EARTh: an Environmental Application Reference Thesaurus in the Linked Open Data Cloud

Editor(s): Pascal Hitzler, Kno.e.sis Center, Wright State University, Dayton, Ohio, USA; Krzysztof Janowicz, University of California, Santa Barbara, USA
Solicited review(s): Natasha Noy, Stanford Center for Biomedical Informatics Research (BMIR), Stanford, California, USA; Tomi Kauppinen, Aalto University School of Science, Finland; Marta Sabou, MODUL University Vienna, Austria.

R. Albertoni[a,b*], M. De Martino[b], S. Di Franco [c], V. De Santis [c] and P. Plini[c]
[a]*Ontology Engineering Group. Dpto. de Inteligencia Artificial Facultad de Informática, Universidad Politécnica de Madrid 28660, Boadilla del Monte, Madrid, Spain, ralbertoni@fi.upm.es*
[b]*CNR-IMATI, Via De Marini, 6, 16149 Genova, Italy, {albertoni, demartino}@ge.imati.cnr.it*
[c]*CNR-IIA-EKOLab - Environmental Knowledge Organization Laboratory, Area della Ricerca di Roma 1, Via Salaria Km 29,300 C.P. 10,I-00016 Monterotondo stazione RM, {difranco, vds, plini} @iia.cnr.it*

Abstract. The paper aims at providing a description of EARTh, the Environmental Application Reference Thesaurus. It represents a general-purpose thesaurus for the environment, which has been published as a SKOS dataset in the Linked Open Data cloud. It promises to become a core tool for indexing and discovery environmental resources by refining and extending GEMET, which is considered the de facto standard when speaking of general-purpose thesaurus for the environment in Europe, besides it has been interlinked to popular LOD datasets as AGROVOC, EUROVOC, DBPEDIA and UMTHES. The paper illustrates the main characteristics of EARTh as a guide to its usage. It clarifies (i) the methodology adopted to define the EARTh content; (ii) the design and technological choices made when publishing EARTh as Linked Data; (iii) the information pertaining to its access and maintenance. Descriptions of EARTh applications and future relevance are highlighted.

Keywords: SKOS, Linked Data, EARTh, Thesaurus, Environment

1. Introduction

Although different directives (e.g. INSPIRE) and policy communications (e.g. SEIS) have been launched at European-scale with the objective of improving the management of heterogeneous environmental data sources, an effective sharing of these resources is still part of the desiderata due to the intrinsic multicultural and multilingual nature of the environmental domain.

Thesauri are widely employed as common ground enabling communication among the different communities working in environment-related domains: they allow users to share and agree upon scientific/technical terms in the target domain and to express them in multiple languages. In the recent years several controlled vocabularies and thesauri have been deployed by different communities having a large spectrum of competencies. They have been created embodying different points of view and based on different ways of conceptualization. Their development reflects different scopes and implies quite a range of levels of abstraction and detail.

Nowadays networked information access to het-

[*] Corresponding author. E-mail: albertoni@ge.imati.cnr.it

erogeneous environmental data sources requires interoperability of these controlled vocabularies [1]. The Linked Data publishing paradigm [2] jointly with Simple Knowledge Organization System (SKOS) [13] provides a promising framework to face with the aforementioned problems: it allows to represent and publish distinct thesauri and their interlinks as a whole enabling a joint exploitation of them.

This paper presents the latest release of EARTh, the Environmental Application Reference Thesaurus (ver. Linked Data 1.4) that takes advantage of this framework providing a SKOS dataset available in the Linked Open Data (LOD) Cloud.

Compared to other environmental thesauri available as Linked Data as AGROVOC[2], EUNIS[3], Geological Survey of Austria (GBA) Thesaurus [4], EARTh[5] provides a more general purpose and thematically neutral terminological support. Compared to the GEneral Multilingual Environmental Thesaurus (GEMET)[6], namely the *de facto* general purpose thesaurus standard, EARTh provides a minor multilingual support, but it extends GEMET with more than 9000 concepts and revises the GEMET concept hierarchy. Being EARTh one of the largest general purpose and structured environmental terminological resources available in the LOD cloud, it aims at providing a bridge for the integration of other terminological resources dealing with environmental topics. It already includes more than 12000 links towards thesauri such as GEMET, AGROVOC, EUROVOC and UMTHES enabling in the traditional thesaurus-based indexing of digital resources, as well as the use of digital resources across multi-thesauri applications [12]. Besides, further interlinkings will be provided as part of the activity committed in the European funded project eENVplus (CIP-ICT-PSP grant no. 325232).

The remainder of this paper is organized as follows. Section 2 describes EARTh in terms of its content, the methodology followed and the extension of GEMET. Section 3 describes how EARTh has been published in the LOD cloud. Section 4 describes the dataset applications. Future steps and conclusions are drawn in Section 5.

2. EARTh thesaurus

EARTh is a project run since 2001 by CNR-IIA-EKOLab aiming at creating a new thesaurus for the environment. It extends the GEMET content and revises its categorical and thematic structure.

Originally GEMET, developed by CNR-IIA-EKOLab and by German Federal Environmental Agency within an international consortium, was intended to be used as an indexing, retrieval and control tool for the European Topic Centre on Catalogue of Data Sources (ETC/CDS) and the European Environment Agency (EEA). The basic idea for the development of GEMET was to use the best of the currently available excellent multilingual thesauri, in order to save time, energy and funds. GEMET was conceived as a “general” thesaurus, aimed to define a common general language, a core of general terminology for the environment. Specific thesauri and descriptor systems (e.g., on Nature Conservation, Wastes, Energy) have been excluded from the first step of development of the thesaurus and have been taken into account only for their structure and upper level terminology.

Since 2001, CNR-IIA-EKOLab performed an overall checking of GEMET in order to improve both its content and its structure. In particular the following activities have been undertaken:

- quality assessment of GEMET structure and content towards ISO standards on mono- and multilingual thesauri;
- assessment of English concept representation vs. source language(s);
- deletion of incorrect terms and removal of about 1000 terms potentially useful for specific lists, such as name of plants, animals, minerals, etc.;
- updating the content with new terms (e.g., land management strategies, pigmy forest, cryodiversity) and extension of the system of non-descriptors;
- management of the correspondence of the terms in British and American English (e.g. sulphur hexafluoride/sulfur hexafluoride);
- revision of the thematic structure and development of a new categorical/hierarchical setup (e.g., Entities, Attributes, Dynamic Aspects, Dimensions) to emphasize the different functions of hierarchy in comparison to themes;
- extension of the horizontal and the vertical relations system;
- representation of the accessory elements: singular and plural forms (e.g., biological index/biological

[2] http://aims.fao.org/website/AGROVOC-Thesaurus/sub
[3] http://eunis.eea.europa.eu/
[4] http://thedatahub.org/dataset/geological-survey-of-austria-thesaurus
[5] http://thedatahub.org/dataset/environmental-applications-reference-thesaurus
[6] http://www.eionet.europa.eu/gemet

indexes), alternate terms (e.g., deoxyribonucleic acid/desoxyribonucleic acid), etc.

EARTh is based on a multidimensional classificatory and semantic model [10]. The vertical structure of the thesaurus is built through a deductive (top-down)–inductive (bottom-up) approach. It is basically mono-hierarchical. It is developed according to a tree semantic model and is based on a system of categories. The first level of categories corresponds to entities, attributes, dynamic aspects, and dimensions. The vertical structure analyses the primary meaning of the terms and places them in the classificatory-hierarchical tree aiming to orientate the users towards the most "essential" characteristics of terms' semantics.

Besides from GEMET, EARTh terminological content is derived from various mono and multilingual sources of controlled environmental terminologies such as UN Environment and Development [6], Italian Thesaurus of Earth Sciences [3], Inland Water terminology (derived from EDEN-IW project), Emergency Management Terms Thesaurus [7] and other terminologies collected from reference documents in specific fields or coming from the daily research activity. EARTh currently contains more than 15.000 terms in English and Italian. Its content includes terms related to earth structure (lithosphere, hydrosphere and atmosphere) and to natural sciences (biosphere), terms dealing with the human society, activities and products (anthroposphere, built environment). In addition the terminology covers physical, chemical, natural and social processes, properties, effects, events, health and safety, productive sectors, data, parameters, methods and techniques. The content is constantly updated following the evolution of environmental terminology.

3. Publishing EARTh as Linked Data

From the technological point of view, the latest release of EARTh has been published combining Virtuoso Open-Source Edition 6.1.6[8] and Pubby[9]. In order to publish EARTh as Linked Data we have adopted the following Linked Data patterns [5]:

- Natural Key for identifier: the internal identifiers for EARTh concepts are assumed as Natural Keys in order to keep coherence with EARTh's previous (not linked data) releases and usages;
- Label Everything: every concept has its English human-readable name expressed as rdfs:label. So human-readable names can be exploited debugging queries and exploring EARTh;
- Preferred Label: every concept has its preferred label expressed as skos:prefLabel. Both English and Italian lexical representations are provided;
- Materialize Inferences: SKOS entailments have been materialized to support clients with limited processing power. In particular, considering the SKOS entailments, entailments indicated in [13] as S7, S8, S11, S17, S21, S22, S23, S25, S26, S39, S40, S41, S42, S43 have been materialized. Besides, rdfs:labels are obtained as materialized inferences of English skos:prefLabel;
- Equivalence Links: more than 4000 skos:exactMatch are provided to indicate equivalent URIs between EARTh and GEMET. That has been possible because EARTh is a significant extension of GEMET and explicit references to the GEMET ID have been maintained for the concepts shared with GEMET. Further equivalences have been created by working out the transitive closure on GEMET's skos:exactMatch. So that, we have been able to import the GEMET's outgoing links to AGROVOC, EUROVOC, DBpedia and UMTHES in EARTh. Unfortunately, links obtained by this procedure only pertain to the subset of concepts that EARTh shares with GEMET. In order to complement that set and find out a more complete connection among EARTh and GEMET's linked datasets, a two-steps process has been put in place: firstly, SILK[10] has been applied to discover new links, then the SILK results have been validated by the expert members of CNR-IIA-EKOLab in order to verify the accuracy of the links and to identify the most suitable types of interlinking property (i.e., skos:exactMatch or skos:closeMatch). The joint exploitation of skos:exactMatch transitive closure and the manually validated SILK link discovery have almost triplicated the number of outgoing links available with respect to the previous EARTh releases. In particular, about 7171 links have been discovered relying on the transitive closure, 465 have been generated deploying SILK. The result is quite reliable: only 3 links have been rejected during the manual validation and only 17 have been switched to skos:closeMatch. That reliability is a consequence

[7] http://library.ema.gov.au/emathesaurus/

[8] http://sourceforge.net/projects/virtuoso/files/virtuoso/6.1.6/

[9] http://wifo5-03.informatik.uni-mannheim.de/pubby/

[10] http://wifo5-03.informatik.uni-mannheim.de/bizer/silk/

of the very selective procedure we have put in place when using SILK: we have been assessing as linkable those concepts whose skos:prefLabel and skos:altLabel had a levenshtein similarity greater than 0.95, preferring a minor number of links with a greater confidence that many links but minor confidence. This new release paves the way for a combined exploitation of EARTh with GEMET, AGROVOC, EUROVOC, DBpedia and UMTHES (about the 33% of EARTh concepts have a link to other thesauri) enabling EARTh adopters in taking advantage of their respective strengths and complementarities.

3.1. EARTh in the LOD cloud

Since the late 2011, EARTh is included in the LOD Cloud. EARTh content is accessible through (i) HTTP dereferenceable URIs [11]; (ii) RDF/XML dump [12]; (iii) SPARQL end point [13]. Moreover EARTh is part of a framework that includes other SKOS linked datasets [4], thus accessing EARTh concepts via SPARQL end point requires to make a good use of EARTh's GRAPH pattern[14], as shown below:

Query 1: SPARQL making use of GRAPH pattern

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT DISTINCT * WHERE {
GRAPH
<http://linkeddata.ge.imati.cnr.it:2020/resource/EARTh/> {
   ?s skos:prefLabel ?o }
}
```

It is worth noting that EARTh concepts can work as a bridge between different thesauri. For example, Query 2 retrieves some EARTh concepts pertaining to "meteorological station". As depicted in Fig. 1, EARTh "meteorological station" concept and its related concepts connect DBPEDIA and AGROVOC to GEMET and UMTHES.

Query 2: Meteorological station's linked entities

```
PREFIX EARTh:
<http://linkeddata.ge.imati.cnr.it:2020/resource/EARTh/>
PREFIX DBPEDIA: <http://dbpedia.org/resource/Category>
PREFIX GEMET: <http://www.eionet.europa.eu/gemet/concept/>
PREFIX AGROVOC: <http://aims.fao.org/aos/agrovoc/>
PREFIX UMTHES: <http://data.uba.de/umt/>
SELECT DISTINCT  * WHERE {
  EARTh:46920 ?property ?hasValue
  OPTIONAL {
   {?hasValue skos:exactMatch ?RConceptMatches } UNION
   {?hasValue skos:closeMatch ?RConceptMatches }}}
```

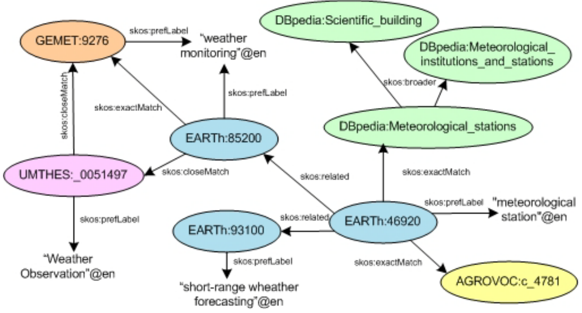

Fig. 1. Meteorological station's linked entities.

Query 3: Third-party thesaurus concepts related via EARTh

```
SELECT ?RConceptMatches1 ?RConceptMatches WHERE {
GRAPH <http://linkeddata.ge.imati.cnr.it:2020/resource/EARTh/>
  {?x ?property ?y}
  OPTIONAL {?x skos:mappingRelation ?RConceptMatches }
  OPTIONAL {?y skos:mappingRelation ?RConceptMatches1}
  FILTER (?RConceptMatches!=?RConceptMatches1    ).
  FILTER(SUBSTR(STR(?RConceptMatches1), 1,20)
!=SUBSTR(STR(?RConceptMatches), 1,20))}
```

Query 3 shows all the couples of concepts from distinct thesauri that can be put in relation by using EARTh as "bridge". EARTh bridges around 9110 couples considering skos:exactMatch and 34000 couples considering both skos:exactMatch and skos:closeMatch. Compared with GEMET bridging potential (respectively 3006 and 28758), this gives an insight into EARTh's higher potential when integrating other terminological resources.

Table 1 provides statistics pertaining to the number of SKOS concepts and the availability of properties for those concepts. The first column of the table provides information about the number of skos:Concept and those having at least one occurrence of the indicated SKOS relations. For example, the first row indicates that 14351 skos:Concept are available, the second row indicates that 14350 concepts have a skos:inScheme property. The second column shows their SKOS lexical representations. For example, the first row shows that 14350 of them have a skos:prefLabel in English and 14002 in Italian. Table 2 shows statistics about EARTh outgoing interlinks respectively towards GEMET, AGROVOC, UMTHES, etc.

Further details pertaining to linkset and accessibility are available in VOID description [15] and Hub[16]. EARTh is available under by-nc-nd creative

[11] http://linkeddata.ge.imati.cnr.it:2020/directory/EARTh
[12] http://purl.oclc.org/net/DumpEarthRDF
[13] http://linkeddata.ge.imati.cnr.it:8890/sparql
[14] http://linkeddata.ge.imati.cnr.it:2020/resource/EARTh/
[15] http://purl.org/NET/EARThVOID
[16] http://thedatahub.org/dataset/environmental-applications-reference-thesaurus

commons licence[17], which grants the right to copy, distribute and transmit it for non-commercial purposes, but implies explicit attribution of work and forbids derived works.

Table 1

Statistics about EARTh SKOS concepts and their properties availability (materialized properties are not listed being easily obtainable from the below).

Property	#	Property	#
skos:concept	14351	**skos:prefLabel**	14350 **(en)** 13813 **(it)**
skos:inScheme	14350	**rdfs:label**	14350 **(en)** 13813 **(it)**
skos:broader or skos:narrower	11664	**skos:definition**	6362 **(en)** 5383 **(it)**
skos:related	4084	**skos:altLabel**	1004 **(en)** 743 **(it)**

Table 2

Statistics about EARTh outgoing interlinks

Interlinks	#	Interlinks	#
skos:exactMatch to GEMET	4365	**skos:exactMatch to EUROVOC**	1337
skos:exactMatch **skos:closeMatch to AGROVOC**	1436 1456	**skos:exactMatch** **skos:closeMatch to DBpedia**	98 1861
skos:closeMatch to UMTHES	2965		

4. EARTh application/relevance

EARTh is exploited in different projects to support both indexing and retrieval of environmental resources. Table 3 summarises EARTh applications at national and international level distinguishing among the different strands in which EARTh is adopted: LOD, RDF or other depending if the applications use EARTh exploiting its published linked data version, its SKOS/RDF dump, or an old fashion access to its relational database version.

At national level, EARTh is deployed by the Italian Environmental Agency (ISPRA) in its portal for Indexing and Networking of Documents on Environmental Knowledge Sharing (INDEKS)[18].

At international level, EARTh is object of a continuous collaborative exchange with FAO Term Portal[19]. EARTh SKOS/RDF dump is often employed in combination with Geonetwork as a terminological source to compile ISO 19115 metadata: for example, it is employed in Geonetwork SDI on Mercury emissions and Geonetwork SDI KnowSeas.

Table 3

EARTh applications at national and international level

Applications	LOD	RDF Dump	Other
ISPRA INDEKS		X	
FAO Term Portal (FAOTERM)			X
SDI Geonetwork on Mercury emissions		X	
SDI Geonetwork KnowSeas		X	
eENVplus (EU Project)	X		
Nature-SDI (EU Project)	X		
GS-Soil (EU Project)	X		
EnvEurope (EU Project)	X		
ExpeER (EU Project)	X		

EARTh availability as linked dataset has raised its relevance at international level, especially in EU funded projects where different controlled vocabularies for specific environmental data themes are employed to support data sharing within Spatial Data Infrastructures, in particular their interoperability is desirable to support the semantic harmonization/integration of heterogeneous data. Specifically the EU projects are:

- NatureSDIplus[20] and eENVplus[21]: EARTh is employed as backbone thesaurus for a thesaurus framework for Nature Conservation [4].
- GS-SOIL[22]: the thesaurus for Soil SoilThes has been created with a domain specific thesaurus with direct outgoing skos:exactMatch links to EARTh.
- EnvEurope[23] and ExpeER: the common controlled vocabulary EnvThes for long term ecological research and monitoring data themes provides explicit outgoing links to EARTh.

The new release of EARTh has empowered its “bridging” nature with respect to other well-known thesauri, and then we expect that EARTh will be even more largely adopted, not only to support finding and understanding environmental data/metadata, but to enable semantic interoperability of data and metadata within the data infrastructure and its services.

In a scenario of a thesaurus-enhanced search for re-

[17] http://creativecommons.org/licenses/by-nc-nd/3.0/
[18] http://www.indeks.isprambiente.it/INDEKS/public/welcome.do
[19] http://www.fao.org/termportal/partnerships1/en/
[20] http://www.nature-sdi.eu/
[21] http://www.eenvplus.eu/
[22] http://gssoil-portal.eu/ingrid-portal/
[23] http://www.enveurope.eu/

source, EARTh can be exploited in analogy with use cases developed in other domains (e.g., Social Sciences, Economics, Medicine) [7, 8, 11, 12, 14, 15]. In particular, it can be deployed for:

- Query expansion and reformulation to supplement additional terms to the original query in order to improve the retrieval performance [7, 8]. Reasoning over thesaurus semantic relationships supports the seeker in finding alternative concepts and expanding or reformulating his queries by automatically suggested term refinements. In particular, links between concepts of different thesauri can be used automatically to expand the search for terms inside the other non-EARTh thesaurus taking advantages of their complementarities in term of domain specificity and multilingualism.
- Index interoperability and Facet Browsing. The thesaurus interlinking eases the heterogeneity coming from the usage of different thesauri in order to index environmental resources. It makes accessible and connectable the content of traditional databases for the applications of the Semantic Web, i.e. as Linked Open Data [11, 15]. Crosswalks between vocabularies can play a relevant role in interoperability, because they serve as a bridging hub for the interlinking of different published and indexed data sets [12].

5. Conclusion and future work

The paper illustrates the main characteristics of the linked dataset EARTh, an environmental thesaurus that promises to become pivotal for environmental data sharing serving as bridging hub for different environmental thesauri. EARTh content continuously evolves as a result of CNR-IIA-EKOLab's research activity, whist EARTh Linked Data releases are provided once a year by CNR-IMATI.

On-going and future activity includes improvement in terms of EARTh content as well as its Linked Data publication. Concerning the content, an overall revision of the thesaurus structure and content is currently undergoing as consequence the recent publication of ISO 25964-1:2011 [9]. The linked data version of EARTh will be maintained and updated in the context of the EU project eENVplus, which aims at establishing semantic interoperability between different existing thesauri for the environment. In particular, the activities planned within eENVplus project are: (i) RT properties, which have been indistinctly mapped into skos:related in order to avoid the adoption of user-defined RDF vocabularies, will be differentiated as in the original version of EARTh; (ii) EARTh connection with other environment-related thesauri will be strengthened providing links to other thesauri (e.g., NALT Agricultural Thesaurus).

Acknowledgements

This activity has been carried out within the EU funded project NatureSDIplus (ECP-2007-GEO-317007) and eENVplus (CIP-ICT-PSP grant No. 325232). Part of Albertoni's contribution to this work has been carried out during the tenure of an ERCIM "Alain Bensoussan" Fellowship Programme supported by the Marie Curie Co-funding of Regional, National and International Programmes (COFUND) of the European Commission.

References

[1] T. Bandholtz, J. Fock, R. Legat, M. Nagy, K. Schleidt and P. Plini, Shared Terminology for the Shared Environmental Information System. In: Environmental Informatics and Industrial Environmental Protection: Concepts, Methods and Tools, 23rd International Conference on Informatics for Environmental Protection, vol. 1, Shaker, pp. 123-127, Aachen, 2009.

[2] T. Berners-Lee, Design Issues: Linked Data, 2006, http://www.w3.org/DesignIssues/LinkedData.html.

[3] A. Carusone and L. Olivetta (eds.), Italian Thesaurus of Earth Sciences, APAT, Rome, 2006.

[4] M. De Martino and R. Albertoni, A multilingual/multicultural semantic-based approach to improve Data Sharing in a SDI for Nature Conservation, International Journal of Spatial Data Infrastructures Research, vol.6, ISSN 1725-0463, pp. 206-233, 2011.

[5] L. Dodds and I. Davis, Linked Data Patterns. A pattern catalogue for modelling, publishing and consuming Linked Data, 2012, http://patterns.dataincubator.org/book/.

[6] Environment and Development United Nations Terminology Bulletin No 344, 1992.

[7] B. Haslhofer, F. Martins, J. Magalhães, Using SKOS vocabularies for improving web search. In: Proceedings of the 22nd international conference on World Wide Web companion. pp. 1253–1258, ACM 2013.

[8] D. Hienert, P. Schaer, J. Schaible, P. Mayr, A novel combined term suggestion service for domain-specific digital libraries. In: S. Gradmann, F. Borri, C. Meghini, and H. Schuldt (eds), Proceedings of Research and Advanced Technology for Digital Libraries – International Conference on Theory and Practice of Digital Libraries, TPDL 2011, Berlin, Germany, September 26-28, 2011, volume 6966 of Lecture Notes in Computer Science, pp. 192-203. Springer Berlin / Heidelberg, 2011.

[9] ISO 25964-1 Information and documentation - Thesauri and interoperability with other vocabularies - Part 1: Thesauri for information retrieval, 2011.

[10] F. Mazzocchi, B. De Santis, M. Tiberi and P. Plini, Relational Semantics in thesauri: Some Remarks at Theoretical and Prac-

tical Levels, J. Knowledge Organization, vol. 34 (4), pp. 197-214, 2007.

[11] J. Neubert, Bringing the "Thesaurus for Economics" on to the Web of Linked Data, In: Proceedings of the Linked Data on the Web Workshop (LDOW2009), Madrid, Spain, April 20, 2009, CEUR Workshop Proceedings, ISSN 1613-0073, online CEUR-WS.org/Vol-538/ldow2009_paper1.pdf.

[12] M. Philipp, and V. Petras. Cross-concordances: terminology mapping and its effectiveness for information retrieval. In: World Library and Information Congress: 74th IFLA General Conference and Council, Québec, Canada. pp 10-14, 2008.

[13] SKOS, Simple Knowledge Organization System Reference, W3C Recommendation, 2009, http://www.w3.org/TR/skos-reference.

[14] J. Tuominen, M. Frosterus, K. Viljanen, and E. Hyvönen, ONKI-SKOS — Publishing and Utilizing Thesauri in the Semantic Web. AI and Machine Consciousness In: T. Raiko, P. Haikonen, and J. Väyrynen (eds.), Proceedings of the 13th Finnish Artificial Intelligence Conference STeP 2008, 2008. http://www.stes.fi/step2008/proceedings/step2008proceedings.pdf.

[15] B. Zapilko, J. Schaible, P. Mayr, B. Mathiak TheSoz: A SKOS Representation of the Thesaurus for the Social Sciences, Semantic Web, vol. 4 (3), pp. 257-263, 2013.