4.0
Landcare Research	NZ	Client PID Service WFS	Bespoke (.NET) CSIRO/AuScope PID Service Snowflake Go Publisher WFS

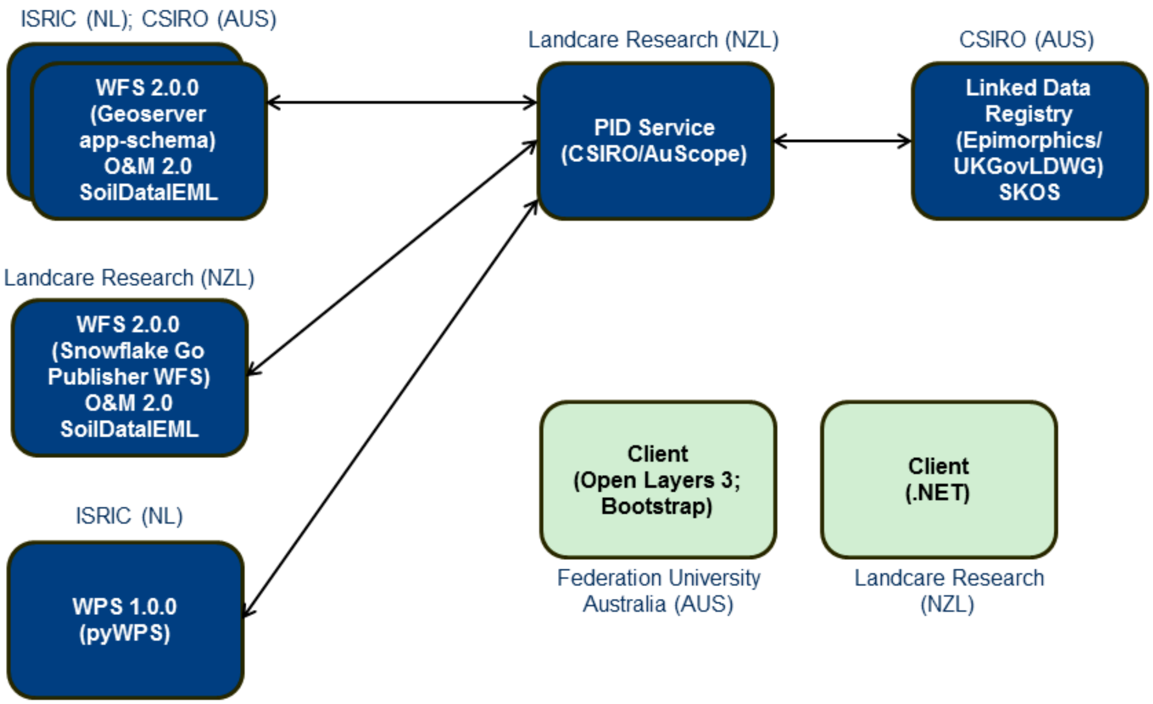


Figure 14 — Overview of Demonstration 1 components.

At any point a client (green) may directly access any one of the data services (blue).

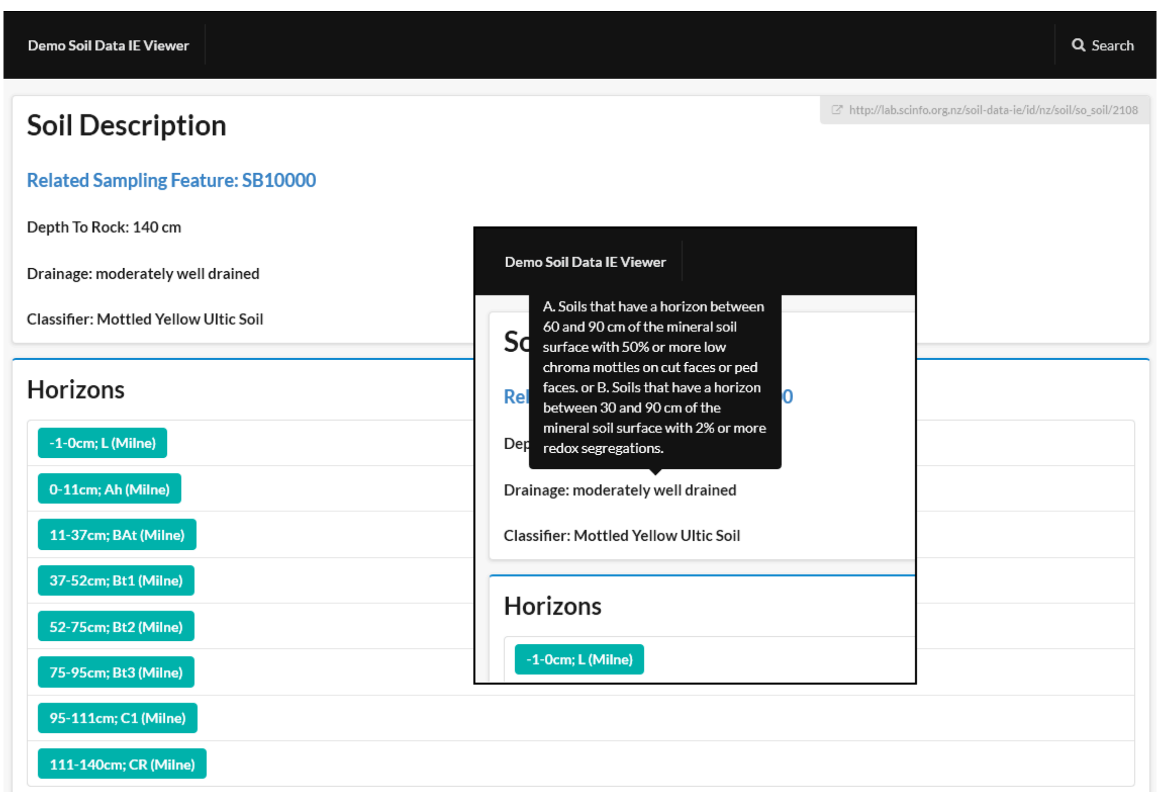
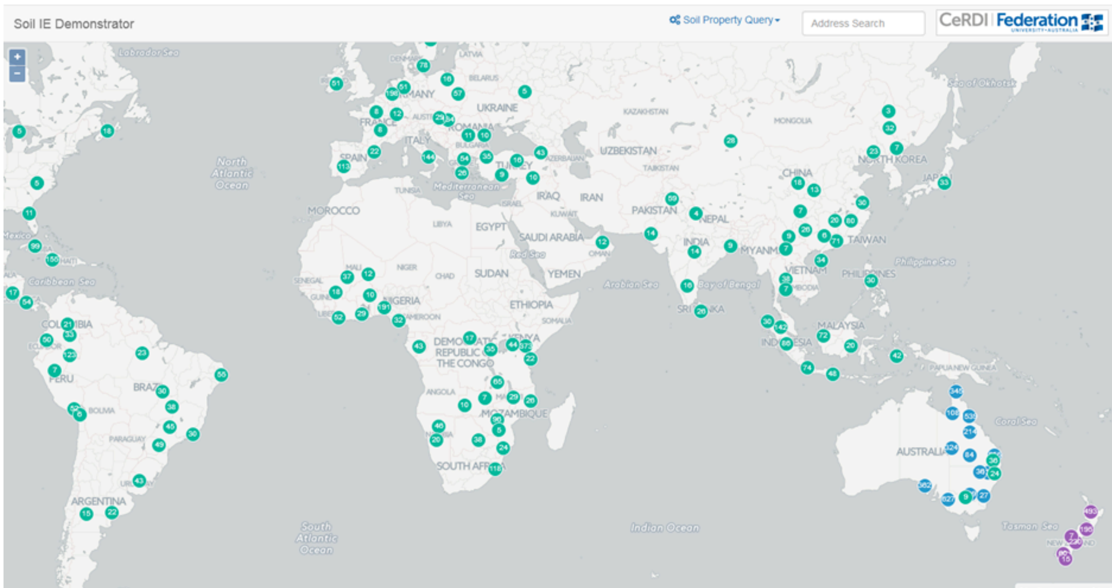


Figure 15 — The Landcare Research Demonstration 1 client.

Users explore the soil description through links to related features (for example, the list of horizons) or use tool tips to access definitions of terms (inset).

a) 

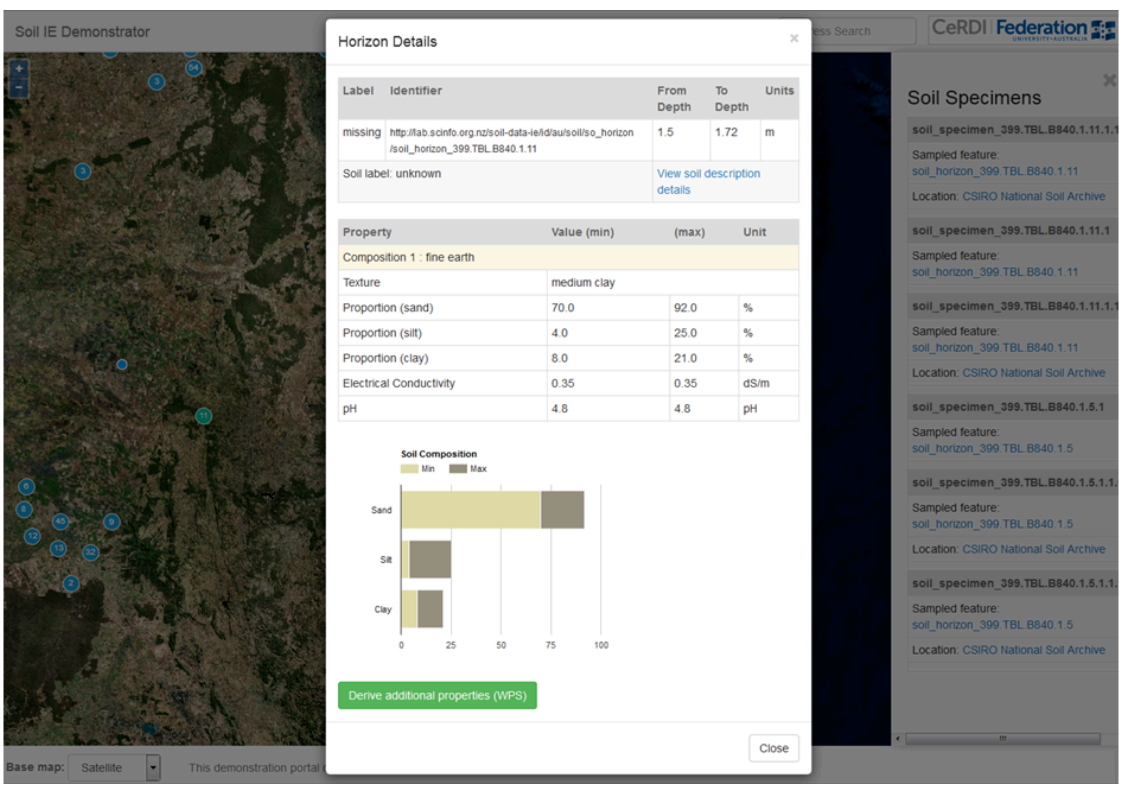
b) 

Figure 16 — The Federation University Australia Demonstration 1 client.

Shows (a) the soil sites provided by CSIRO (blue), ISRIC (green) and Landcare Research (magenta); and (b) horizon details, including a visualization of sand, silt and clay distribution in the fine earth fraction of the soil. Users could add predictions of soil wilting point, field capacity and bulk density to the property list by selecting the green ‘Derive additional properties (WPS)’ button.

11.2 Demonstration 2 – Use case 2

This demonstration deployed a client and services that implemented Use Case 2 (Soil Sensor Data). Table 6 summarizes the participants and their contribution and Figure 17 shows the components of the Soil IE architecture that were used.

Participants deployed:

— existing Sensor Observation Services that provided soil moisture monitoring data;

— an XSLT Mediator service to transform the SOS service data into TimeseriesML; and

— a Register of soil property definitions.

As with Demonstration 1, a Persistent Identifier Service was used to manage resource URIs.

Landcare Research deployed a simple client to plot the temporal variation in soil moisture state and rainfall (Figure 18). An overview of user interaction, and the services and feature types used to facilitate this, is included in Annex D.3. Data were provided by Horizons Regional Council (catchment soil hydrology monitoring services), Landcare Research (precision irrigation monitoring at a research farm) and the USGS (region flood hazard monitoring).

As it was not a formal standard at the time of the IE, no TimeseriesML SOS services were available for testing. Instead, each participant made available SOS endpoints conforming to SOS 1.0 (O&M 1.0) or SOS 2.0 (WaterML 2.0) and a mediator service was deployed to broker access to the services and transform their responses into TimeseriesML.

Table 6 — Demonstration 2 participants and components

Participant	Country	Component	Implementation
Horizons Regional Council	NZ	SOS	Hilltop Server
Landcare Research	NZ	Client PID Service SOS	Bespoke (R/Shiny) CSIRO/AuScope PID Service Hilltop Server
USGS	US	SOS	NcSOS-THREDDS

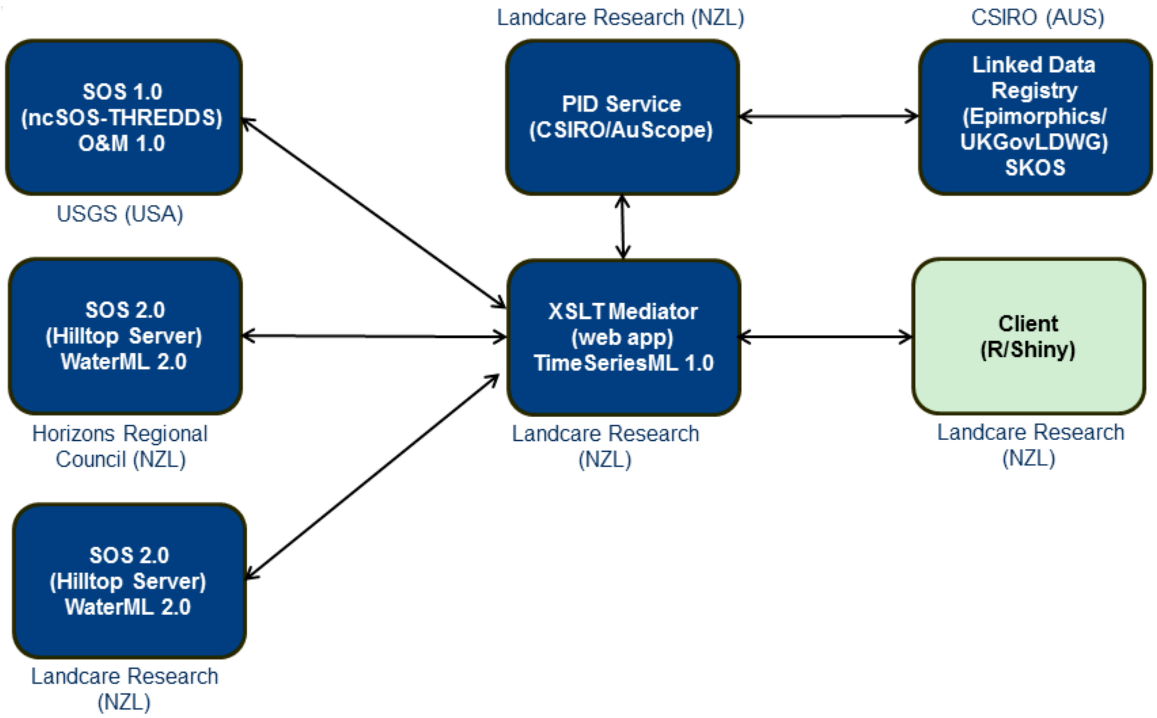


Figure 17 — Overview of Demonstration 2 components.

As the source data services were deployed to differing specifications an XSLT mediator was used to harmonize data for the client (green).

Figure 18 — Landcare Research Demonstration 2 client.

Shows soil moisture (top graph) and rainfall (bottom graph) timeseries data from a monitoring station (red circle) in the Manawatu Region of New Zealand. Data provided by Horizons Regional Council.

11.3 Demonstration 3 – Use case 3

This demonstration used an existing Web Coverage Service to implement Use Case 3 (Soil Property Modeling and Predictions). Table 7 summarizes the participants and their contribution.

As part of the consortium of researchers working on the Soil and Landscape Grid of Australia (http://www.clw.csiro.au/aclep/soilandlandscapegrid), CSIRO has deployed a Web Coverage Service that includes coverages predicting the national variation of a set of soil properties. Meanwhile ISRIC has deployed a Web Coverage service predicting the global distribution of an equivalent set of properties. In both cases, the predictions were consistent with the specifications of the GlobalSoilMap project and therefore covered soil properties that had been prioritized for the Soil IE. Table 8 summarizes the coverages provided and the Soil Data IE UML properties to which they correspond.

The soil properties are derived properties in that they are conditional on other properties. For example, the clay prediction in the soParticleSizeDistribution value where the soParticleSize is <2µm and the SO_Component describes the fine earth fraction. Depending on the particle size classification scheme used by the data provider, the source data may be processed to match these criteria for the predictions.

With the exception of ‘Depth of Soil’, coverages of predictions were provided for six depth intervals (0-5cm, 5-15cm, 15-30cm, 30-60cm, 60-100cm and 100-200cm) for each Soil Property. In turn, each prediction is accompanied by two coverages representing the lower 5% and upper 95% confidence intervals for the prediction.

The coverages can be viewed as layers in the Federation University Australia client deployed for Demonstration 1, or, juxtaposed with site data from SoilIEML WFS, in QGIS (Figure 19).

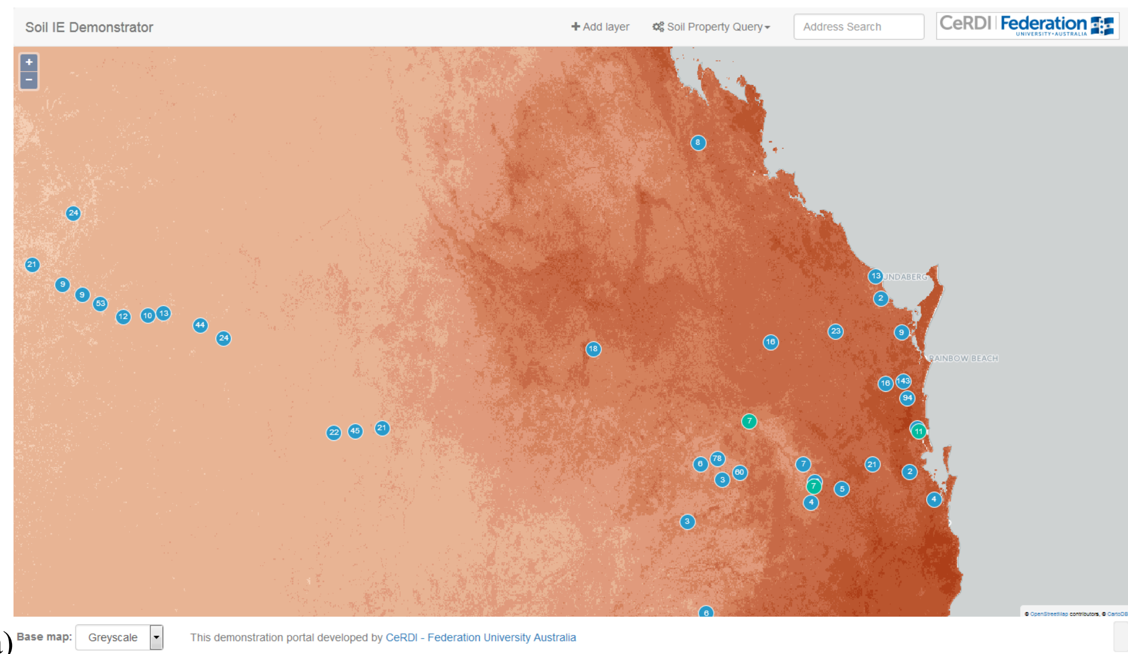
Table 7 — Demonstration 3 participants and components

Participant	Country	Component	Implementation
CSIRO	AU	WCS/WMS	ArcGIS Server
ISRIC	NL	WCS/WMS	Geoserver

Table 8 — Demonstration 3 coverages and their corresponding SoilIEML properties.

GlobalSoilMap Property	Soil Data IE Property
Bulk Density (whole earth)	soBulkDensity (where soParticleSizeFraction = ‘whole soil’)
Organic Carbon	soOrganicCarbon (where soParticleSizeFraction = ‘fine earth’)
Clay	soParticleSizeDistribution (where soParticleSize < 2µm and soParticleSizeFraction = ‘fine earth’)
Silt	soParticleSizeDistribution (where soParticleSize = 2-20µm and soParticleSizeFraction = ‘fine earth’)

Sand	soParticleSizeDistribution (where soParticleSize = 20µm-2mm and soParticleSizeFraction = 'fine earth')
pH (CaCl2)	pH (where om:procedure is the CaCl2 method and soParticleSizeFraction = 'fine earth')
Available Water Capacity	soAvailableWaterCapacity
Effective Cation Exchange Capacity	soEffectiveCationExchangeCapacity (where soParticleSizeFraction = 'fine earth')
Depth of Soil	soDepthToRock



a)

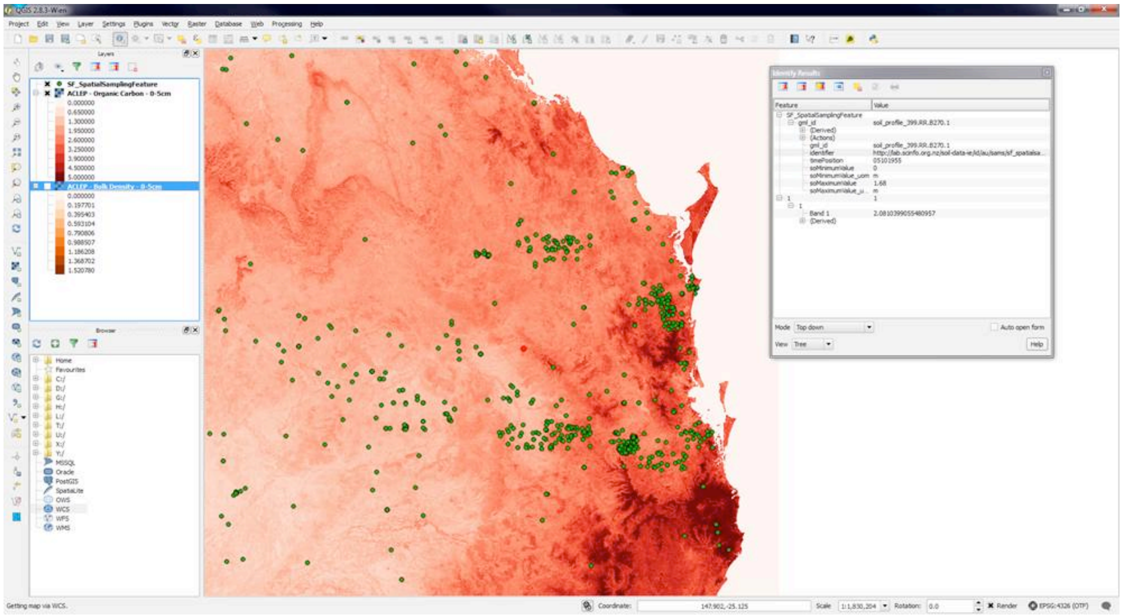
b) 

Figure 19 —Demonstration 3 clients.

Shows GlobalSoilMap Organic Carbon predictions (a) as WMS layers in the Demonstration 1 client and (b) WCS coverages juxtaposed with CSIRO soil sampling sites in QGIS.

12 Results

This Soil IE successfully met the stated objectives. It developed a core soil information model, implemented standardized services against the model from a number of different sources, demonstrated the ability to satisfy the defined use cases, and delivered this IE engineering report and a set of recommendations to further progress this work.

12.1 Information modeling

The following were discovered during the modeling process.

1. Soil is a complex entity and therefore is often difficult and time consuming to conceptualize. This is reflected in the existing candidate models: they are either relatively simple, and defer a lot of the work to implementers, or extensive, but complex and difficult to comprehend. Both approaches have their limitations.

2. Simple models rely on extension through soft-typing and risk proliferation of extensions that may not interoperate.

3. Explicit, complex models can be difficult to implement and result in complex and cumbersome physical implementations.

4. When defining a soil data exchange standard, we need to strive for a continuum from the conceptual to physical models where simplicity and complexity are used carefully, but not uniformly. For example, by defining an explicit conceptual model to guide and constrain implementations but using a more flexible physical encoding utilizing soft/virtual typing. Regardless, the models need to be defined in modules and developed in increments. Also, for a given application, not every aspect of a soil may be described, instead the end user will compile a view of the soil using only those modules that are needed.

5. The new model developed in the IE is simple, but robust, defining a soil as per the profile model (a collection of horizons) but reconciling it with the existing OGC standards baseline by decoupling the profile as a combined soil and soil sampling artefact. However, while the current model for soil composition (SO_Component) works well for a description of soil properties according to the GlobalSoilMap project, it is unlikely to be extensible. A more robust means of describing the material composition of the soil, and the interaction of those components should be defined.

6. The use of OM_Observation as a datatype was very useful for annotating property values with useful provenance and quality metadata, but expensive in terms of volume of XML generated; not all OM_Observation properties are needed in this context. An ability to switch between a terse view of an observation (when inline) and a comprehensive view, when provided as a feature would be valuable.

12.2 Implementation

The following were discovered during implementation.

1. The IE did not have sufficient resources to develop the model further, especially with the limited time for face-to-face meetings and development work.

2. Most participants were inexperienced information modelers and this further hampered the modeling exercise. Future agreement on a standard soil data model will require increased access to experienced information modelers with capacity to mentor other participants and soil domain specialists.

3. The use of Linked Data principles for identification and linking of features was helpful, simplifying much of the interaction with services on RESTful lines; these were easily implemented using the PID service.

4. Participants reported difficulty deploying services. Although the physical model is relatively simple, implementation was complicated by:

    a. Limited experience with web service applications and object relational mapping techniques;
    b. The complexity of mapping from the source database tables to the target schema; and

    c. The reliability of the application used to provide the service.

5. Simple desktop clients for 'quick and dirty' testing of services are rare. WFS 2.0 and GML Simple Features Level 1 application schema support in GIS tools is poor. When restructuring the data to provide layer driven GML/GeoJSON formation information can be lost or compromised.

6. TimeseriesML can be used to deliver soil moisture and rainfall data and WaterML 2.0 can be translated to TimeseriesML 1.0 with minimal effort.

7. As a result of the IE we believe GML is missing some useful core datatypes:

    a. A Measure and Reference range data type (also a known requirement for geology and groundwater data) to consistently provide interval values; we expect most schema that work with interpreted/aggregated data would benefit from such a type; and
    b. A register of sampling feature types building on the informative O&M common sampling features [ISO 19156:2011, clause 10.4] would be useful; this should be a cross domain vocabulary as features like plots are not unique to soil, e.g. ecology.

8. There were conceptual difficulties implementing O&M observations. The 'featureOfInterest' as used in the O&M Abstract Specification is inconsistently applied in implementations, representing either the immediate sampling feature or the ultimate domain specific sampled feature.

9. The use of SKOS for data definitions is well established but they are not well integrated into the OGC suite of standards. Moving from GML to RDF/XML and OWS to SPARQL when working with vocabulary data is an irritant for software engineers. It would also be useful to filter across the GML/SKOS disconnect, for example a filter to find all features classified with a Concept that also discovers those classified with its narrower (transitive) concepts. Finally, there are inconsistencies between RDF implementations, for example comparing RDF from SISSVoc services and the Linked Data Registry, each uses properties from different namespaces for the same purpose. At the very least an OGC profile for SKOS/RDF should be defined.

10. The ability to use coverages to provide predictions of the variation in soil properties in space was expected and proven. In a useful distributed soil information system, data from soil observation and description services would be used to generate and test predictions. Further testing is required to investigate and demonstrate the utility of the IE services in this context.

12.3 Conclusions

This Interoperability Experiment has shown that the problem of soil data interoperability can be addressed and the definition of solutions has already been well advanced by existing initiatives. Overall, the information model defined by the Soil IE:

- Matches the conceptual needs of the GlobalSoilMap specification;
- Exceeds the coverage of the ISO SoilML, IS0/IUSS and eSoTerML schema;
- Is broadly equivalent to the INSPIRE Soil model;
- Falls well short of the ANZSoilML model; and
- Requires reworking to better allow class and property extensions.

Future work will need to reconcile the need for a comprehensive and unambiguous conceptual model and a flexible, lightweight implementation model.

Meanwhile, current technology allows the deployment of web services to support soil data exchange. The main impediment to progress is the absence of a standards community that is well enough resourced and mentored to develop the necessary technical and modeling skills required to develop a model and systems that describe a complex and nuanced natural resource. A resource that we must describe and understand in as much detail as possible to ensure we can support the Earth's human population and biosphere well into the future.

13 Recommendations

We make the following recommendations for related OGC standards:

1. Create feature or data type to support attributed property values, either by defining an XML implementation of OM_Observation where all properties except result are optional, or a new datatype that uses OM_Observation as the pattern for structure.

2. Add a MeasureRangeType to GML Basic Types.

3. Develop best practices for the use of SKOS in OGC services, perhaps including a GML encoding of SKOS (consistent with ISO 19150-2:2015, clauses 6.2.7 and 6.7.2 [6]) with a view to defining SKOS operators in the Filter Encoding Specification.

4. Change the O&M abstract specification to add 'samplingStrategy' to O&M Observation and restrict 'featureOfInterest' for the feature with the observed property (as per [9], Table 13).

14 Future work

We make the following recommendations for future work.

1. The OGC establishes a Standards Working Group to undertake the modular development of a set of abstract and implementation models for soil data in coordination with activities of the IUSS WGSIS and the FAO Global Soil Partnership (Pillars 5 and 4).

2. An initial task of the SWG would be to review the recommendations of this report and initiate change requests, or other activities, where appropriate.

3. Initiate a Soil Coverage Interoperability Experiment to test the integration of soil description, observation, monitoring data (plus other environmental covariates) for the development of soil property predictions to support farm and environmental management activities.

Annex A

Use Cases

A.1 Use Case 1 — Soil Data Integration

Participants	CSIRO (AU), Federation University (AU), ISRIC (NL), Landcare Research (NZ)
Synopsis	This use case involves using the soil schema as a canonical structure into which heterogeneous soil data (for sites, laboratory results and historical reports) formats are transformed and published. The focus is on data generated by an observation, or similar process, at a given location. For data that describe interpreted or modelled data see Use Case 3.
Scope	Site registration; soil sampling, observation, classification and mapping; vocabularies.
Benefits/Value	This use case provides a basic mechanism for exchange of soil data which has been captured for a variety of purposes and is managed in many disparate data systems. Publication of standardized data services will greatly benefit soil research, analysis and reporting mechanisms at individual, local, national and international levels.
Objective	The objective of this use case is to have a number of different soil data management organizations delivering a set of agreed soil data services compliant to an agreed data model. Many applications could then be conceived to access and utilize this standardized data for a variety of purposes, such as those above. This would also facilitate: — Serving soil data so it can be used for commercial purposes such as to inform farming decisions; and — Serving soil data so it can be used to determine temporal trends in soil health for farm and environmental management and reporting.
Actors	Human: — farmer-consultant — researcher/analyst — archive manager Machine: — portals — web services — modeling software
Actions	An 'actor' accesses multiple services to determine if soil sampling has: — occurred in a certain area (such as farm, paddock, region); — measure certain properties (such as texture, EC, pH, carbon content); — determined the soil type (soil classification); or — has been stored as an archived specimen.
Implementation	Web Feature Service

A.2 Use Case 2 — Soil Sensor Data

Participants	Horizons Regional Council (NZ), Landcare Research (NZ), USGS (US)
Synopsis	This use-case involves identifying the location of soil monitoring sensors and accessing dynamic soil properties at these sensor locations as well as related observations from these sensors to inform farmers, scientists, managers, and the public about local soil conditions. The test case is the provision of soil moisture data.
Scope	Site registration; soil monitoring.
Benefits/Value	Soil moisture timeseries data is of particular interest to drought coordinators for predicting and evaluating drought conditions, to flood forecasters to improve base conditions used for modeling flood conditions, emergency managers for controlling wildfires, and farmers to manage crop and pasture irrigation.
Objective	— Serve in-situ soil moisture timeseries data in concert with soil sensor metadata and soil parameters. — Serving data aggregated as part of a national monitoring network of in-situ sensors (e.g. soil moisture monitoring network). — To deliver a temporal feed of actual soil moisture data using of in-situ, site probe based data through related services.
Actors	Human: — drought modelers — emergency managers — water resource managers — flood forecasters — farmers Machine: — portals — web services — modeling software
Actions	An actor discovers soil sensors and accesses soil properties at the locations of those sensors: — at one or more depths/depth intervals beneath the surface; or — as a continuous vertical function for a soil profile or pedon. An actor accesses related timeseries observations at sensors
Implementation	Sensor Observation Service Web Feature Service

A.3 Use Case 3 — Soil Property Modeling and Predictions

Participants	CSIRO (AU), ISRIC (NL)
Synopsis	Provision high resolution estimates of functional soil properties across the globe. These estimates will be provided as coverages describing variation in a property over a horizontal extent (either at the earth's surface or subsurface intervals) or vertically (for example continuous variation within a soil profile). An example application is the provision of the data products generated by participants in the GlobalSoilMap project (a global map of regularly distributed soil property predictions over six depth intervals).
Scope	Soil description, classification and mapping.
Benefits/Value	Supports the development and implementation of policies for the management of the soil resources and to respond to natural resource emergencies (e.g. floods, droughts, wildfires).
Objective	Serving soil data: — so it can be used for commercial purposes such as to inform farming decisions; — for integration into farm management software, then delivering outputs for monitoring purposes e.g. phosphorus, management for maximum yields, subsequent impact; — for environmental monitoring (e.g. drought monitoring), soil moisture coverage would fit this; — aggregated as part of a national monitoring network of in situ sensors (e.g. soil moisture network); or — describing vertical variation in soil properties (e.g. profile slices for pH, EC and organic carbon).
Actors	Human: — farmer-consultant — fertilizer company — researcher — emergency manager — natural resource manager — flood forecaster — agronomists — catchment managers Machine: — portals — web services modeling software

Actions	An actor accesses multiple services (potentially provided to meet Use Cases 1, 2 and 4) to get input data for a process that will generate the coverages addressed by this Use Case. An actor may use a pedo-transfer function from Use Case 4 to generate the coverage. An actor accesses a service providing a coverage describing the variation in a soil property: — across the surface of the earth; — at one or more depths/depth intervals beneath the surface; and/or — as a continuous vertical function for a soil profile or pedon.
Implementation	Web Coverage Service

A.4 Use Case 4 — Pedo-transfer Functions

Participants	ISRIC (NL)
Synopsis	Soil pedo-transfer functions (PTF) are predicted functions of certain soil properties using data from soil surveys and normally used for the purpose of translating between soil properties, attributes and/or agronomic values. The use case will focus on running specific functions on soil data obtaining values that were not measured in the field, new agronomical significant values, soil classification parameters.
Scope	Soil description, classification.
Benefits/Value	Most soil data is collected for a specific purpose and while some has detailed analytical results, often only a few properties of interest are analyzed in laboratories. Pedo-transfer functions allow development of mathematical relationships between values of different properties, such that the properties of interest can be estimated from the results of those analyses that are available.
Objective	— Data conversion / harmonization — New derived soil attributes from field measured
Actors	Human: — researcher — agronomists Machine: — web services — modeling software
Actions	An actor accesses multiple services (potentially provided to meet Use Cases 1, 2 and 4) to get input data for a process that will generate the coverages addressed by this Use Case. An actor may use a pedo-transfer function from Use Case 4 to generate the coverage. An actor accesses a service providing a coverage describing the variation in a soil property: — across the surface of the earth; — at one or more depths/depth intervals beneath the surface; and/or — as a continuous vertical function for a soil profile or pedon.
Implementation	Web Processing Service Web Feature Service (input)

Annex B

Soil Information Models

B.1 Australia and New Zealand Soil Mark-up Language (ANZSoilML)

Agencies	CSIRO (Australia); Landcare Research (New Zealand)
Synopsis	ANZSoilML was originally developed in Australia (as OzSoilML) to capture soil descriptions created according to the Australian Soil and Land Survey Field Handbook. It was intended to support the exchange of soil description and laboratory data between Australian state and federal government agencies. The design of the model made reference to the SoTerML, INSPIRE Soil and ISO SoilML models. It became ANZSoilML when it was adopted in NZ, both for data exchange and also as the conceptual model for the NZ National Soil Data Repository. This work included the extension of the model to define a profile of Observations and Measurements to support the GlobalSoilMap.net consortium's metadata requirements for the individual cells in its global grid of soil property predictions. It is currently at version 2.0.1
Documentation	http://anzsoil.org/anzsoilml/
Scope	Site registration; soil formation, description, sampling, observation, classification and mapping; vocabularies.
Modeling approach	Comprehensive; hard-typed, additional properties can be provided using related Observations and Measurements observations.
Imported models	Observations and Measurements 2.0 (sampling and observation data) and GeoSciML 3.1 (extends Earth Material for soil material descriptions).
Accessibility	Publically available.
XML Schema	http://anzsoil.org/def/schema/
Implementation readiness	Medium-high. Complex model requires skilled/experienced technical staff (database and service configuration).
Implementation	Yes

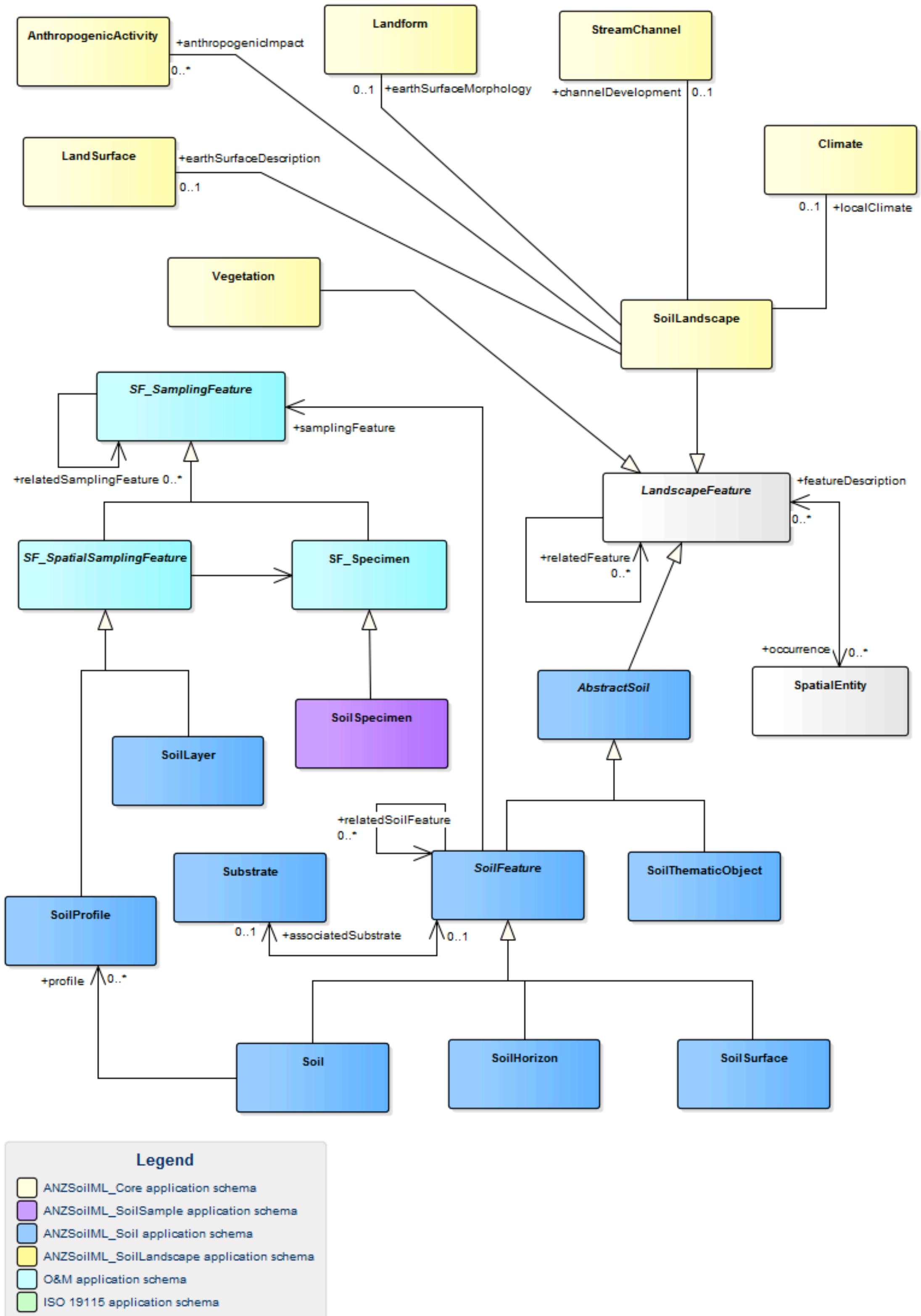


Figure B.1 — Australia and New Zealand Soil Mark-up Language overview

B.2 e-SOTER Soil and Terrain Mark-up Language (SoTerML)

Agencies	ISRIC (Netherlands); University of Miskolc (Hungary); BGR (Germany); JRC (EU); Cranfield University (UK); Szent Istvan University (Hungary); Scilands GmbH (Germany); INRA (France); University of Nottingham (UK); Czech University of Lfe Sciences (Czech Republic); Chinese Academy of Sciences (China); INRA-Maroc (Morocco); Wageningen University (Netherlands)
Synopsis	"The e-SOTER project aims at providing web services for soil and terrain data based on […] quantitative mapping of landforms; soil parent material and soil attribute characterisation using pattern recognition on remote sensing data; and standardisation of methods and measures of soil attributes to convert legacy data." "The design of SoTerML had to take into account two competing aspects. The first is the various soil attribute data profiles or classifications already used in SOTER databases. The three main data designs are: the SOTER profile (Oldeman and van Engelen, 1993) [9], WRB (2006) (World Reference Base of Soil Resources) (WRB, 2006) [8] and the FAO classification schemas (FAO, 1988) [2]. The second aspect is related to the interoperability developments led by OGC (Open Geospatial Consortium) concerning geographical datasets. On the one hand the goal is to be able to facilitate understanding and transfer from existing formats and data models, whilst on the other the aim is to provide exchange of datasets necessitating harmonisation and standardisation with compliancy to existing standards." "The approach in designing SoTerML data attributes was to incorporate as much flexibility and reuse as possible. Different elements in the class hierarchy of SoTerML require attributes to be associated with them without restricting their numbers or their data types or specificities." *From: http://dx.doi.org/10.1016/j.cageo.2011.11.026*
Documentation	http://www.esoter.net/content/standards-and-services-soil-and-terrain-data-exchange-soterml
Scope	Site registration; soil formation, description, sampling, observation, classification and mapping.
Modeling approach	Framework; soft-typed, no soil properties specified – added using a structured attribute data type.
Imported models	None
Accessibility	Publically available.
XML Schema	http://www.isric.org/specification/SoTerML.xsd
Implementation readiness	Medium. Requires development of property lists.
Implementation	Yes

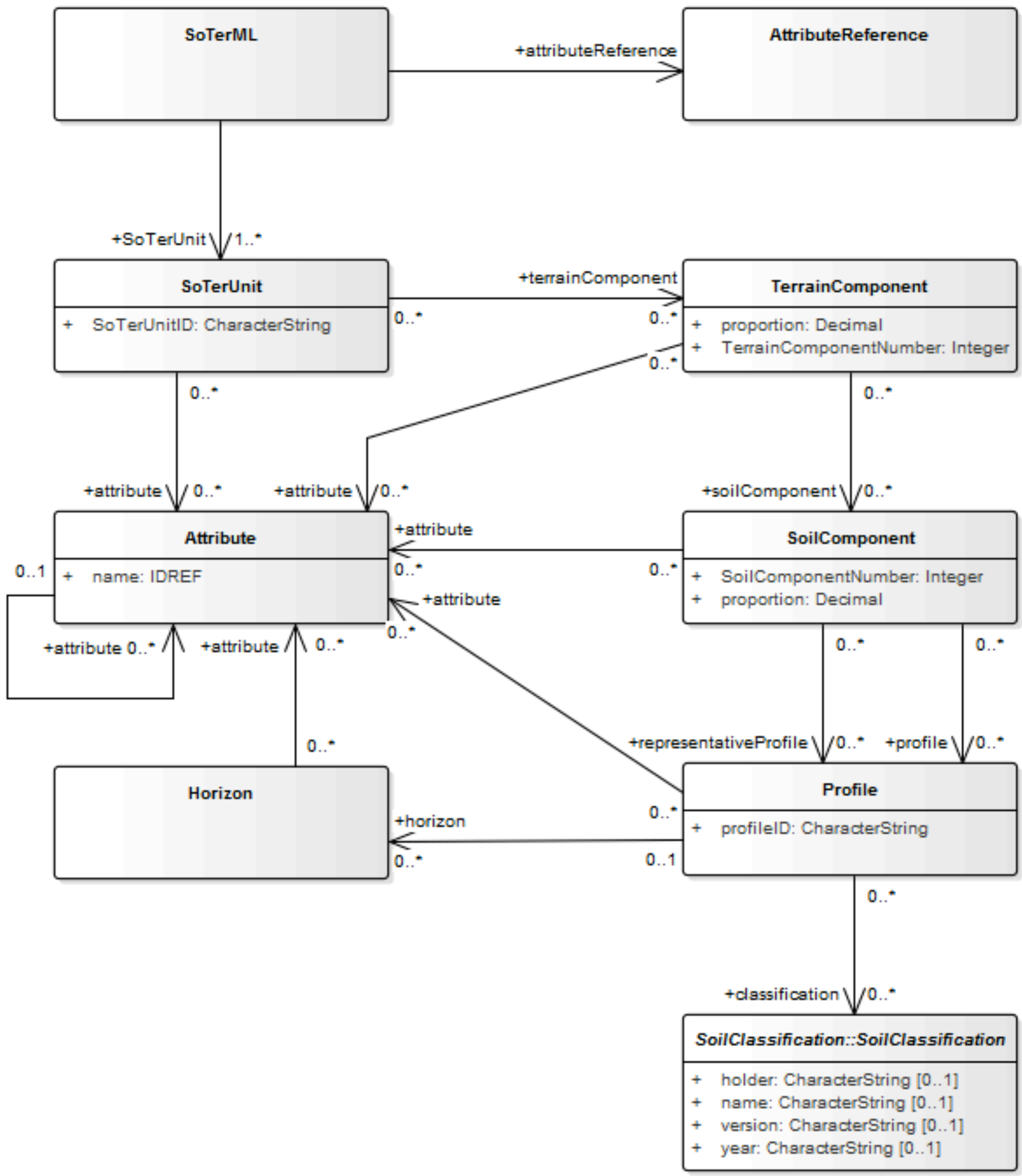


Figure B.2 — Soil and Terrain Mark-Up Language (SoTerML) overview

B.3 INSPIRE D2.8.III.3 Data Specification on Soil (INSPIRE Soil)

Agencies	INSPIRE Thematic Working Group, Soil
Synopsis	"Based on the definition given by the Directive (2007/2/EC), the scope for the soil theme covers: a) Soil inventories, providing one-off assessments of soil conditions and/or soil properties at certain locations and at a specific point in time, and allow soil monitoring, providing a series of assessments showing how soil conditions and/or properties change over time. b) Soil mapping, providing a spatial presentation of the properties linked to the soils, including soil types; typically, soil maps are derived with the help of data available in soil inventories. Also other soil related information derived from soil properties, possibly in combination with non-soil data are within the scope." *From: see Documentation.*
Documentation	http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_SO_v3.0rc3.pdf
Scope	Site registration; soil description, sampling, observation, classification and mapping; vocabularies.
Modeling approach	Targeted; 'essential' properties are hard-typed, all others soft-typed – values provided using associated O&M observations (certain properties are from defined code lists that may be extended).
Imported models	Observations and Measurements 2.0 (sampling and observation data).
Accessibility	Publically available.
XML Schema	http://inspire.ec.europa.eu/schemas/so/
Implementation readiness	Medium. Requires development of property lists.
Implementation	Yes

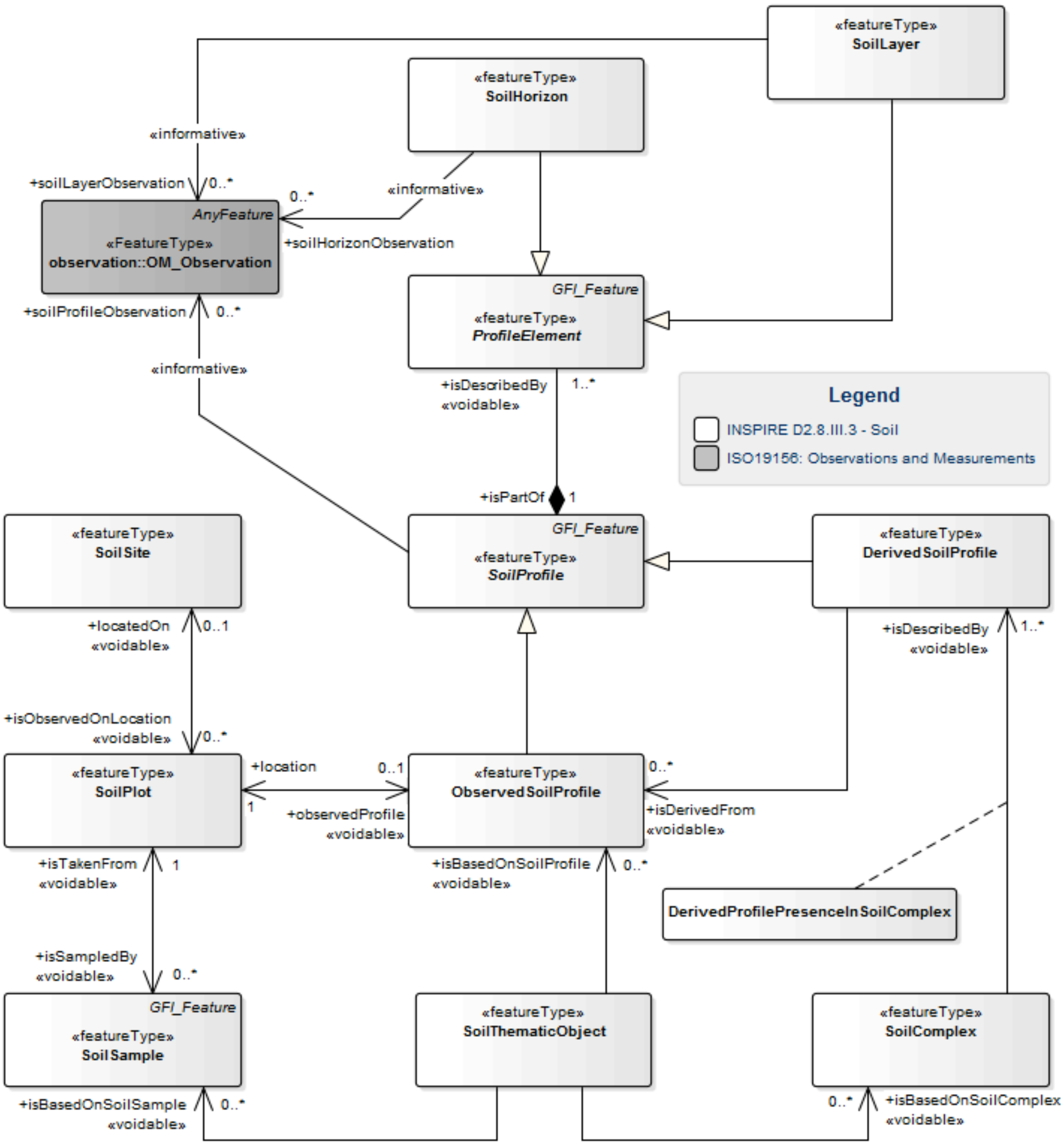


Figure B.3 — INSPIRE D2.8.III.3 Data Specification on Soil overview

B.4 ISO 28258:2013 Soil quality – Digital exchange of soil-related data (SoilML)

Agencies	ISO Technical Committee 190 (ISO/TC 190)
Synopsis	"This International Standard contains definitions of features, several parameter specifications and encoding rules that allow consistent and retrievable data exchange. It also allows the explicit geo-referencing of soil data by building on other International Standards, thus facilitating the use of soil data within geographical information systems (GIS). Because soil data are of various origins and are obtained according to a huge variety of description and classification systems, this International Standard provides no property catalogue, but a flexible approach to the unified encoding of soil data by implementing the provisions of ISO 19156 observations and measurements (OM) for use in soil science." *From: https://www.iso.org/obp/ui/#iso:std:iso:28258:ed-1:v1:en*
Documentation	https://www.iso.org/obp/ui/#!iso:std:44595:en
Scope	Site registration; soil sampling, observation, classification and mapping.
Modeling approach	Framework; soft-typed, no soil properties specified – added by creating application schema that extend this model.
Imported models	Observations and Measurements 2.0 (sampling and observation data).
Accessibility	Restricted, charge for access.
XML Schema	No normative instance (example in informative annex to the standard)
Implementation readiness	Low. Requires development of schema to extend the model and define properties; no available core GML schema (see above).
Implementation	Yes (portal specific extension)

B.5 IUSS/ISO 'Wageningen Proposal'

Agencies	International Union of Soil Sciences Working Group – Soil Information Standards (IUSS WG-SIS); ISO Technical Committee 190 (ISO/TC 190)
Synopsis	This model was developed to draft stage by representatives of IUSS WG-SIS and ISO/TC 190. It extends the scope of the ISO 28258:2013 model and addresses known issues. At this stage it has a similar scope to, and is subject to the limitations of, the ISO 28258:2013 model but with reorganized classes relating to soil profiles and descriptions.
Documentation	Contact: Peter Wilson, Chair, IUSS WG-SIS (Peter.Wilson@csiro.au)
Scope	Site registration; soil sampling, observation, classification and mapping; vocabularies.
Modeling approach	Framework; soft-typed, no soil properties specified.
Imported models	Observations and Measurements 2.0 (sampling and observation data).
Accessibility	Restricted, draft.
XML Schema	None.
Implementation readiness	Low. Incomplete; no properties; no GML schema.
Implementation	No

Annex C

Utility Classes

A set of utility classes was defined in this model to provide placeholders for complex data types and classes (Figure C.1). Wherever possible the semantics of these classes was aligned with their equivalents in the ISO19100 series. Definitions are included in Table C.1.

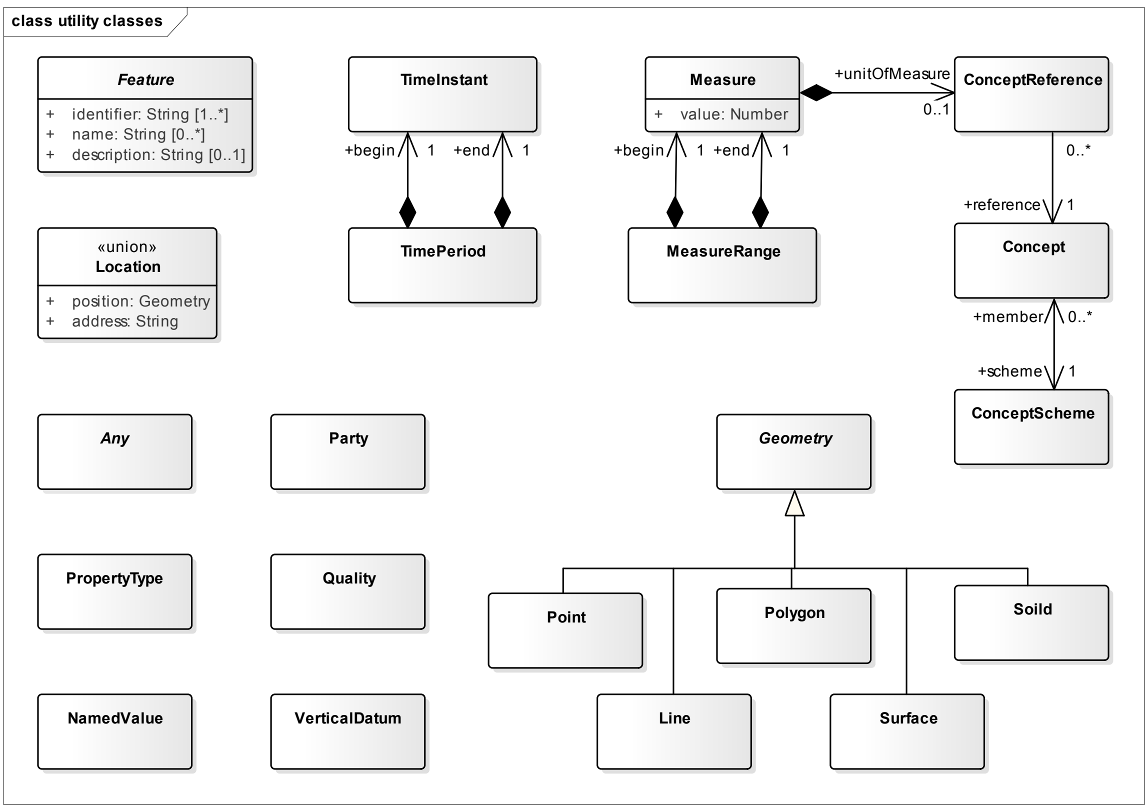


Figure C.1 — Conceptual utility class model

Table C.1 — Utility class definitions.

Class	Definition
Feature	as per General Feature Model.
Any	a wildcard class used when the target class is not known.
TimeInstant	an identifiable position in time. TimeInstants are used to mark the beginning or end of a TimePeriod.
Measure	a quantity value with a unit of measure. Measures are used to mark the beginning or end of a MeasureRange.
Concept	a term with a definition that is accessed via a ConceptReference and organized into a ConceptScheme.
Party	an individual or organization.
Location	a location of a Feature specified using an address or geometry.
PropertyType	a property of a Feature.
Quality	a representation of the quality of a value. A Quality could be expressed as a category (for example 'poor') or a representation of uncertainty (for example a probability distribution curve).
NamedValue	a soft-typed property value pair.
VerticalDatum	the elevation used as the datum for height related measurements.
Geometry	a set of geometric primitives including **Points**, **Lines**, **Polygons**, **Surfaces** and **Solids**.

Annex D

Demonstration Client Workflows

D.1 Demonstration 1: Landcare Research

Step	Description	Components	Feature/Resource Type
1	Home, map of soil sites	WFS	sams:SF_SpatialSamplingFeature
	Implemented using a bounding box (based on the map extent) filtered WFS GetFeature request.		
2	Click on link for individual sampling feature to view full site description. Choice: click on link to view specimens taken at the site (step 3) or the description of the soil at the site (step 4)	PID Service; WFS	sams:SF_SpatialSamplingFeature
	Implemented using an HTTP GET request to the feature URI (provided as an @xlink:href value). The PID service redirects the client to a WFS request to a service that serves the feature instance.		
3	View list of specimens at the site	PID Service; WFS	sams:SF_SpatialSamplingFeature
	Implemented as per Step 2.		
3.1	Click on a link to view an individual specimen and its related observations.	PID Service; WFS	spec:SF_Specimen; om:OM_Observation
	Implemented as per Step 2.		
3.2	Hover over an observed property or procedure value to view a definition.	PID Service; Linked Data Registry	skos:Concept
	Implemented using an HTTP GET request to the Concept URI (provided as an @xlink:href value). The PID service redirects the client to the Linked Data Registry which responds with data encoded according to the MIME type specified in the original GET request's HTTP Accept header.		
4	View soil description	WFS	soil:SO_Soil
	Implemented as per Step 2.		
4.1	Hover over a drainage class value to view a definition.	PID Service; Linked Data Registry	skos:Concept
	Implemented as per Step 3.2.		

4.2	Click on a link to view an individual soil horizon	PID Service; WFS	soil:SO_SoilHorizon
	Implemented as per Step 2.		
4.3	Hover over a soil texture, particle size class, observed property or procedure value to view a definition.	PID Service; Linked Data Registry	skos:Concept
	Implemented as per Step 3.2.		

D.2 Demonstration 2: Federation University of Australia

D.2.1 Map Selection

Step	Description	Components	Feature/Resource Type
1	Home, map of soil sampling sites	WFS	sams:SF_SpatialSamplingFeature
UC1	*Implemented using a bounding box (based on the map extent) filtered WFS GetFeature request.* *GML Features are converted to GeoJSON objects before being drawn as markers.*		
2	Click on a site to view the individual samples taken at this location. Choice (for each specimen) : click on link to view the sampled soil horizon (step 3) or the observations made against the specimen (step 4)	PID Service; WFS	spec:SF_Specimen
UC1	*Implemented using an HTTP GET request to the feature URI (provided as an @xlink:href value). The PID service redirects the client to a WFS request to a service that serves the feature instance.*		
3	Click on the sampled horizon link for individual specimens at each site to view the related horizon description.	PID Service; WFS	soil:SO_Horizon
UC1	*Implemented as per Step 2.*		
3.1	Click on a link to view the summary of the description of the soil at the site. (Click 'Back' button when finished.)	PID Service; WFS	soil:SO_Soil
UC1	*Implemented as per Step 2.*		
3.2	Click on the 'Derive additional properties (WPS)' button to generate additional soil properties (soil wilting point, field capacity and bulk density)	WPS; PID Service; WFS	soil:SO_Horizon
UC4	*A WPS Execute request is sent to the ISRIC WPS, the target soil horizon URI is a parameter value. The WPS dereferences the URI as per Step 2, runs the process on the horizon data and returns the set of property values as a soil:soParticleSizeDistribution element in a wps:Output.*		
4	Click on the link to view all related specimen observations,	PID Service; WFS	om:OM_Observation
	Implemented as per Step 2.		

D.2.2 Property Filter Selection

Step	Description	Components	Feature/Resource Type
1	Home, map of soil sampling sites	WFS	sams:SF_SpatialSamplingFeature
UC1	*Implemented using a bounding box (based on the map extent) filtered WFS GetFeature request.* *GML Features are converted to GeoJSON objects before being drawn as markers.*		
2	Click on ‘Soil Property Query’ menu button. Select the feature type (observation or soil horizon); depth (if horizon); target service provider; and property (bulk density, organic carbon, pH, electrical conductivity or effective cation exchange capacity). Enter a value with an appropriate comparison operator (<, >, =).	Client	N/A
UC1	*Implemented using an HTTP POST request to the target server(s) that contains an XML encoded WFS GetFeature request including an appropriate ows:Filter element.*		
3	Explore the result set as per steps 2 to 4 above.	PID Service; WFS	soil:SO_Horizon
UC1	*Implemented as per Steps 2 to 4 above.*		

D.3 Demonstration 3: Landcare Research

Step	Description	Components	Feature/Resource Type
1	Select data provider and monitoring station.	Client	N/A
	Configures the service to make a GetObservation request to an appropriate SOS endpoint for the defined feature of interest. *Site locations were taken from cached data due to difficulties getting monitoring site data from the SOS instances used for this demonstration.*		
2	Draw graphs of time variation of soil moisture and rainfall measures at selected site.	XSLT Mediator; SOS	om:OM_Observation/om:result: tsml:TimeseriesTVP or tsml:TimeseriesDomainRange
	Implemented using SOS GetObservation requests for each of the observed properties.		

Annex E

GML Instance Documents

Example GML XML Schema Documents (XSD) and instance documents are published in the OGC Soil Data IE GitHub Repository.

SoilIEML Schema Document:

https://github.com/opengeospatial/SoilDataIE/schema/soil-data-ie.xsd

Instance documents:

https://github.com/opengeospatial/SoilDataIE/examples/

Bibliography

[1] Eberhardt, E., Kozel, J., Fitzner, D., Templer, S., Hudson, G.: Data exchange with the GS Soil Generic application schema for soil information – a soil data cookbook using ISO DIS 28258. Last accessed October 2013 at http://www.gssoil-portal.eu/best_practice/gs_soil_d4.2_cookbook.pdf (2012)

[2] FAO: Soil Map of the World Revised Legend, World Soil Resources Report 60. Food and Agriculture Organization of the United Nations, Rome (1988)

[3] FAO: Guidelines for Soil Description, 4th edition. Food and Agriculture Organization of the United Nations, Rome (2006)

[4] Global Soil Map: Specifications Tiered GlobalSoilMap products. Release 2.4. Last accessed May 2016 at http://globalsoilmap.net/system/files/GlobalSoilMap_specifications_december_2015.pdf (2015)

[5] Isbell, R.F and National Committee on Soil and Terrain: The Australian Soil Classification. 2nd edition. CSIRO Publishing (2016)

[6] ISO: ISO 19150-2:2015 – Geographic information -- Ontology -- Part 2: Rules for developing ontologies in the Web Ontology Language (OWL) (2015)

[7] ISO: ISO 28258:2013. Soil quality — Digital exchange of soil-related data (2013)

[8] IUSS: World reference base for soil resources 2006. A framework for international classification, correlation and communication. World Soil Resources Reports 103. Food and Agriculture Organization of the United Nations, Rome (2006)

[9] OGC: OGC 15-100r1 – OGC Observations and Measurements – JSON implementation Discussion Paper (2015)

[10] Oldeman, L.R. and V. van Engelen: A world soils and terrain digital database (SOTER) – an improved assessment of land resources. Geoderma, 60 (1993)

[11] Schoeneberger, P.J., D.A. Wysoki, E.C. Benham, and Soil Survey Staff: Field book for describing and sampling soils, Version 3.0. Natural Resources Conservation Service, National Soil Survey Center, Lincoln, NE (2012)

[12] Soil Science Glossary Terms Committee: Glossary of Soil Science Terms 2008. Soil Science Society of America (2008)

[13] Soil Survey Staff: Soil Taxonomy: a Basic System of Soil Classification for Making and Interpreting Soil Surveys. 2nd edition. Agricultural Handbook 436, Soil Conservation Service, US Department of Agriculture, Washington , DC (1999)