The original publication is available http://www.springerlink.com/.

Riccardo Albertoni, Monica De Martino, Paola Podestà,
Linkset Quality Assessment for the Thesaurus Framework LusTRE

MTSR 2016: Metadata and Semantics Research pp 27-39

Series: Volume 672 of the book series Communications in Computer and Information Science (CCIS)

Garoufallou E., Subirats Coll I., Stellato A., Greenberg J. (eds)
2016
ISBN: 978-3-319-49156-1
Doi: 10.1007/978-3-319-49157-8_3

Linkset Quality Assessment for the thesaurus framework LusTRE

Riccardo Albertoni, Monica De Martino, and Paola Podestà

Istituto di Matematica Applicata e Tecnologie Informatiche
Consiglio Nazionale delle Ricerche,
Via De Marini, 6, 16149 Genova, Italy
{albertoni,demartino,podesta}@ge.imati.cnr.it

Abstract. Recently a great number of controlled vocabularies (e.g., thesauri) covering several domains and shared by different communities, have been published and interlinked using the Linked Data paradigm. Remarkable efforts have been spent from data producers to make their thesauri compliant with Linked Data requirements both for the content encoding and for the connections (aka, linkset) with others thesauri. Also in our experience in the creation of the framework of multilingual linked thesauri for the environment (LusTRE), within the EU funded project eENVplus, the development of the interlinking among thesauri, have required significant efforts, thus, the evaluation of their quality in term of usefulness and enrichment of information became a critical issue. In this paper, to support our claim, we discuss the results of the quality evaluation of several linksets created in LusTRE. To this purpose, we consider two quality measures, the *average linkset reachability* and the *average linkset importing*, able to quantify the linkset-accessible information.

Keywords: Linkset quality, SKOS, Linked Data, environmental thesauri, metadata

1 Introduction

The continued expansion of the Web as a medium for the exchange of interoperable data and the sustained growth of the Linked Open Data Cloud[1], represent important factors for the Linked Data paradigm success. In this context, where data sharing and consumption are accessible to a large number of actors, the quality of the exposed data become one of the most critical issues, since as widely known and accepted, data is only worth its quality [16].

Linked Data paradigm [10] is based on two core characteristics: data sources and connections of information belonging to different sources through the linksets. Thus, not only data but also the connections among data are essential to keep the Linked Data promise to "evolve the current web data into a Global Data Space" [10]. In fact, through a link a consumer can navigate in a seamless way between objects belonging to different datasets, possibly of different domains,

[1] http://lod-cloud.net/

accessing to richer and more complete information than the data at hand. The quality of connections (in the following, *linkset quality*) should be as important as the quality of data, but, since now, the research on Linked Data quality has been mainly focused on datasets [16].

During the EU funded project eENVplus (CIP-ICT-PSP grant No. 325232) supporting the Infrastructure for Spatial Information in the European Community (INSPIRE)[2] directive implementation, we have developed a multilingual linked thesaurus framework for the environment (LusTRE)[3]. The framework aims to provide shared standard and scientific terms for a common understanding of environmental data among the different communities operating in different fields of the Environment. It supports better cross-domain and multilingual metadata compilation and information discovery. In order to construct the framework, we have analysed several environmental thesauri exposed on the Web according to the Simple Knowledge Organization System (SKOS)[4] Vocabulary [4]. Then, we have spent considerable efforts to publish in Linked Data some of them (ThiST, EARTh [3]) and to interlink each other GEMET, AGROVOC [8], EUROVOC, TheSoz, DBpedia, ThiST, EARTh; but, until now, there is no way to calculate the information reachable navigating the interlinks.

In this paper, we evaluate the quality of the connections created in LusTRE. We rely on the notion of dataset and linkset provided in the Vocabulary of Interlinked Datasets (VoID)[5]. Thus, a *dataset* is a set of Resource Description Framework (RDF)[6] triples published, maintained or aggregated by a single provider; a *linkset*, is a special kind of dataset containing only RDF links between two datasets, the *subject*, and the *object* of the linkset. We focus on the `skos:exactMatch` linksets (i.e., linkset composed only by `skos:exactMatch` mapping relation) developed in LusTRE and we consider two measures: the *average linkset importing*, based on a measure presented in [5], and a new measure, the *average linkset reachability*. Using the average linkset importing, we evaluate LusTRE linksets assessing the multilingualism enrichment in terms of new translated labels reachable in the object through a linkset. Average linkset importing can help to address the incomplete language coverage issue, that is, when `skos:prefLabel` and `skos:altLabel` are provided in all the expected languages only for a subset of the thesaurus concepts (in [14]), affecting many popular SKOS thesauri. Average linkset reachability evaluates the number of new concepts reached by crossing a linkset. It can be exploited to evaluate the potential enrichment of the space of concepts that can be browsed (aka, the thesaurus browsing space). The application of both measures to the linksets of LusTRE shows that, in average, linksets bring an effective enrichment in terms of multilingualism and browsing space.

[2] http://inspire.ec.europa.eu/
[3] http://linkeddata.ge.imati.cnr.it/
[4] https://www.w3.org/TR/skos-reference/
[5] https://www.w3.org/TR/void/
[6] https://www.w3.org/RDF/

The paper is organized as follows. Section 2 presents the related work on quality of Linked Data, while, Section 3 presents the LusTRE framework. *Average linkset importing* and *average linkset reachability* measures, are presented in Section 4 and the evaluation of LusTRE linksets is discussed in Section 5. Finally, Section 6 illustrates conclusions and future work.

2 Related Work

A recent systematic review of quality assessment for Linked Data can be found in the SWJ submission [16]. This paper reviews quality dimensions traditionally considered in data quality (e.g., availability, timeliness, completeness, relevancy, availability, consistency) and Linked Data specific dimensions, such as licensing and interlinking, considering, for the latter, the framework LINK-QA [9] and the works [15], [2]. LINK-QA defines two network measures specifically designed for Linked Data (i.e., Open SameAs chains, and Description Richness) and three classic network measures (i.e., degree, centrality, clustering coefficient) for determining whether a set of links improves the overall quality of Linked Data. Whilst, [15] and [2] detect the quality of interlinking via crowd-sourcing. Recently, logic detection of invalid SameAs has been proposed in [12]. The main differences with respect to the two linkset measures deployed in this paper are: (i) [9], [15], [2], [12] work on links independently from the fact that links are part or not of the same linksets; (ii) [9], [15], [2], [12] address the correctness of links, and not the gain in terms of multilingualism or browsing space. A set of scoring functions measuring the gain obtained when complementing a dataset with its `owl:sameAs` linksets is proposed in our previous work [6], that, we extend here deploying two new measures based on `skos:exactMatch` linkset among environmental SKOS thesauri. A set of quality measures specific for SKOS thesauri have been proposed in [14]. The paper summarizes a set of 26 quality issues for SKOS thesauri and shows how these can be detected and improved by deploying qSKOS [11], PoolParty checker, and Skosify [13]. Unfortunately, an analysis on linksets among thesauri is not included in [14]: missing out-links and in-links are adopted as indicators of SKOS thesaurus quality, but, their potential in terms of reachability of new concepts or importing of `skos:prefLabel` and `skos:altLabel` values is not considered. A first attempt to evaluate linkset quality is the linkset importing measure presented in our previous work [5]. Average linkset importing differs from linkset importing, since the former considers the absolute number, while the latter considers the percentage of "new values" of an RDF property accessed through a linkset.

3 LusTRE Framework

Multilingual linked thesaurus framework for the environment (LusTRE) is an interesting example of the exploitation of Linked Data to support metadata compilation and information discovery, for describing and for finding INSPIRE data and INSPIRE data services (or, environmental geodata in general). The

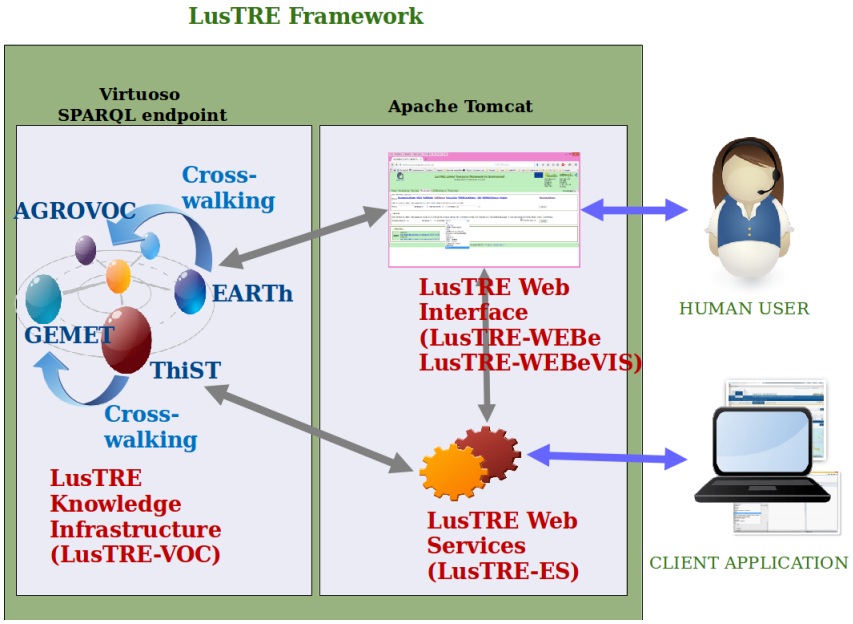


Fig. 1. LusTRE components and their interactions.

main goal of having such a framework is to be able to preserve and retrieve the information based on the semantic definitions, rather than just lexical keywords. This would guarantee the uniformity of the persisted metadata information, as well as discovery of metadata based on the semantic meanings even if metadata include diverse and dissimilar keywords [1].

The main components of LusTRE and their interactions, illustrated in Fig. 1, are the following:

- **LusTRE knowledge infrastructure (LusTRE-VOC)** contains different environmental thematic vocabularies, the interlinking among them, and the access to them as one virtual integrated Linked Data source. It is deployed on Virtuoso server and accessible by SPARQL endpoint.
- **LusTRE Exploitation Services (LusTRE-ES)** is a set of end-user oriented web services with a REST interface. It allows to exploit the knowledge contained in the LusTRE-VOC for improving client applications such as a metadata editor or a geodata portal. In particular, LusTRE-ES supports the automatic navigation among vocabulary (that is, *the cross-walking* explained in the following).
- **LusTRE web interface** provides a human-accessible interface to manually search and navigate the interlinked knowledge infrastructure using a textual (LusTRE-WEBe) or a visual browsing (LusTRE-WEBeVIS).

Table 1. Linksets in LusTRE.

Linkset Name	*Subject Thesaurus*	*Object Thesaurus*	*Linkset Cardinality*
AGROVOC2EARTH	AGROVOC	EARTH	1438
AGROVOC2EUROVOC		EUROVOC	1269
AGROVOC2GEMET		GEMET	1188
AGROVOC2THIST		THIST	1695
EARTH2AGROVOC	EARTH	AGROVOC	1438
EARTh2EUROVOC		EUROVOC	1346
EARTH2GEMET		GEMET	4365
EARTH2THIST		THIST	1140
EUROVOC2AGROVOC	EUROVOC	AGROVOC	1269
EUROVOC2EARTh		EARTH	1346
EUROVOC2GEMET		GEMET	1683
EUROVOC2THIST		THIST	733
GEMET2AGROVOC	GEMET	AGROVOC	1188
GEMET2EARTH		EARTH	4365
GEMET2EUROVOC		EUROVOC	1683
GEMET2THIST		THIST	792
THIST2AGROVOC	THIST	AGROVOC	1695
THIST2EARTH		EARTH	1140
THIST2EUROVOC		EUROVOC	733
THIST2GEMET		GEMET	792

To be concrete, as shown in Fig. 1, there are two kinds of usage provided by LusTRE: (i) for the direct manual interaction, LusTRE offers to the end user a number of Web GUI elements for browsing, inspecting, searching and translating thesaurus concepts and for visualizing the knowledge structures hidden in the interlinked thesauri; and (ii) for the transparent access, it offers the possibility to extend the functionalities of existing third-party tools by accessing the web services. Currently, LusTRE web services are being integrated in client applications such as the EUOSME[7] and QSPHERE[8] metadata editors and the under development version of the INSPIRE Geoportal[9].

A relevant feature provided by LusTRE is the *cross-walking*: the possibility to automatically navigate among matching concepts belonging to different linked thesauri. It supports easier working beyond the scope and limitations of a single thesaurus, possibly enriching data at hand, and, thus, improving user satisfaction in data consuming process. Considering on one hand the important efforts spent to construct all the linksets in LusTRE, and on the other hand the potentiality of the exploitation of the linksets through the cross-walking, we decide to investigate the quality for LusTRE linksets. In particular, we focus on: (i) the multilingualism improvement and (ii) the widening of the browsing space. We consider in LusTRE the following SKOS thesauri, with the number of concepts and languages indicated within brackets: ThiST (34150 concepts , 2 languages), AGROVOC (32310 concepts, 24 languages), EARTh (14350 concepts, 2 languages), EUROVOC (6883 concepts, 23 languages), GEMET (5223 concepts, 32 languages). Concerning linksets, we consider the twenty `skos:exactMatch` linksets, presented in Table 1, describing the couple of thesauri involved and the number of links in each linkset (that is, linkset cardinality). Linksets have been created by working out the transitive closure on existing `skos:exactMatch` and applying specific linkset discovery tasks. Every linkset discovery task has

[7] http://showcase.eenvplus.eu/client/editor.htm
[8] http://plugins.qgis.org/plugins/qsphere/
[9] http://showcase.eenvplus.eu/client/geoportal.htm

followed a two-steps process: firstly, SILK[10] has been applied to discover new links, then the link correctness has been validated by some domain experts. Linkset completeness is reasonably ensured by having applied different and not very restrictive matching functions during the discovery task. SKOS entailments have been materialized to support clients with limited processing power. As a consequence of the such materialization and the `skos:exactMatch` symmetry, reciprocal linksets (e.g., EARTH2GEMET and GEMET2EARTh or EUROVOC2AGROVOC and AGROVOC2EUROVOC in Table 1) have exactly the same links but inverted.

4 Quality Measures for Linkset

In this section, we present the two linkset quality measures, evaluating the information accessed cross-walking the linksets of LusTRE. To this purpose, we identify in the following with L the linkset between the subject thesaurus T_s and the object thesaurus T_o, the cardinality of a thesaurus T with $|T|$. The two linkset measures address different aspects of linkset quality: (i) the average linkset reachability, that estimates the enrichment of thesaurus browsing space. It evaluates, for each linkset, the average number for link of concepts in the object thesaurus browsable starting from the concepts involved in the linkset; (ii) the average linkset importing, that focuses on the average number for link of new values of a certain RDF property, reachable by the linkset in the object thesaurus. In this paper, the average linkset importing evaluates the average number of new `skos:prefLabel` and `skos:altLabel` reachable through the linkset, and it can help in addressing the incomplete language coverage issue, which affects many popular SKOS thesauri [14].

The application of the two measures requires the correctness of thesauri. Note that correctness is not the focus of our measures, in fact, our objective is to evaluate the additional information collected by the subject SKOS thesaurus from different object SKOS thesauri through different linksets. For reachability, we assume also *completeness*, described before, for `skos:exactMatch` linkset, that is, *any concept in the subject thesaurus having an exact equivalent concept in the object thesaurus must be involved in a* `skos:exactMatch` *link and of course in the linkset.* Nevertheless, if correctness and completeness are not satisfied our measures might take into account duplicated information and the final evaluation might differ from the real one. In any way, these assumptions seem reasonable since: (i) currently, all applications consuming Linked Data implicitly assumes at least correctness (trusting on publisher reliability); (ii) there are several tools (SILK, LIMES and qSKOS) that can help to reach linkset correctness and completeness. Finally, we consider the set of SKOS properties for average linkset importing $SKOSlabel$ = {`skos:prefLabel`,`skos:altLabel`} and for average linkset reachability $SKOSrel$ ={`skos:narrower`, `skos:related`, `skos:broader`}.

[10] http://silkframework.org/

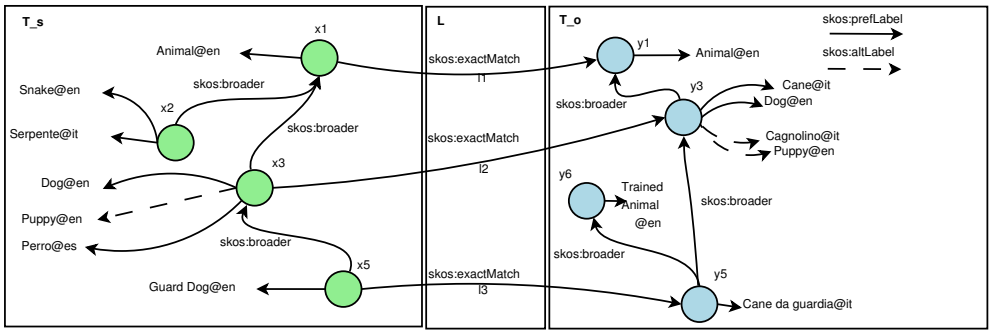


Fig. 2. Example of RDF/SKOS thesauri and `skos:exactMatch` linkset.

4.1 Average Linkset Importing Measure

The average linkset importing measure evaluates the average number of "new values" of a RDF property p, accessible through a link. Generalising to the entire linkset, "new values" are those not already present in the subject thesaurus T_s, but reachable through the linkset L in object thesaurus T_o.

Definition 1 ***(Average linkset importing).*** *Let ln be an ISO language tag or _ (which stands for all languages), l be a link in L of the form (t_s, `skos:exactMatch`, t_o) and p be in SKOSlabel. Let TsVal(p,l,ln) be the values of p for concept t_s and ToVal(p,l,ln) be the values of p for concept t_o. The average linkset importing measure can be defined as follows:*

$$ALI^{p}_{ln} = \frac{1}{|L|} \sum_{l \in L} |ToVal(p,l,ln) - TsVal(p,l,ln)|$$

Example 1. Considering the thesauri T_s, T_o and linkset L in Fig. 2, and p = `skos:prefLabel`, TsVal(`skos:prefLabel`,l2,en)={Dog@en}, whilst TsVal(`skos:prefLabel`,l2, _) ={Dog@en, Perro@es}, since in the latter there is no constraint on the language tag. Considering y_3, the `skos:exactMatch`-linked concept for x_3, all the translations for the `skos:prefLabel` in the object thesaurus are ToVal(`skos:prefLabel`,l2,_)={Dog@en, Cane@it}, while for the subject the translations are TsVal(`skos:prefLabel`,l2,_) ={Dog@en,Perro@es}. Thus, $ALI^{\texttt{skos:prefLabel}}_{\_} = \frac{1}{3} * (0 + 1 + 1) = \frac{2}{3} = 0.67$. A value of 0.67 represents the average number of new `skos:prefLabel` values, in any language, accessible through each link of the linkset. The result means that every three links, two new translations of `skos:prefLabel` are accessible.

4.2 Average Linkset Reachability Measure

The *average linkset reachability* evaluates the average number of concepts reachable when cross-walking a link of the linkset and exploring the object thesaurus T_o until a certain depth, identified with the number of hops. Concepts in the T_o directly involved in the linkset are not considered. .

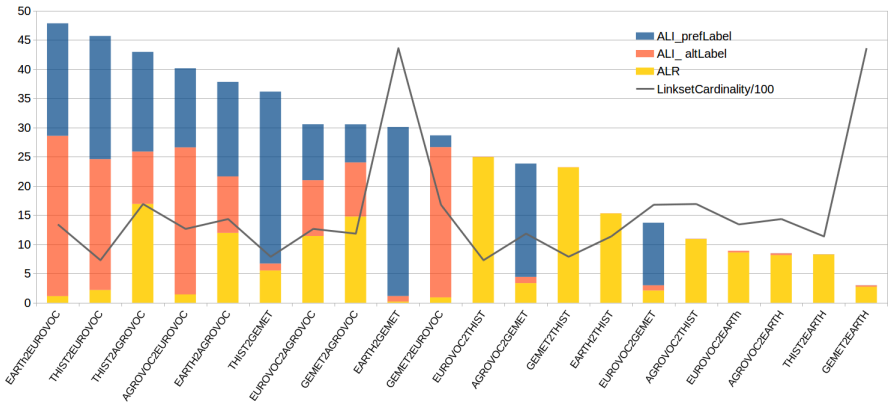


Fig. 3. Average linkset measures results (histograms) and cardinality/100 (line) of LusTRE linksets.

Definition 2 ***(Average linkset reachability).*** *Let L_o be the set of object concepts of each link in L, $ConceptsT_o(k,SKOSrel,L_o)$ be the set of concepts in T_o, from the concepts in L_o, through the relations in SKOSrel in a number of hops <= k. The average linkset reachability is defined as follows:*

$$ALR_k^{SKOSrel} = \frac{|ConceptsT_o(k, SKOSrel, L) - L_o|}{|L|}$$

Example 2. Considering the thesauri T_s and T_o and linkset L in Fig. 2, L_o is represented by $\{y_1, y_3, y_5\}$, while considering a number of hop k equal to 2 and p=`skos:broader`, $ConceptsT_o(2, \texttt{skos:broader}, L_o) = \{y_1, y_3, y_5, y_6\}$. The $ALR_2^{SKOSrel} = \frac{|\{y_1,y_3,y_5,y_6\} \setminus \{y_1,y_3,y_5\}|}{3} = \frac{1}{3} = 0.33$. The 0.33 represents the average number of concepts in object thesaurus reachable, in 4 hops, to each link in the linkset, which implies one new concept every three links.

5 Linksets Quality Evaluation of LusTRE

We have developed a prototype in JAVA/JENA that implements both average linkset importing and reachability, and applied it to all the linksets developed among within SKOS thesauri of LusTRE (see Table 1 for details). LusTRE linksets satisfy both requirements of correctness and completeness, since they have been validated by domain experts. As already discussed, the considerable efforts spent during the project eENVplus to create and validate LusTRE linksets, suggest that an evaluation of the linkset in terms of usefulness and information enrichment is important. In this application case, the *average linkset importing* measure evaluates the average number of new values for `skos:prefLabel` and `skos:altLabel`, in different languages, reachable through each link, in the object thesaurus. While, the *average linkset reachability* represents the average number of concepts accessible in the object thesaurus, through each link, in a number

of hops less or equal to 4, considering the SKOS relations `skos:narrower`, the `skos:broader` and `skos:related`. We chose k=4, since it seems a reasonable number of steps that a user should performs, for example, during the search of a specialization/generalization of a SKOS concept in a thesaurus. The results obtained by both measures allow to analyse the quality of the linkset at different levels of detail. An overall evaluation of the linksets is shown in Fig. 3, where the x axis shows the linksets and the y axis shows the average linkset reachability and the average linkset importing. In each bar in the graph, we have piled the results of importing for `skos:altLabel` and `skos:prefLabel`, and of reachability for number of hops k=4. All the linksets provide a set of new values in terms of `skos:prefLabel`, `skos:altLabel` or new concepts. Thus, globally we can say that the efforts spent in the development of LusTRE linksets have paid off, in fact, the multilingualism and/or the browsing space of a single thesaurus are enriched by its linksets. After that, we can also observe another important fact: the substantial independence between the quantity of new label translations/concepts accessed and the cardinality (divided by 100) of the linkset shown as a line in Fig. 3. Often, the quality of a linkset is identified with its cardinality, more link means more quality, but, as shown in the graph the two facts are not related. In particular, the linkset EUROVOC2GEMET has about 1600 links, and it provides for each link a total of about 15 new translations and concepts, while the linkset EARTH2EUROVOC has about 1300 links, but it brings a total greater than 45 of new translations and concepts for each link, almost three times EUROVOC2GEMET. As a consequence, it seems clear that the quality of a linkset is substantially independent and more complex than its cardinality.

More in details, we analyse the result of each measure for a single link. We notice that, as obvious, the two measures captures different aspects of the linkset quality, thus, a linkset can have two different evaluation in the two measures. Considering Fig. 4, for example, we notice that linksets EARTH2EUROVOC and THIST2EUROVOC are "good" for average linkset importing and "poor" for average linkset reachability, and vice versa GEMET2THIST and EUROVOC2THIST are "good" for average linkset reachability and "poor" for average linkset importing. These results can be used, also, to formulate some hypothesis concerning the content of the involved thesauri, for example, Fig. 4 shows that EARTH thesaurus should not have to many translations or it should have but in languages different from GEMET, since it adds, in average, about 25 new `skos:prefLabel` for each link.

Finally, it is possible to deepen the analysis of the average link importing focusing on the average number of values for each distinct languages accessible for each link of the linksets. As an example, we focus on EARTH2AGROVOC and EARTH2GEMET, and we report the result in the radial graph in Fig. 5, where radial axes include one axis for each language (41 different languages are considered) representing the average importing for link. In Fig. 5, we observe that, both linksets import the same set of 10 languages (i.e., ar, ru, es, tr, pt, pl, de, fr, hu, cs). Besides, 19 languages (e.g., bg, ga, fi, sl, eu, ro) can be imported only through EARTH2GEMET, and 9 only through EARTH2AGROVOC. Discussing

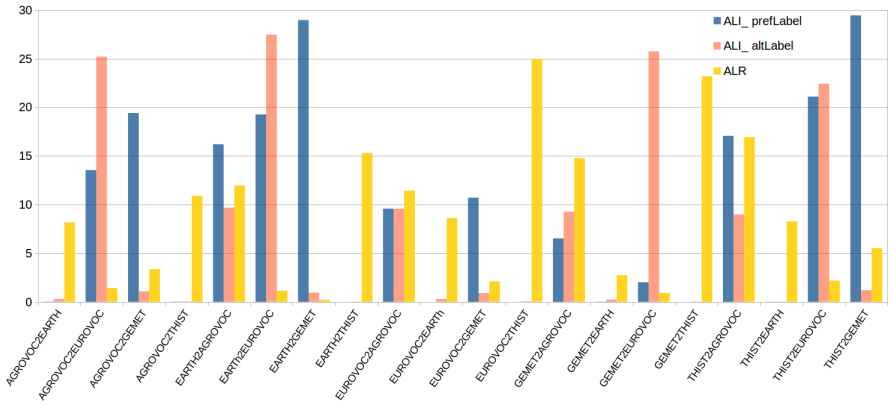


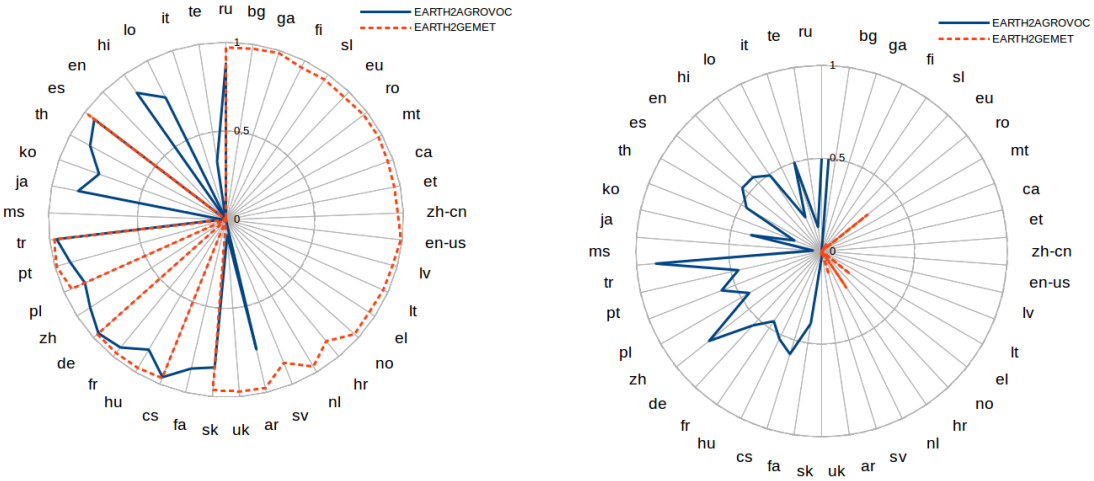
Fig. 4. Average linkset importing and reachability in details.

the graphs in detail, we have two opposite situations: (i) for `skos:prefLabel`, Fig. 5(a), it is evident that both the average importing and the number of languages of the linkset EARTH2GEMET are higher than in EARTH2AGROVOC; (ii) for `skos:altLabel`, Fig. 5(b), we have exactly the opposite situation. In fact EARTH2AGROVOC imports more `skos:altLabel` and also more different languages than EARTH2GEMET. In fact, we import about 20 languages from EARTH2AGROVOC and only 4 from EARTH2GEMET.

As a synthesis , we can say that in average the linksets developed in LusTRE bring an advantage in term of multiligualism and browsing space, as a consequence all the actors (human or software) of the data consuming process can benefit from this information enrichment, exploiting the cross-walking feature provided by LusTRE. Focusing on the linkset quality issue, it is evident that, the choice of the best linkset is not straightforward, but it depends on the specific point of view of the analysis. In our case, we have demonstrated the ability of a linkset to enrich the thesaurus multilingualism and browsing space. However, the linkset quality is a very complex issue, thus, its full characterizations requires the definition of further measures capturing different aspects. On the other hand, it is evident that considering only the number of links is not enough.

6 Conclusions and Future Work

This paper presents the evaluation of the linksets developed in the multilingual linked thesaurus framework for the environment LusTRE. Our purpose, in the assessment of linkset quality of LusTRE, is to evaluate if the efforts spent in the creation of such connections results in some improvement in consuming LusTRE data. This is a quite new point of view in Linked Data quality assessment since almost all the existing methods focus on datasets. Our approach is based on two considerations: (i) linkset is as important as data in the evolution of the Web of Data into the Global Data Space; (ii) the creation of linksets require efforts, thus, the evaluation of their quality is pivotal. To this purpose we consider



(a) skos:prefLabel (b) skos:altLabel

Fig. 5. Average linkset importing for EARTH2AGROVOC and EARTH2GEMET.

two measures evaluating the information accessible through a linkset: average linkset importing and average linkset reachability. Average linkset importing can be exploited to evaluate the multilingual enrichment, in terms of translated labels, while average linkset reachability assesses the browsing space extension in terms of new concepts accessible through the linkset. Both measures are used to evaluate the `skos:exactMatch` linksets developed in LusTRE. The results show that the efforts spent in the creation of the linksets are paid off, in fact, all the linksets provide an increase of the multilingualism and of the browsing space. As future work, we plan to extend the quality framework adding other measures able to capture other aspects of linkset quality. Moreover we foresee to encode the quality results according to the Data Quality Vocabulary (DQV) [7], developed by the Data on the Web Best Practices Working Group, and to face the challenge to apply the measures "into the wild", that is to the `skos:exactMatch` linksets exposed in the LOD Cloud.

Acknowledgements. The paper activity has been carried out within the EU funded project eENVplus (CIP-ICT-PSP grant No. 325232). The authors would like to thank all partners and, in particular, Paolo Plini (IIA-CNR) for the important collaboration, and the team of the European Commission's Joint Research Centre (Italy) for the valuable contribution.

References

1. Abecker, A., Wössner, R., Schnitter, K., Albertoni, R., De Martino, M., Podestà, P.: Latest developments of the linked thesaurus framework for the environment (LusTRE). In: 29th EnviroInfo and 3rd ICT4S Conference 2015, Copenhagen, Denmark, September 7-9 (2015)

2. Acosta, M., Zaveri, A., Simperl, E., Kontokostas, D., Auer, S., Lehmann, J.: Crowdsourcing linked data quality assessment. In: Alani, H., Kagal, L., Fokoue, A., Groth, P.T., Biemann, C., Parreira, J.X., Aroyo, L., Noy, N.F., Welty, C., Janowicz, K. (eds.) The Semantic Web - ISWC 2013 - Sydney, NSW, Australia, October 21-25, 2013. LNCS, vol. 8219, pp. 260–276. Springer (2013)
3. Albertoni, R., De Martino, M., Di Franco, S., De Santis, V., Plini, P.: EARTh: An environmental application reference thesaurus in the linked open data cloud. Semantic Web 5(2), 165–171 (2014)
4. Albertoni, R., De Martino, M., Podestà, P.: Environmental thesauri under the lens of reusability. In: Ko, A., Francesconi, E. (eds.) EGOVIS 2014. LNCS, vol. 8465, pp. 222–236. Springer (2014)
5. Albertoni, R., De Martino, M., Podestà, P.: A linkset quality metric measuring multilingual gain in SKOS thesauri. In: Rula, A., Zaveri, A., Knuth, M., Kontokostas, D. (eds.) Proceedings of the 2nd Workshop on Linked Data Quality co-located with 12th Extended Semantic Web Conference (ESWC 2015), Portorož, Slovenia, June 1, 2015. CEUR Workshop Proceedings, vol. 1376. CEUR-WS.org (2015)
6. Albertoni, R., Gómez-Pérez, A.: Assessing linkset quality for complementing third-party datasets. In: Guerrini, G. (ed.) EDBT/ICDT Workshops. pp. 52–59. ACM (2013)
7. Albertoni, R., Isaac, A., Debattista, J., Dekkers, M., Guret, C., Lee, D., Mihindukulasooriya, N., Zaveri, A.: Data on the Web Best Practices: Data Quality Vocabulary (2016), `http://www.w3.org/TR/vocab-dqv/`, W3C Working Draft, accessed: 28-07-2016
8. Caracciolo, C., Stellato, A., Morshed, A., Johannsen, G., Rajbhandari, S., Jaques, Y., Keizer, J.: The AGROVOC linked dataset. Semantic Web 4(3), 341–348 (2013)
9. Guéret, C., Groth, P.T., Stadler, C., Lehmann, J.: Assessing linked data mappings using network measures. In: Simperl, E., Cimiano, P., Polleres, A., Corcho, Ó., Presutti, V. (eds.) ESWC 2012. LNCS, vol. 7295, pp. 87–102. Springer (2012)
10. Heath, T., Bizer, C.: Linked Data: Evolving the Web into a Global Data Space. Synthesis Lectures on the Semantic Web, Morgan & Claypool Publishers (2011)
11. Mader, C., Haslhofer, B., Isaac, A.: Finding quality issues in skos vocabularies. In: Zaphiris, P., Buchanan, G., Rasmussen, E., Loizides, F. (eds.) TPDL 2012. LNCS, vol. 7489, pp. 222–233. Springer (2012)
12. Papaleo, L., Pernelle, N., Saïs, F., Dumont, C.: Logical detection of invalid sameas statements in RDF data. In: Janowicz, K., Schlobach, S., Lambrix, P., Hyvönen, E. (eds.) EKAW 2014, Linköping, Sweden, November 24-28, 2014. LNCS, vol. 8876, pp. 373–384. Springer (2014)
13. Suominen, O., Hyvönen, E.: Improving the quality of skos vocabularies with skosify. In: ten Teije, A., Völker, J., Handschuh, S., Stuckenschmidt, H., d'Aquin, M., Nikolov, A., Aussenac-Gilles, N., Hernandez, N. (eds.) EKAW 2012. LNCS, vol. 7603, pp. 383–397. Springer (2012)
14. Suominen, O., Mader, C.: Assessing and improving the quality of skos vocabularies. J. Data Semantics 3(1), 47–73 (2014)
15. Zaveri, A., Kontokostas, D., Sherif, M.A., Bühmann, L., Morsey, M., Auer, S., Lehmann, J.: User-driven quality evaluation of dbpedia. In: Sabou, M., Blomqvist, E., Noia, T.D., Sack, H., Pellegrini, T. (eds.) I-SEMANTICS 2013, Graz, Austria, September 4-6, 2013. pp. 97–104. ACM (2013)
16. Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J., Auer, S.: Quality assessment for linked open data: A survey. Semantic Web 7(1), 63–93 (2016)