

STWO: An Ontology for Soil Food Web Reconstruction

Nicolas Le Guillarme¹, Mickaël Hedde² and Wilfried Thuiller¹

¹Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA, Laboratoire d'Ecologie Alpine, F-38000 Grenoble, France

²INRAE, UMR Eco & Sols, Montpellier, France

Abstract

While food webs are pivotal tools to understand the structure, dynamics and functioning of ecosystems, their reconstruction is not trivial since feeding relationships are not always known. To this end, soil ecologists often simplify the problem by either grouping morphologically similar organisms into trophic groups with known interactions or by assuming that feeding relationships are predictable from consumer diets (e.g. frugivore or bacterivore). Interestingly, the scientific community has collected a considerable amount of information on trophic interactions and feeding habits. However, the large-scale exploitation of these data for food web reconstruction is hampered by the lack of standards for representing and reasoning upon trophic knowledge. The goal of our work is to propose an ontology that will support the automatic reconstruction of soil food webs from community composition data.

Keywords

soil ecology, food web, ontology development

1. Introduction

Food webs (also called trophic webs, trophic interaction networks) encode both the composition of ecological communities as well as the feeding relationships within the community. In soil ecology, food webs are often used to understand the structure and dynamics of soil assemblages, and their impact on decomposition processes and nutrient cycling [1]. Yet, reconstructing soil food webs is not a straightforward task. The nature of soil as a black-box ecosystem makes direct observations of most trophic interactions fairly impossible, and knowledge of resource preferences of many taxonomic groups of soil fauna are not well known. These preferences may be inferred from morphological similarities with species of known feeding habits or by phylogenetic proximity. This results in a more or less fine categorization of soil flora, fauna, fungi and microbes into a multitude of trophic groups (e.g. bacterivorous nematodes, arbuscular mycorrhizal fungi or saprotrophic fungi...). These trophic groups allow reconstructing simplified food webs that are expected to have the same structural properties as real soil food web. The development of high-throughput species identification methods (e.g. eDNA metabarcoding), and the availability of massive amounts of data about trophic interactions and feeding habits collected by researchers over the past decades have paved the way for new knowledge-based


S4BioDiv 2021: 3rd International Workshop on Semantics for Biodiversity, held at JWOW 2021: Episode VII The Bolzano Summer of Knowledge, September 11–18, 2021, Bolzano, Italy

✉ nicolas.leguillarme@univ-grenoble-alpes.fr (N. Le Guillarme); mickael.hedde@inrae.fr (M. Hedde)

ORCID 0000-0003-4559-7579 (N. Le Guillarme); 0000-0002-6733-3622 (M. Hedde); 0000-0002-5388-5274 (W. Thuiller)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

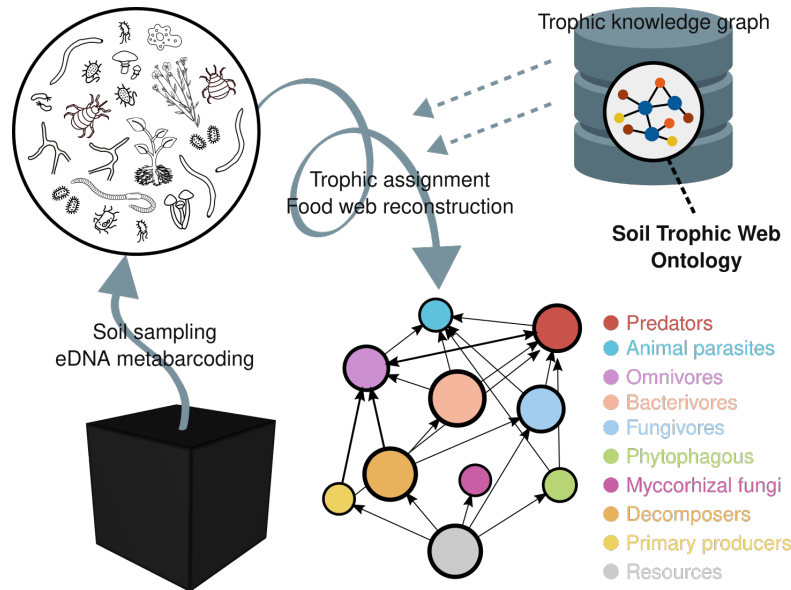


Figure 1: STWO formally describes knowledge about trophic groups and trophic interactions. Combined with trophic interaction records into a trophic knowledge graph, it can support the automatic reconstruction of food webs from compositional data.

methods to automate the reconstruction of food webs on an unprecedented scale [2]. Still, some problems need to be solved first, the most important of which is probably the lack of consensus on what exactly a trophic group is and how to classify soil organisms. This absence of formal description of trophic knowledge prevents the consistent mapping of consumers (e.g. *Acrobeloides sp.*) to their trophic group(s) (e.g. bacterivore) and resources (e.g. bacteria belonging to *Alphaproteobacteria*), which is needed to automatically assign a species to a trophic group based on its known interactions, or conversely, to predict the species potential interactions based on the trophic group(s) to which it belongs.

To address this issue, we are developing the Soil Trophic Web Ontology, whose role is to provide formal definitions of trophic groups that are consistent across all taxonomic groups of soil organisms, together with an ontological structure that enables inference in support of our main objective which is to automate the reconstruction of food webs from community inventories (Fig. 1).

2. The Soil Trophic Web Ontology

The Soil Trophic Web Ontology (STWO) is a domain ontology which represents and maps together knowledge about trophic interactions (consumer-resource relationships) and trophic groups (feeding habits, diets). Following good practices of ontology development, STWO is built by leveraging existing resources as much as possible. In particular, STWO extends a "trophic subset" of the ECOCORE ontology of core ecological entities with additional classes for missing trophic groups and resources. As much as possible, classes for resources are imported from existing OBO ontologies (Fig. 2). STWO includes classes to represent trophic groups at different

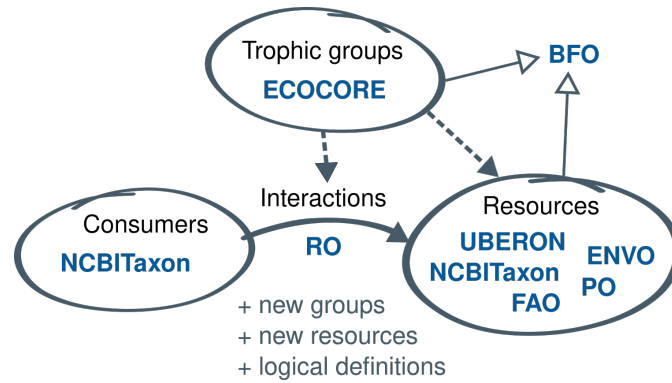


Figure 2: STWO imports modules from existing OBO ontologies and adds new classes for missing trophic groups and resources. Ontology reuse reduces modeling efforts and promotes interoperability.

resolutions (e.g. *heterotroph*, *decomposer*, *saprophylophage*). Trophic resources may be of different types: an organism represented by a taxonomic unit (e.g. *Bacteria*, *Fungi*, *Viridiplantae*), an anatomical part of an organism (e.g. *leaf*, *mycelium*, *blood*), and any type of environmental material (e.g. *carbon dioxide*, *soil organic matter*). STWO also reuses object properties from the Relation Ontology (RO) to describe trophic interactions (e.g. *eats*, *acquires nutrients from*). Finally, STWO provides logical definitions of trophic groups in the form of OWL equivalence axioms, which makes it possible, using an OWL reasoner, to infer the trophic group(s) an organism belongs to based on the resources it consumes, as well as to predict potential trophic interactions from its feeding regime (Fig. 3).

STWO development process is both collaborative and iterative. The initial version of the ontology was created from a list of relevant terms (trophic groups and resources) and their definitions created by a small specialized group of soil ecologists. A subset of these terms could be mapped to existing resources using the Ontobee search engine. These resource identifiers (URIs) were used as "seeds" for ROBOT's MIREOT extraction method [3]. This method enables to extract a subset of terms (a module) from an external ontology instead of importing the whole ontology, while preserving the subclasses/subproperties hierarchy. The resulting modules were merged to form the backbone of STWO. Missing classes and their logical definitions were added manually using the Protégé editor. The whole development workflow (extraction, merging, validation, release, versioning) is managed using the Ontology Development Kit [4]. Each new release of STWO is submitted to a group of >20 international experts in soil ecology to collect their feedback as well as suggestions for revisions and new terms. Debatable points are discussed and agreed upon using collaborative decision-making tools. The ontology is thus progressively corrected and enriched with each new iterations.

3. Project Status and Future work

STWO development is now in its second iteration. The first draft of the ontology includes over 60 newly-defined resource and trophic group classes. Experts are in the process of agreeing on the revisions to be made to the current version. Once the terms and structure of the ontology

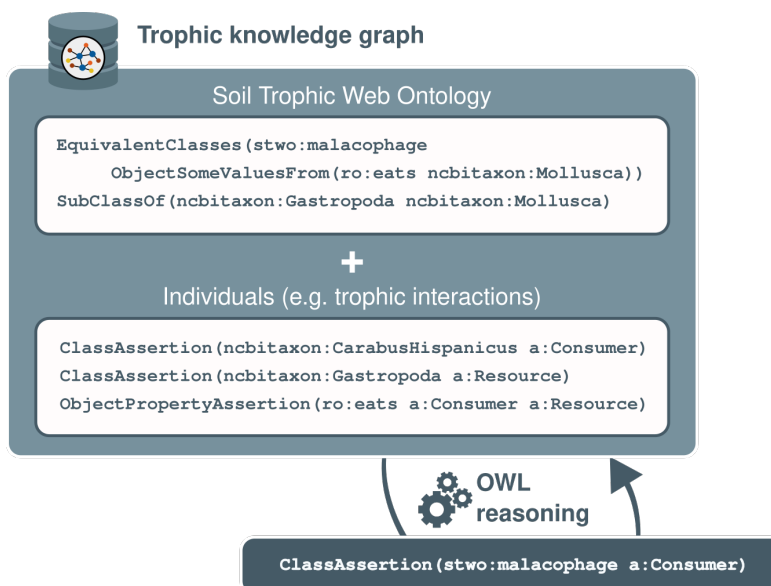


Figure 3: Logical definitions of trophic groups are key to STWO automated reasoning capabilities. In this example, the fact that *Carabus hispanicus* is a malacophagous organism is derived from the fact that *C. hispanicus* trophically interacts with gastropods, a subclass of molluscs.

are stabilized, we plan to submit a request for new content and revision to the ECOCORE team, so that STWO new classes and axioms for trophic groups and resources are made publicly accessible as part of ECOCORE. It is our wish that the work of our team of soil ecology experts benefits to a large community of users. We will also consider adding new properties to describe potential trophic interactions, which would be useful to distinguish between documented and inferred interactions (e.g. using trait matching).

Acknowledgments

The research received funding from the French Agence Nationale de la Recherche (ANR) through the GlobNets (ANR-16-CE02-0009) project and through MIAI@Grenoble Alpes (ANR-19-P3IA-0003).

References

- [1] S. Scheu, The soil food web: structure and perspectives, *European journal of soil biology* 38 (2002) 11–20.
- [2] Z. G. Compson, W. A. Monk, C. J. Curry, D. Gravel, A. Bush, C. J. Baker, M. S. Al Manir, A. Riazanov, M. Hajibabaei, S. Shokralla, et al., Linking DNA metabarcoding and text mining to create network-based biomonitoring tools: A case study on boreal wetland macroinvertebrate communities, *Advances in ecological research* 59 (2018) 33–74.
- [3] R. C. Jackson, J. P. Balhoff, E. Douglass, N. L. Harris, C. J. Mungall, J. A. Overton, ROBOT: a tool for automating ontology workflows, *BMC bioinformatics* 20 (2019) 1–10.

[4] N. Matentzoglou, INCATools/ontology-development-kit: June 2020 release (2021). doi:10.5281/zenodo.4973944.