



HAL
open science

State-of-the-art browsing of SKOS resources in AgroPortal

Clement Jonquet, Syphax Bouazzouni

► **To cite this version:**

Clement Jonquet, Syphax Bouazzouni. State-of-the-art browsing of SKOS resources in AgroPortal. NKOS Consolidated Workshop (NKOS 2023), J. Busch, D. Tudhope, M. Zeng, Nov 2023, Daegu, South Korea. hal-04102864

HAL Id: hal-04102864

<https://hal.science/hal-04102864>

Submitted on 22 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

State-of-the-art browsing of SKOS resources in AgroPortal

Clement Jonquet^{1,2}[0000-0002-2404-1582] and Syphax Bouazzouni¹[0000-0002-2069-8739]

¹ LIRMM, University of Montpellier & CNRS, France

² MISTEA, University of Montpellier, INRAE & Institut Agro, France
{jonquet, syphax.bouazzouni}@lirmm.fr

Abstract. Using standard formats such as OWL or SKOS has facilitated the development of a wide range of tools and applications for semantic artefacts. The challenge is now to package them inside semantic artefact catalogues where the cohabitation, interoperation and appropriate use of each type of semantic artefact (SKOS and OWL) will be possible. In this paper, we present the new functionalities to support SKOS resources in the AgroPortal (<https://agroportal.lirmm.fr>) vocabulary and ontology repository. We have created state-of-the-art innovative browsing approaches to discover and navigate SKOS concepts in thesauri that made an extensive use of scheme and collections. We have also enabled the use of mappings declared with SKOS properties.

Keywords: SKOS thesauri, semantic artefacts, ontology browsing, ontology repository, vocabulary service, SKOS schemes and collections, mappings.

1 Context and motivations

In the realm of semantic artefacts –a broader term to include ontologies, terminologies, taxonomies, thesauri, vocabularies, metadata schemas and standards– developed with semantic web technologies, SKOS resources are becoming more and more prominent. Typically, since its release in 2009, SKOS (Simple Knowledge Organization System) [1], the W3C Recommendation for thesauri and vocabularies, has been widely adopted in multiple domains from environmental sciences e.g., GEMET, to ecosystems research e.g., ANAEE Thesaurus, to agriculture e.g., AGROVOC, social sciences e.g., TheSoz, cultural heritage e.g., UNESCO Thesaurus, or even institutional standardization e.g., EuroVoc Thesaurus.

SKOS resources can be semantically very rich –not at the same level that OWL ontologies– especially in the way they regroup and classify concepts with sub-schemes and collections; therefore, specific software needs to be developed to handle and valorize this richness and complexity. One challenge is also to facilitate the cohabitation, interoperation and appropriate use of each type of semantic artefact (SKOS/OWL) in a common shared platform.

Often specific “terminology/vocabulary services” are designed to support a given SKOS thesaurus or set of thesauri such as the NERC vocabulary server, the CESSDA Vocabulary service or Humanum OpenTheso.¹ SKOSMOS [2] was certainly the first, and for sure is the most adopted, generic technology to easily deploy vocabulary services for SKOS (multiple) resources. For instance, Loterre, Finto, SHOC Vocabs services, BARTOC (original version) are services based on this technology.² Since 2013, the BioPortal ontology repository [3]

¹ Respectively: <http://vocab.nerc.ac.uk> ; <https://vocabularies.cessda.eu> ; <https://opentheso.humanum.fr>.

² <https://www.loterre.fr> ; <https://finto.fi> ; <https://vocabs.sshopencloud.eu/vocabularies> ; <https://skosmos.bartoc.org>.

offers basic support for SKOS and now hosts around one hundred of such resources. However, because SKOS was not really adopted in biomedicine, we have not seen a full support of this format in BioPortal.³

Handling SKOS in BioPortal and in the generic version of the OntoPortal technology provided by the OntoPortal Alliance (<https://ontoportals.org>) being limited, we wanted to enhance this support especially in the context of AgroPortal –a reference vocabulary and ontology repository for agri-food [4] based on the OntoPortal technology– as SKOS is a representation language that has been widely adopted in agri-food. The vision was to also extend this to other projects relying on the OntoPortal technology such as EcoPortal (ecology/biodiversity) and the coming EarthPortal (earth sciences).

2 Related work and objectives

One limitation of the BioPortal technology was not supporting schemes, collections and SKOS-XL. The portal also is not multilingual [5] and not capable of integrating mappings between semantic artefacts declared with SKOS mappings properties in the multiple mapping services offered by the technology. The main strength of BioPortal is really to offer a wide range of services (from indexing to browsing, versioning, visualizing, annotation, recommendation, mappings, community notes, etc.) for various types of semantic artefacts, but because the portal was originally developed “for ontologies” with OWL at its core, it was not trivial to make it more SKOS compliant. The richness of supporting multiple formats in addition to OWL (e.g., OBO, UMLS RRF and SKOS) came with some limitations in how these formats were supported. For instance, BioPortal is built to handle classes (`owl:Class`) and instances but was not necessarily ready to support resources in SKOS where this distinction does not exist (everything is an instance of `skos:Concept`).

On its side, the SKOSMOS technology, certainly the most advanced for browsing/using SKOS resources at this time supports schemes, collections and SKOS-XL. Plus, its multilingual and the interface simply displays SKOS mappings (but no other mapping services are supported by SKOSMOS). The main strength of SKOSMOS is definitely to support SKOS very well –as it is design for this language– but it comes with the biggest limitation that it is impossible to use SKOSMOS to store and serve semantic artefacts in various representation languages, typically OWL. In some SKOSMOS installations e.g., Loterre, additional possibilities to browse concepts by temporality were added (based on dates metadata at the concept level). This is especially helpful when knowledge resources editors –such as VocBench does– directly incorporate `det:created` or `det:modified` properties when concept are edited.

One big limitation (in BioPortal but also SKOSMOS) that we wanted to address was the capability to browse concepts in an intuitive and easy manner when thesauri would make an extensive use of both schemes and collections. Indeed, there was no innovative way to browse concepts taking into accounts the multiples schemes and collections constructs that are inherently at the core of SKOS.

In the SKOS specification, a scheme (`skos:ConceptScheme`) is “an aggregation of one or more SKOS concepts” and is typically used to group concepts together with a specific hierarchy (integration to the scheme is done at the concept level with the property `skos:inScheme`). Typically, the thesaurus itself is a concept scheme but it can also include sub-schemes that would be only part of the thesaurus that one could browse and interact with as

³ See for details: <https://www.bioontology.org/wiki/SKOSupport>

with the main thesaurus. SKOS collections “are labeled and/or ordered groups of SKOS concepts” and are typically used to group concepts that “share something in common” (integration to the collection is done at the collection level with the `skos:member` property). From these constructs and other aspects in the SKOS specification, several questions raised:

- How to distinguish the main concept scheme from other sub-schemes in a unique SKOS source file? This is absent from the SKOS specification but is required typically to extract the right metadata and to distinguish schemes.
- How to display the hierarchies of any scheme independently or mixed? Considering some hierarchy paths could go through concepts that are not necessarily included in the scheme(s) being browsed.
- How to display concepts in collections (especially non ordered one)? The typical alphabetical approach (e.g., the approach proposed by SKOSMOS) is not necessarily sufficient when a collection contains a large number of members.
- How to browse concepts in a thesaurus (or even classes in an OWL ontology) ordered by their dates of creation/modification? So that someone can have a quick look at the temporal evolution of the content of a semantic artefact.
- How to include SKOS mappings in the “mapping repository” service offered by ontology repositories?

We wanted to address these needs while staying as close as possible to the official SKOS specification and not developing any extension or new constructs. Our vision was really to stay as generic and standard as possible. For instance, Loterre has addressed the issue of browsing multiple collections by relying on SKOS resources that would use constructs from SKOS extensions such as ISO-THES or UNESKOS.⁴

3 SKOS browsing in AgroPortal

For SKOS resources, we have completed the AgroPortal model and API to include all the objects in the SKOS specification such as `skos:ConceptScheme`, `skos:Collection`, and `skosxl:Label`.⁵ We have created new user interfaces to visualize these objects in AgroPortal and re-invented how to browse concepts.⁶ In summary, users can now:

- Visualize the list of schemes defined in a SKOS resource.⁷ AgroPortal will also distinguish one main scheme from the others. We recommend developers to use the SKOS resource URI to identify the “main” `skos:ConceptScheme` and also to type it with `owl:Ontology` to allow us to distinguish it from the other schemes and extract metadata.
- Visualize the lists of collections defined in a SKOS resource and count the numbers of concept members. (**Fig. 1**).
- Jump to concepts with a autocomplete search box (pre-existing) and visualize concepts with 3 different approaches:



Hierarchical, by selecting the scheme(s) to browse. The hierarchy tree will always refresh to display only the selected scheme(s) and selected collection members

⁴ Respectively <http://purl.org/iso25964/skos-thes> and <https://skos.um.es/TR/uneskos>

⁵ See AgroPortal’s API documentation: <https://data.agroportal.lirmm.fr/documentation>.

⁶ The following features are illustrated in the Supplementary Figures section or can be consulted in the AgroPortal release notes (<https://doc.jonquetlab.lirmm.fr/share/d77f1b99-60fb-4fdb-982b-927204f1c1e3>).

⁷ Good resources in AgroPortal to illustrate are BIODIVTHES, AGROVOC or INRAETHES (available very soon).

will be displayed with specific colors (**Fig. 2**). Plus, concepts traversed by a hierarchy path that are not necessarily included in the selected scheme(s) are displayed in italic.



By collections, showing alphabetically all the concept members of a selected collection.



Temporal, showing all concepts/classes sorted by date of creation or modification (using dcterms properties). This feature is also available for OWL ontologies (if the source file provides the information).

- Visualize SKOS-XL labels to provide information about a reified label (**Fig. 3**).
- See for every concept, which scheme and collection the concepts belong to. This is trivial for schemes (as defined in the resources by `skos:inScheme` property) but needs to be computed for collections (we use the property `uneskos:memberOf` to encode this additional information in our triple-store).
- See inside the class/concept mappings tab (**Fig. 4**) the SKOS mappings explicitly defined in the source file, aside of the other ones in the mapping repository (automatically generated e.g., `SAME_URI`, `LOOM` or uploaded by users e.g., `REST`). They are retrieved in the triple-store and also available in the API as any other mappings.

4 Conclusion and perspectives

A complete technical trace of what was implemented is available in AgroPortal documentation (releases 2.3 and 2.4).⁸ We are currently working with the OntoPortal Alliance to include these features in the main code branch so they become available for other semantic artefact catalogues based on this technology. All together, these features make AgroPortal a fully SKOS compliant platform which is certainly one of the most advanced SKOS browsing technology. Plus, it does that in a platform that can seamlessly handle a large variety of semantic artefacts (typically OWL and SKOS) enabling use cases where multiple knowledge resources are needed. One big step now for AgroPortal is to become multilingual (planned for 2023) as this is required in various communities, especially the ones developing (SKOS) thesauri.

Acknowledgement

This work has been supported by the *Data to Knowledge in Agronomy and Biodiversity* project (D2KAB – www.d2kab.org – ANR-18-CE23-0017) and the NUMEV Labex (grant ANR-10-LABX-20).

References

1. Miles, A., Bechhofer, S.: SKOS simple knowledge organization system reference. W3C Recom. 18, (2009).
2. Suominen, O., Ylikotila, H., Pessala, S., Lappalainen, M., Frosterus, M., Tuominen, J., Baker, T., Caracciolo, C., Retterath, A.: Publishing SKOS vocabularies with Skosmos. Manuscript submitted for review. (2015).
3. Noy, N.F., Shah, N.H., Whetzel, P.L., (...), Musen, M.A.: BioPortal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Res.* 37, 170–173 (2009). <https://doi.org/10.1093/nar/gkp440>.
4. Jonquet, C., Toulet, A., Arnaud, (...), Larmande, P.: AgroPortal: A vocabulary and ontology repository for agronomy. *Comput Electron Agric.* 144, 126–143 (2018). <https://doi.org/10.1016/j.compag.2017.10.012>.
5. Jonquet, C., Emonet, V., Musen, M.A.: Roadmap for a multilingual BioPortal. 4th Workshop on the Multilingual Semantic Web, MSW4'15. pp. 15–26. , Portoroz, Slovenia (2015).

⁸ <https://github.com/agroportal/project-management>. Corresponding GitHub issue for traceability and reproducibility is: <https://github.com/agroportal/project-management/issues/262>.

Supplementary Figures

Biodiversity Thesaurus
Last updated: September 27, 2022

Summary Concepts Properties Notes Mappings Schemes Collections

Activity

- Activity / Behavior / Society
- Agent / Factor
- Agronomy sciences / Agriculture
- Analysis / Measure
- Animal
- Animal ecology
- Aquatic ecology
- Bacteria
- Behavior
- Biological compound
- Biological phenomenon / Physiological phenomenon
- Biological sciences
- Biological sciences_Miscellaneous
- Chemical compound
- Chemical compound_Miscellaneous
- Classification / Nomenclature
- Community composition
- Discipline
- Document title / Work title
- Earth and universe sciences
- Ecology
- Ecology_Miscellaneous
- Economic concept or system
- Ecosystem function
- Ecosystem structure
- Environment
- Environment / Pollution / Sustainable development
- Food
- Food industry / Food science
- Forestry / Silviculture
- Function
- Fungi
- Gene / Genome
- Genetic composition
- Genetics

ID	http://data.loterre.fr/ark:/67375/BLH-ACT1
Preferred Name	Activity
Members count	76
Type	http://www.w3.org/2004/02/skos/core#Collection
comment	Number of members : 76 Nombre de membres : 76
modified	2022-02-11
superGroup	http://data.loterre.fr/ark:/67375/BLH-ACSO
definition	Number of members : 76 Nombre de membres : 76
prefLabel	Activity Activité
inScheme	http://data.loterre.fr/ark:/67375/BLH

Fig. 1. Collections defined in the Biodiversity Thesaurus (<https://agroportal.lirmm.fr/ontologies/BIODIVTHES/?p=collections>).

Jump to Filter

Schemes

BLH (main) x

Collections

Animal x Analysis / Measure x Bacteria x Behavior x
Discipline x Animal ecology x

- applied ecology
- biological substance
- cultural heritage
- ecological parameter
- ecological process
- ecological system
- ecology
 - animal ecology
 - behavioral ecology
 - sensory ecology
 - zoogeography
 - autecology
 - bioclimatology
 - chemical ecology
 - demeology
 - disease ecology
 - evolutionary ecology
 - fire ecology
 - forest ecology
 - functional ecology
 - integrative ecology

Details Visualization Notes (0) Concepts Mappings (1) Access Concepts JSON

ID	http://data.loterre.fr/ark:/67375/BLH-WP8V9ZL-L
Preferred Name	behavioral ecology
Synonyms	ecoethology behavioural ecology
Member of	Discipline ↗ Animal ecology ↗
In Schemes	http://data.loterre.fr/ark:/67375/BLH
Type	http://www.w3.org/2004/02/skos/core#Concept
memberOf	http://data.loterre.fr/ark:/67375/BLH-DISC http://data.loterre.fr/ark:/67375/BLH-EKA
altLabel	ecoethologie behavioural ecology
closeMatch	http://www.eionet.europa.eu/gemet/concept/2984
prefLabel	behavioral ecology écologie comportementale
created	2021-10-25
broader	animal ecology ↗

Fig. 2. Hierarchy of the main scheme in BIODIVTHES while highlighting concepts in 6 different collections with colors.

The screenshot shows a modal window with the following details for the SKOS-XL label 'gaeabionta':

ID	http://aims.fao.org/aos/agrovoc/xl_uk_58e9db08
Preferred Name	gaeabionta
Type	http://www.w3.org/2008/05/skos-xl#Label

literalForm	gaeabionta
notation	58e9db08
created	2017-10-13T16:14:02Z
modified	2017-10-17T13:34:49Z
inDataset	http://aims.fao.org/aos/agrovoc/void.ttl#Agrovoc

Below the modal, a list of related concepts is visible, including 'life on earth', 'Lebewesen', 'gaeabionta', and 'пріонні білки'.

Fig. 3. The SKOS-XL label “gaeabionta” (synonym of “organisms”) in AGROVOC.

The screenshot displays the 'Concepts Mappings (16)' section of the BIODIVTHES interface. On the left, a tree view shows the hierarchy of concepts, with 'hydrosphere' selected. The main area shows a table of mappings:

MAPPING TO	ONTOLOGY	RELATIONS	SOURCE	TYPE	ACTIONS
hydrosphere	ANAEETHES		SAME_URI	Internal	
4124	http://www.eionet.europa.eu/gemet/concept		SKOS-EXACT_MATCH	External	
hydrosphere	GEMET	skos:exactmatch	REST ANAEETHES	Internal	
hydrosphere	GEMET		LOOM	Internal	

Fig. 4. Mappings from different sources (including declared in the source file with `skos:exact-match`) for the concept “hydrosphere” in BIODIVTHES.