An Ontology of Soil Properties and Processes

Heshan Du[1], Vania Dimitrova[1], Derek Magee[1], Ross Stirling[2], Giulio Curioni[3], Helen Reeves[4], Barry Clarke[1], Anthony Cohn[1]

[1] University of Leeds
[2] Newcastle University
[3] University of Birmingham
[4] British Geological Survey

Abstract. Assessing the Underworld (ATU) is a large interdisciplinary UK research project, which addresses challenges in integrated inter-asset maintenance. As assets on the surface of the ground (e.g. roads or pavements) and those buried under it (e.g. pipes and cables) are supported by the ground, the properties and processes of soil affect the performance of these assets to a significant degree. In order to make integrated decisions, it is necessary to combine the knowledge and expertise in multiple areas, such as roads, soil, buried assets, sensing, etc. This requires an underpinning knowledge model, in the form of an ontology. Within this context, we present a new ontology for describing soil properties (e.g. soil strength) and processes (e.g. soil compaction), as well as how they affect each other. This ontology can be used to express how the ground affects and is affected by assets buried under the ground or on the ground surface. The ontology is written in OWL 2 and openly available from the University of Leeds data repository: `http://doi.org/10.5518/54`.

Keywords: OWL ontology, soil property/process, asset maintenance

1 Introduction

Assessing the Underworld (ATU) project is a large interdisciplinary UK research project, which addresses challenges in asset maintenance, especially how to reduce the economic, social and environmental costs or impacts of streetworks required in the maintenance of roads and buried assets (e.g. pipes and cables). Existing asset management systems (e.g. UK Pavement Management Systems) can help local authorities or utility companies with financial reporting or assessing the economic costs of the construction, repairing and replacement of assets, but provide less support in assessing the environmental/social impacts and the impacts on other assets, which are also important for enabling better informed decisions in asset maintenance. To establish the total cost (social and environmental impacts along with economic costs) of asset maintenance activities, it is essential to understand how assets affect each other and how they affect or are affected by the natural environment and human activities. Within the ATU project, we are developing a series of ontologies based on the knowledge and expertise in multiple areas for describing how buried assets, soil, roads, the natural environment and human activities affect each other. The ontology of soil properties and processes (OSP) is a central ontology in this series.

The condition of an asset can affect and be affected by assets close to it. For example, a water pipe burst increases the water content of the soil (including man-made fills) surrounding it. As a consequence, the strength of soil decreases, which has negative effects on the function of soil for supporting buried pipes and cables nearby, as well as the roads or pavements above. In many interactions between assets, as shown in the example above, soil plays a role as a medium. In other words, the conditions of roads and buried assets affect each other through soil, and soil directly affect and is affected by the conditions of roads and buried assets. Therefore, when assessing the impacts of decisions (e.g. whether to fix a water pipe leakage immediately) for surface or buried asset maintenance, it is important to have a basic knowledge of soil properties and processes.

Soil has many properties and is involved in various processes, as described in [16, 15]. The relationships between soil properties and processes are complicated. Whilst it is often easy to know what a process or property affects directly, it is more difficult for people to answer questions like 'what are the factors affecting or affected (directly or indirectly) by a particular property or process of soil?'. To answer such questions, it is necessary to handle information about soil properties, processes and their relationships automatically. This requires a proper ontological model and automated reasoning.

There exist several ontologies [17, 9, 20, 12, 7, 18, 8, 1, 2], where soil is defined. Some of them [17, 7, 1, 2] are general environmental or agricultural ontologies, whilst others are specialized for describing soil. These soil ontologies [9, 20, 12, 18, 8] are not publicly available. Most of the existing soil ontologies [9, 20, 18, 8] describe classifications of soil or different types of soil but do not elaborate the various soil properties and processes. The ontologies described in [12, 2] define some soil physical/chemical properties, but limited to those relevant to farming or agricultural applications. None of the existing ontological models provides a systematic and comprehensive description of soil properties and processes, and none of them defines how soil properties and processes affect each other.

The paper addresses this gap. We present a new ontology for describing soil properties and processes, as well as how they affect each other. It reuses and specifies high-level classes in NASA's Semantic Web for Earth and Environmental Terminology (the SWEET ontology) [17], which is widely adopted and extended. The ontology is developed using the NeOn methodology [19]. It is written in OWL 2 Web Ontology Language Manchester Syntax [14], which is based on description logic (DL) [6]. The DL expressivity of the ontology is $\mathcal{SRI}$, allowing transitive relations and inverse relations[5]. The ontology contains 592 concepts and 2243 relation statements (OWL logical axioms), which are based on the knowledge of soil experts[6], the SWEET ontology [5, 17], English dictio-

[5] To avoid confusion, we call 'OWL object properties' relations.

[6] Soil experts were involved in the development and evaluation of the OSP ontology. They checked 1407 relation statements for describing 193 main classes.

naries [3, 4] and a textbook on soil physics [16]. The ontology, together with a tutorial of viewing and querying it, a translation of its main relation statements in natural language and a feedback form, is publicly available [11] from the University of Leeds data repository: `http://doi.org/10.5518/54`, under the license Creative Commons Attribution 4.0 International (CC BY 4.0)[7].

The rest of the paper is structured as follows. Section 2 describes how the main concepts and relations are defined in the ontology. Section 3 explains how to reason with and query the ontology. Section 4 illustrates the extensibility of the ontology. Section 5 discusses its applications. Section 6 concludes the paper.

2 Defining Soil Properties and Processes

The ontology of soil properties and processes (OSP) defines two main high-level classes or categories: *SoilProperty* and *SoilProcess*. The classes *SoilProperty* and *SoilProcess* are specifications of the classes *Property* and *Process* in the SWEET ontology [5, 17] for soil. A property is an attribute, quality, or characteristic of something [4]. A soil property is a property of soil. A process refers to a series of changes that happen naturally over time [3, 4]. A soil process is a process involving soil. Following the style of the SWEET ontology, we define different kinds of soil properties and processes as classes in the OSP ontology and classify them into physical, chemical and spatial/biological categories. Table 2 summarizes the number of subclasses of *SoilProperty* and *SoilProcess* in different categories. Though the OSP ontology covers classes in all these categories, it mainly describes soil physical properties and processes.

Table 1. Number of subclasses of *SoilProperty* and *SoilProcess* in the OSP ontology

SoilProperty	Subclasses	SoilProcess	Subclasses
SoilPhysicalProperty	176	SoilPhysicalProcess	111
SoilChemicalProperty	16	SoilChemicalProcess	29
SoilSpatialProperty	4	SoilBiologicalProcess	8

In addition to classifying soil properties and processes into different categories as the SWEET ontology does, we also define soil properties and processes regarding how they affect each other. The main types of relations defined in the OSP ontology are *hasImpactOn* and its inverse *influencedBy*, meaning ‘affects or changes’ and ‘is affected or changed by’ respectively. The relations *hasImpactOn* and *influencedBy* are both defined as transitive. A process q *hasImpactOn* a property p, if q causes a change in p. A property p *hasImpactOn* a process q, if a change in p changes how the process q goes. A property p_1 *hasImpactOn* a property p_2, if a change in p_1 causes a change in p_2. A process q_1 *hasImpactOn*

[7] https://creativecommons.org/licenses/by/4.0/

a process q_2, if q_1 changes how the process q_2 goes. Other relations defined in the OSP ontology include *hasPossibleCause* and its inverse *hasPossibleEffect*, meaning 'has a possible reason' and 'has a possible consequence' respectively. The word 'possible' means the cause/effect exists in some situation, but may exist or not exist in a particular real world situation considered.

In the OSP ontology, we define relationships between soil properties and processes at the concept level rather than at the individual level. For example, we express 'soil water content *hasImpactOn* soil strength' in description logic as $SoilWaterContent \sqsubseteq \exists hasImpactOn.SoilStrength$. Such relation statements are not defined for a particular type of soil at a certain location and depth, but generally applicable to any soil. The statement 'soil water content *hasImpactOn* soil strength' means that for any soil, a change in its water content causes a change in its strength[8]. Similarly, the statement 'soil compaction *hasImpactOn* soil air content' means for any soil, a compaction applied to it changes the air content of it. The statement 'soil compaction *hasImpactOn* soil air movement' means for any soil, a compaction applied to it changes how the air in it moves. The statement 'soil air movement is *influencedBy* soil porosity' means for any soil, a change in its porosity changes how the air in it moves. Note that inverse relations are not always reciprocally asserted in the manually-written ontology.

We rank subclasses of *SoilProperty* and *SoilProcess* based on their usages or how many statements a class is involved. If a class is involved in a statement, then any class equivalent to it is also considered to be involved in the same statement. Table 2 lists the top 20 classes in this ranking, which illustrates the main soil properties and processes defined in the OSP ontology well. For each set of equivalent classes, only one class in it is included in the list to represent the whole set. When naming a soil property/process class, 'Soil' is added to the front to avoid potential name conflicts. A class name indicates its intended meaning. For instance, the class name *SoilWaterInfiltration* is intended to represent the concept of 'infiltration of water into soil'.

3 Querying the OSP ontology

Description logic (DL) reasoners can be used to reason about statements in the OSP ontology. A basic question a reasoner can answer is whether a statement (e.g. '*SoilWaterEvaporation hasImpactOn SoilStrength*') can be inferred from statements in a given ontology. A justification based explanation framework [13] can be used together with a DL reasoner to explain how a new statement is inferred from existing statements. By reasoning with the OSP ontology, more advanced questions can be answered, such as, for given a class C, list all the classes that C *hasImpactOn* or *influencedBy*.

[8] According to DL/OWL semantics, it means 'every soil water content *hasImpactOn* some soil strength'. The interpretation adopted here is more strict, however, this makes more sense for soil experts and the standard DL inference rules are still sound. In practice, we only expect the model to be applied to a local spatial context.

Table 2. Ranking the subclasses of *SoilProperty* and *SoilProcess* in the OSP ontology by the number of statements involved (usages) — the Top 20

	SoilProperty	Usages	SoilProcess	Usages
1	SoilMoistureContent	93	SoilAeration	70
2	SoilStructure	93	SoilCompaction	59
3	SoilTexture	67	SoilWaterInfiltration	40
4	SoilPorosity	62	SoilCrusting	40
5	SoilTemperature	59	SoilWaterEvaporation	39
6	SoilStrength	52	SoilAggregation	39
7	SoilClayMineral	49	SoilWaterMovement	32
8	SoilBulkDensity	40	SoilErosion	29
9	SoilOrganicMatterContent	39	SoilDrying	29
10	SoilHydraulicConductivity	36	SoilShrinkage	28
11	SoilParticleSpecificSurfaceArea	31	SoilWaterRetention	28
12	SoilClayContent	30	SoilSoluteTransport	26
13	SoilPoreSizeDistribution	28	SoilMicrobialActivity	21
14	SoilCohesion	23	SoilHardsetting	20
15	SoilPlasticity	22	SoilWaterTransmission	19
16	SoilMoisturePotential	18	SoilFreezing	19
17	SoilParticleArrangement	18	SoilDeformation	18
18	SoilBearingCapacity	18	SoilSwelling	18
19	SoilOxygenConcentration	18	SoilDispersion	18
20	SoilAirContent	17	SoilGaseousEmission	17

A simple way to reason with and query the OSP ontology is to use an ontology editor Protégé[9] and its reasoner plugins. Within the DL query tab of Protégé, we can execute queries like getting all subclasses of the class expression '(*hasImpactOn* some *SoilStrength*) and *SoilProperty*'. For each class listed in the query results, its explanations can be obtained easily. We provide a tutorial [10] illustrating (with figures and examples) how to query the OSP ontology. It is available at: `http://doi.org/10.5518/54`, the same DOI of the OSP ontology.

4 Extensibility of the OSP ontology

The OSP ontology extends three other top-level classes, *Phenomena*, *HumanActivity* and *Substance*, of the SWEET ontology [5, 17]. A phenomenon refers to a fact/situation that exists and can be observed [3, 4]. A human activity refers to a thing that a person/group does or has done [4]. The concept *Substance* [5] covers living entities and non-living entities, such as animals, plants, material things, etc. By including these concepts, the OSP ontology can express how soil properties and processes affect/are affected by other environmental factors (such as water, air, weather and trees) and human activities.

[9] http://protege.stanford.edu

For each of the five top-level classes, its next-level classes in the OSP ontology are shown in Fig. 1. Some of the next-level classes (e.g. *PlanetaryPhenomena*, and *MaterialThing*) are inherited from the SWEET ontology, whilst others (e.g. *PlantProcess* and *SoilSubstance*) are created by specifying high-level concepts in the SWEET ontology for a particular type of objects. All the high-level categories can be extended and enriched further easily by reusing concepts in the SWEET ontology and following similar ways as defining soil properties and processes in Section 2. In addition, more high-level categories can be added easily and linked to the existing classes. We are developing ontologies for describing pipe/road properties, processes and phenomena (e.g. defects or failures) in a similar way and linking them to the OSP ontology.

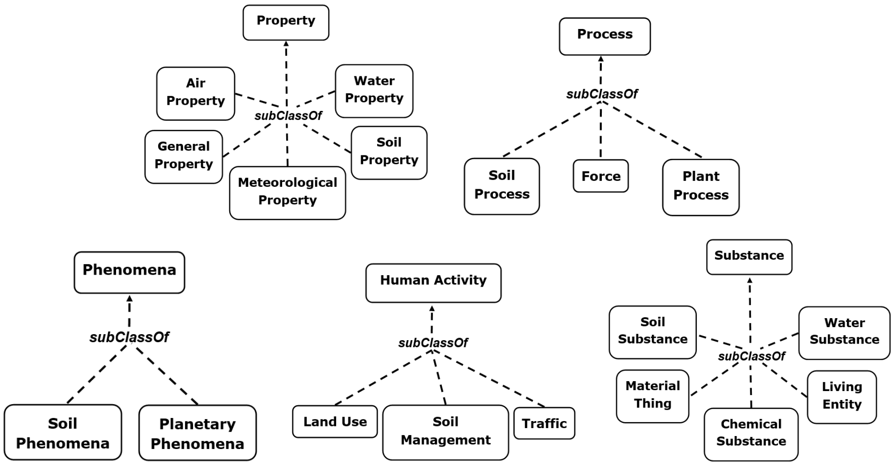

Fig. 1. Class hierarchies of the five top-level classes in the OSP ontology

Note that the OSP ontology mainly describes soil properties and processes from a soil science perspective, but also captures main terminologies (e.g. soil compaction) used in engineering applications. We are aware that terminological differences exist between soil scientists and geotechnical engineers. We follow the concepts defined in soil science mainly because they are more understandable by non-specialists. The OSP ontology can be extended and enriched with parameters[10] (e.g. soil shear strength) and calculations defined in soil mechanics [15].

5 Application

The OSP ontology is intended to be used in a software system that supports inter-asset management/maintenance decisions (ATU decision support system).

[10] The OSP ontology describes several parameters used for assessing soil structure, under a class called *SoilStructureIndex*.

When making decisions for maintaining a single asset, people (e.g. local authorities and utility companies) often consider the economic cost but may not be aware or take into account the impacts of their decisions or actions on other assets, the natural environment and our daily life. The ATU decision support system will prompt people into asking a series of questions, the answers to which will lead to a complete understanding of the economic, social and economic impacts of different options for dealing with an asset defect/failure. By developing the OSP ontology, we establish a general framework, which extends the SWEET ontology, for expressing how properties, processes, phenomena, human activities and substances affect each other. In this way, many environmental and social impacts can be represented in an ontology or inferred from it using automatic reasoning. The OSP ontology can help users understand the environmental impacts of a process or a human activity related to soil. For example, knowing that a water pipe leakage affects soil water content, we may ask the OSP ontology about the impacts of a change in soil water content. By executing a DL query of getting subclasses of '*influencedBy* some *SoilWaterContent*' in Protégé, a list of 264 classes is returned[11], including various soil properties and processes, as well as factors like building stability, plant growth, air quality, water quality, etc. For classes listed, explanations are provided, which are useful for identifying and understanding indirect or hidden impacts.

The soil properties and processes described in the OSP ontology have many agricultural, engineering and environmental applications (see the subclasses of *SoilFunction*). For example, the OSP ontology describes how the growth of a plant is affected by soil and other environmental factors. The OSP ontology can be applied to a range of contexts that require descriptions of soil properties or processes. For example, the EU project NeTTUN[12] utilises soil properties and processes to provide contextual factors for tunnel construction and maintenance. By extending the OSP ontology and linking it to ontologies for describing the environment, infrastructure assets, human activities and economic cost models, we envisage that the OSP ontology will make a useful contribution to ontological models for environmental, social and economic sustainability.

6 Conclusion

We present a new ontology OSP for describing soil properties, processes and how they affect each other. It is created using reliable knowledge sources and extends the SWEET ontology. It can be reasoned with automatically using DL reasoners and queried using Protégé. The OSP ontology helps people to understand indirect and complicated relationships between soil properties and processes, as well as many environmental impacts of a process or a human activity related to soil. The OSP ontology is easy to (re)use, easily extensible, generally applicable, publicly available and findable. As the ATU decision support system develops, we will maintain and enhance this ontology, making new versions publicly available.

[11] Different types of factors can be filtered out easily using DL query.
[12] http://nettun.org

Acknowledgments. This research is supported by EPSRC under grant no. EP/K021699/1 which we gratefully acknowledge.

References

1. Eionet GEMET Thesaurus. http://www.eionet.europa.eu/gemet, 2015.
2. AGROVOC Multilingual agricultural thesaurus. http://aims.fao.org/standards/agrovoc, 2016.
3. Cambridge Dictionaries Online. http://dictionary.cambridge.org, 2016.
4. Oxford Dictionaries. http://www.oxforddictionaries.com, 2016.
5. SWEET: Semantic Web for Earth and Environmental Terminology. https://sweet.jpl.nasa.gov, 2016.
6. F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. *The Description Logic Handbook*. Cambridge University Press, 2007.
7. P. L. Buttigieg, N. Morrison, B. Smith, C. J. Mungall, and S. E. Lewis. The environment ontology: contextualising biological and biomedical entities. *Journal of Biomedical Semantics*, 4:43, 2013.
8. C. Deb, S. Marwaha, P. Malhotra, S. Wahi, and R. Pandey. Strengthening soil taxonomy ontology software for description and classification of USDA soil taxonomy up to soil series. In *Proceedings of the 2nd International Conference on Computing for Sustainable Global Development*, pages 1180–1184, 2015.
9. A. dos Santos Aparício, O. L. M. de Farias, and N. dos Santos. Integration of Heterogeneous Databases and Ontologies. *Cadernos do IME-Série Informática*, 21:4–10, 2006.
10. H. Du and A. Cohn. A Tutorial of Viewing and Querying the Ontology of Soil Properties and Processes. Technical report, University of Leeds, 2016.
11. H. Du and A. Cohn. An Ontology of Soil Properties and Processes. University of Leeds. [Dataset]. http://doi.org/10.5518/54, 2016.
12. T. Heeptaisong and A. Shivihok. Soil Knowledge-based Systems Using Ontology. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, pages 1–5, 2012.
13. M. Horridge. *Justification Based Explanation in Ontologies*. PhD thesis, University of Manchester, 2011.
14. M. Horridge and P. F. Patel-Schneider. OWL 2 Web Ontology Language Manchester Syntax. https://www.w3.org/TR/owl2-manchester-syntax, 2012.
15. J. Knappett and R. Craig. *Craig's Soil Mechanics*. CRC Press, 2012.
16. R. Lal and M. K. Shukla. *Principles of Soil Physics*. CRC Press, 2004.
17. R. G. Raskin and M. J. Pan. Knowledge representation in the semantic web for Earth and environmental terminology (SWEET). *Computers & Geosciences*, 31(9):1119–1125, 2005.
18. P. Shivananda and P. S. Kumar. Building Rules Based Soil Classification Ontology. *International Journal of Computer Science and Information Technology & Security*, 3(2), 2013.
19. M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López. The NeOn Methodology for Ontology Engineering. In *Ontology Engineering in a Networked World*, pages 9–34. 2012.
20. M. Zhao, Q. Zhao, D. Tian, P. Qian, and X. Zhang. Ontology-based intelligent retrieval system for soil knowledge. *WSEAS Transactions on Information Science and Applications*, 6(7):1196–1205, 2009.