Towards the reuse of standardized thesauri into ontologies[1]

Elena Cardillo[1], Antonietta Folino[2], Roberto Trunfio[2], Roberto Guarasci[2]

[1] Institute for Informatics and Telematics, Italian National Research Council, Pisa, Italy
[2] Laboratorio di Documentazione (LabDoc), University of Calabria, Rende, Italy
elena.cardillo@iit.cnr.it
{antonietta.folino,roberto.trunfio,roberto.guarasci}@unical.it

Abstract. One of the main holdbacks towards a wide use of ontologies is the high building cost. In order to reduce this effort, reuse of existing Knowledge Organization Systems (KOSs), and in particular thesauri, is a valuable and much cheaper alternative to build ontologies from scratch. In the literature tools to support such reuse and conversion of thesauri as well as re-engineering patterns already exist. However, few of these tools rely on a sort of semi-automatic reasoning on the structure of the thesaurus being converted. Furthermore, patterns proposed in the literature are not updated considering the new ISO 25964 standard on thesauri. This paper introduces a new application framework aimed to convert thesauri into OWL ontologies, differing from the existing approaches for taking into consideration ISO 25964 compliant thesauri and for applying completely automatic conversion rules.

Keywords: Thesauri; Ontology; Knowledge Organization System; Conversion Framework; ISO 25964; OWL.

1 Introduction

Nowadays, Knowledge Organization Systems (KOSs) (e.g. thesauri, taxonomies, ontologies, classification systems) are assuming a key role in knowledge and information management. Even though a KOS can be used as a standalone application, to take benefit of its features it is generally integrated in larger information systems (search engines, content management systems, etc.). Thus, in the literature a consolidated issue is the interoperability of different KOSs for a cost-effective exchange of structured information. Driven by the philosophy of the Semantic Web [4] that widely relies on ontologies for knowledge sharing, reuse and inference, researchers are concentrating their effort towards the reuse of less formal terminological systems, and in particular thesauri, into ontologies. This is motivated by the fact that a huge amount of thesauri have been already realized in the last decades for almost every knowledge

[1] Authors have jointly collaborated in the drafting of the article. However, each has been involved in one or more paragraphs: E. Cardillo of Section 1, 2.1, and 3; A. Folino of Section 2.2., and 4.2; R. Trunfio of all Section 4; and R. Guarasci of Section 5.

domain [18] and some of them are steadily updated. So, there could be no benefit to replicate a domain modeling from scratch, but recovering and reusing the knowledge embedded into a thesaurus could speed-up the activity of building ontologies.

In order to address this challenging topic, this paper introduces an application framework for reusing standardized thesauri[2] into ontologies as defined by [7]. In particular, the framework uses new re-engineering rules formulated to be applied on an ISO 25964 compliant thesaurus [12, 13].

The paper is organized as follows: a brief description of the main similarities and differences between thesauri and ontologies, and between thesaurus entities before and after the introduction of ISO 25964 is given in Section 2; an updated overview of the literature is provided in Section 3; the description of the proposed re-engineering approach along with some examples of rules for the reuse of a standardized thesaurus into an OWL ontology are given in Section 4; finally, conclusions and future works are outlined in Section 5.

2 Thesauri and Ontologies: an overview

In this section, first a brief description of the similarities and differences between thesauri and ontologies is given; secondly, some important novelties introduced by the ISO 25964 standard in terms of thesaurus principles, interoperability and mapping with other types of controlled vocabularies are briefly reported.

2.1 Thesauri vs Ontologies

Some similarities between thesauri and ontologies can be identified. In particular, both describe a domain, include concepts and relations between them, use hierarchies.

Furthermore, both are used in applications of information management for cataloguing and in search engines. However, several differences must be taken into account, specifically depending on their origin and purposes. Thesauri have been used for a long time in librarian contexts as indexing tools and controlled vocabularies. As such they are thought to represent knowledge in a less formal and comprehensive manner than ontologies. They are not characterized by a level of conceptual abstraction as ontologies, which are originated from philosophy and conceived as accepted and formal ways of describing knowledge domains[3]. So difference between a concept and its lexicalization in a thesaurus is not clearly established.

Moreover, other fundamental differences are related to their structures. As clearly described in [10], with respect to thesauri, ontologies are characterized by: *(i)* an explicit representation of the types of relationships; *(ii)* the use of powerful formalisms

---

[2] Thesauri are defined as "controlled and structured vocabulary in which concepts are represented by terms, organized so that relationships between concepts are made explicit, and preferred terms are accompanied by lead-in entries for synonyms or quasi-synonyms" [12].

[3] Ontologies consist of a set of definitions of classes, properties, and individuals, together with a set of axioms expressing the relations between classes and properties, and a set of facts about particular individuals.

(e.g. axioms, relationships cardinality). These differences enable ontologies to reuse knowledge of a domain, make domain assumptions explicit, and allow access to and evolution of legacy data, as well as automated reasoning.

2.2 ISO 25964 vs other standards on thesauri

In this section we briefly analyze structural and conceptual differences between ISO norms on thesauri in order to allow for the proposal of updated and more standardized re-engineering rules for building ontologies from thesauri.

From a more comprehensive perspective, the need for a new norm about thesauri and other controlled vocabularies was due to the technological evolution and to the new role recognized to thesauri as information retrieval tools and not only as indexing resources. This required new specifications about data models, exchanging formats, interoperability with other vocabularies, etc., all present in the ISO 25964 [12, 13]. ISO 25964 completely replaces the previous ISO 2788:1986 [15] and ISO 5964:1985 [14] standards, and almost reuses the content of two other national norms: ANSI/NISO Z39.19:2005 [2] and BS 2783-5:2008 [5].

Given the aim of the paper, we focus here only on three main changes that contribute to approach thesauri towards ontologies and that are of interest for re-engineering purposes (see [6] for a complete analysis):

- ***Term-based vs Concept-based thesaurus.*** Considering thesaurus structure, ISO 25964 provides a data model (absent in ISO 2788:1986) that formally represents its objects and relationships. The major difference that can be observed in this data model is the clear distinction between concepts and terms[4] (*ThesaurusConcept* is separated from *ThesaurusTerm* and from *PreferredTerm*, *hasPreferredLabel*, *hasNonPreferredLabel*, etc.). In fact, similarly to ontologies, concepts in ISO 25964 are defined as units of thoughts lexically represented by terms. Thus, a thesaurus is considered as a concept-based resource, rather than a term-based one.

- ***Thesaurus structure.*** The ISO 25964 adds important elements for organizing the structure of a thesaurus, such as the *thesaurus array*[5] and *concept group*[6] elements, that were not explicitly illustrated in the previous ISO norm. This implies that in a conversion process these constructs should be taken into account, so new patterns have to be introduced in order to reengineer them into ontology elements.

- ***Semantic relationships.*** One of the innovations introduced by the current norm is the possibility to make explicit the nature of semantic relationships, by further specifying the standard ones. This possibility contributes to make thesauri more similar to ontologies, as well as the clear distinction between *concept* and *term.* In

[4] This does not mean that in the ISO 2788:1986 the same meaning was recognized to terms and concepts, but for practical aims they were treated as synonyms.

[5] A *thesaurus array* is defined in [13, p.4] as “group of sibling concepts”.

[6] A *concept group* is defined in [12, p.1] as “a group of concepts selected by some specified criteria, such as relevance to a particular subject area”.

particular, some changes regarding the *equivalence* and the *hierarchical* relationships are worth mentioning:

- *Equivalence relationship*. The ISO 25964 clarifies, first of all, that it relates terms rather than concepts (more than one term for expressing a unique concept). In addition to the standard tags USE/UF, the norm establishes that other tags could be used, depending on the kind of terms to be put into relation: AB for abbreviations, FT for the full form of a term, SP for spelling variant.
- *Hierarchical relationship.* It holds between two or more concepts that express subordinate and super-ordinate meanings at different levels and is depicted through the tags BT/NT (i.e. narrower term and broader term, respectively). To extend the richness of thesauri, hierarchical relationships can be further divided into generic (NTG/BTG), partitive (NTP/BTP) and instantial (NTI/BTI). Both ISO 2788 and 25964 clarify that concepts can be hierarchically related only if they belong to the same conceptual category (objects, actions, properties, etc.) and if the so called *all-and-some test* is verified. Moreover, the ISO 25964 specifies that this relationship holds "*between a pair of concepts when the scope of one of them falls completely within the scope of the other*" [12]. This criterion is undoubtedly important to guarantee the correctness of the hierarchy.

3 Related Works

In the last decades the conversion of thesauri into ontologies has taken the attention of the scientific community, that produced several stimulating works. Nonetheless, few of them rely on a computer-aided conversion process. Moreover, most of the existing approaches were conceived before the publication of ISO 25964.

Two pioneering works described in [22] and [19] show the creation of RDFS Ontologies reusing domain specific thesauri: the former in particular proposes highly structured semantic descriptions for a subset of art-object (antiquate furniture) from the Art and Architecture Thesaurus (AAT)[7]; the latter approach is tested on the well-known AGROVOC thesaurus[8]. In both cases BT/NT relationships are mapped to the is-a relation in the ontology and terms of the thesaurus to classes.

A semiautomatic approach is proposed in [3] where the TERMINAE method for ontology extraction from texts is applied to a thesaurus. Here three hypotheses about the reuse of terms relations are proposed: *(i)* preferred terms in the thesaurus are indexed to constitute terms identifying domain concepts in the ontology; *(ii)* relations between terms and preferred terms are synonym relations in the thesaurus, that means putting together terms as label of the same concept in the ontology; *(iii)* relations between preferred terms are indexed to identify concepts relations. Even if some interesting conversion rules are presented, the application of this methodology to particular use cases is very limited (i.e. an application for the medical domain).

[7] http://www.getty.edu/research/tools/vocabularies/aat/
[8] http://aims.fao.org/standards/agrovoc/about/

More close to our work is the conversion approach proposed by [9], where they formalize an ISO 25964 compliant thesaurus into an OWL ontology[9], assuming that a standard-compliant thesaurus is *concept-based*, rather than *term-based*, as observed in the norm. This implies that thesaurus concepts and facets are treated as classes in the corresponding ontology, while thesaurus terms (synonyms or quasi-synonyms introduced by the equivalence relationships) are represented in the ontology as class labels. In addition, checking and refinement of thesaurus relationships are performed in order to: explicitly distinguish between different types of hierarchical relationships, manage cycles, orphans, polyhierarchy, etc. (e.g. the *generic relation* (NTG/BTG) are translated into *is-a* relation in the ontology, and the *instantial* one (NTI/BTI) into the distinction between class and individuals). Moreover, in order to guarantee similarity in design choices, an alignment of the thesaurus (i.e. AGROVOC) with top-level ontologies is proposed. However, also in this case, the approach "has not been tested and refined on the scale of re-engineering a complete thesaurus" [9].

In addition to these approaches, in the last few years different patterns for the re-engineering of non-ontological resources (including thesauri) into ontologies have been proposed. We mainly refer here to those proposed in [20] and [21], whose objective is the conversion of both terms/concepts and semantic relationships provided by thesauri (USE/UF; BT/NT; RT) into ontology classes, individuals and relations.

Other approaches that try to convert thesauri, but also classification systems into more formal resources as ontologies can be found in [8, 16, 17].

Considering the literature, the aim of our research is to demonstrate that, although these studies proved usefulness and applicability, some of the existing patterns need to be revised/integrated in the light of the mentioned ISO norm 25964 [12, 13].

Furthermore, due to the common issues related to the conversion process (e.g. the ambiguity in the distinction between a class and an instance, difficulties in the specification of the nature of thesaurus relationships) in our opinion, a challenge for the future is the development of new methodological and application frameworks that allow for more specific, standard based, and automatic conversions.

4 A Framework for Reusing ISO 25964 Thesauri into OWL Ontologies

By taking advantages of the analysis provided in Section 2, this section briefly outlines the concepts at the basis of our conversion framework. The framework's aim is to export the knowledge embedded into an ISO 25964 compliant thesaurus and to reuse it as the skeleton for a formalized ontology. Specifically, the application constructs an ontology in OWL 2 DL[10] starting by an RDF-based schema for modelling an ISO 25964 standardized thesaurus. The framework is implemented in Java 7 and uses APIs taken from the well-known *Apache Jena*[11] framework for building Seman-

[9] https://www.w3.org/TR/owl2-overview/
[10] http://www.w3.org/TR/owl-features/
[11] https://jena.apache.org/

tic Web and Linked Data application. In particular, it uses the `Model Loader` and the `Model Translator` APIs. The former is devoted to enable the loading of a thesaurus formalized through a model specified via RDF[12] which is handled by the `RDF` API from Apache Jena.

The `Model Translator` API applies a set of rules designed to identify classes, properties, instances and semantic relations from the RDF graph and consequently to extract an ontological schema. For this purpose, the `Ontology` API from Apache Jena is used to represent a formal logical descriptions in OWL 2 DL. The use of OWL 2 DL is motivated by the fact that it is decidable and its tableau reasoners prove to be tractable in practice on several scales of problems (despite the poor theoretical complexity). Moreover, for reasons of performance, all the conversion rules are hard coded in the Java library. The Java API is available upon request from the authors.

4.1 Model Loader API

As mentioned before, the `Model Loader` is devoted to represent a standardized thesaurus in an RDF-graph based on a SKOS (Simple Knowledge Organization System)[13] and SKOS-XL[14] extension for modelling ISO 25964 thesauri, known as iso-thes-25964 [11] (for convenience, we define the prefix `it: <http://purl.org/iso25964/skos-thes>`). It must ensure the consistency of the RDF graph with respect to the domains of the semantic relations and entailment rules provided by the iso-thes-25964 and those inherited by SKOS. To this extent, the iso-thes-25964 extension promotes the reuse of some of the standard SKOS classes (i.e. `skos:Concept`, `skos:ConceptScheme`). However, it introduces important classes like `it:ConceptGroup`, aimed to represent group of concepts selected by some specified criteria, such as relevance to a particular subject area (see Sec. 3.18 from [13]) and `it:ThesaurusArray`, used in our framework to represent a node label with characteristic of division (see Sec. 11 from [12]).

Moreover, it provides new semantic relations (e.g. `it:broaderGeneric`, `it:broaderInstantial`, `it:broaderPartitive`, `it:narrowerGeneric`, `it:narrowerInstantial`, `it:narrowerPartitive`, etc.), which integrate and specialize the existing SKOS relations intended to be used for thesauri representation (e.g.: `skos:broader`, `skos:broaderTransitive`, `skos:hasTopConcept`, `skos:inScheme`, `skos:narrower`, `skos:narrowerTransitive`, `skos:related`, `skos:topConceptOf`, `skos:member`).

Finally, following ISO 25964, we extended iso-thes-25964 by introducing the class `Facet` as sub-class of `skos:Collection`. `Facet` is defined as disjoint with classes `it:Concept`, and `it:ThesaurusArray`. Moreover, it is in domain of `it:superGroup` and in range of `it:inScheme`. In our opinion, the choice to introduce a new class is supported by the need to ensure a non-ambiguous representation of facets in the RDF graph.

[12] http://www.w3.org/TR/1999/REC-rdf-syntax-19990222
[13] http://www.w3.org/2004/02/skos/
[14] http://www.w3.org/TR/skos-reference/skos-xl.html

4.2 The Model Translator API: Conversion rules

The `Model Translator` applies a set of rules in order to extract an OWL 2 DL ontology from an ISO 25964 standardized thesaurus formalized according with the `Model Loader` specifications. The `Model Translator` can translate in an OWL class or property elements of a thesaurus such as concepts, concept groups, facets, thesaurus arrays, as well as associative (RT), hierarchical (TT, BT/NT, BTG/NTG, BTP/NTP, BTI/NTI) and equivalence relationships (USE/UF). The RDF graph is explored hierarchically starting by the nodes identified as facets, top terms and terms with no broader definition. Finally, concept groups nodes are explored. For space reasons, we give examples only for those rules that need clearer explanations.

Facets and Top Terms. Facets are used to group together concepts from the same category, that means concepts sharing similar properties, such that members of a facet are mutually exclusive with the members of another. According to the standard, the following set of facets should be used: *things, types, parts, properties, materials, processes, activities, products, by products, patients, agents, space and time*. In this perspective, although facets can be personalized, in our framework facets are considered as the classes of a foundation ontology. Hence, rule *(1)* is applied:

(1) "*Once a* `Facet` *node is found, it must be mapped as a class that is subclass of* `owl:Thing` *and has no other parent relations*".

When a `skos:Concept,` identified as a *Top term*, is found in the RDF graph, then it must be mapped to an ontological class, as stated by the following rule:

(2) "*Given a* skos:Concept *<A> and a skos:ConceptScheme <C> such that <A> skos:topConceptOf <C>, then: if exists one and only one Facet <F> such that <F> skos:member <A>, then <A> is mapped as subclass of the ontological class defined for <F>; otherwise <A> is mapped as a subclass of owl:Thing and it has no other parent relations". Moreover, if another Facet <E> exists, then a relation* owl:disjointWith *must be used between the resulting OWL classes for nodes <A> and each class under facet <E>*".

For translation convenience, whenever no TT relation are explicitly provided in the thesaurus, we assume that a concept that has not broader concepts is a top concept.

Hierarchical relationships. Generally speaking, whenever a BT/NT or a BTG/NTG relationship is found in the RDF graph, the following rule comes:

(3) "*Given two Concept nodes <A> and <B> such that <A> skos:narrower <B>, then <A> and <B> are mapped as two OWL classes such that <B> rdfs:subClassOf <A>*".

Further rules are needed for the hierarchical relationships BTP/NTP and BTI/NTI. Specifically, a BTP/NTP relation in our framework entails rule *(4)*:

> *(4)* "*Given three `Concept` nodes <A>, <B1> and <B2>, such that <A> `it:broaderPartitive` <B1>, <B2>, then <B1> and <B2> are each represented in OWL either as classes linked to node <A> by a part of relationship in the case they have further narrower concepts in the thesaurus, or as individuals of node <A>*".

A narrower concept represented via the BTI/NTI relation generally becomes instance of the OWL class obtained from the broader concept. However, as stated in [23], the following exception may occur: if at a certain level of the hierarchy in the thesaurus an instance of a NTI/BTI relation is also the broader concept of one or more concepts in other BT, BTP, or BTG relations, then it cannot be converted in the ontology as an individual. To clarify this exception, the following example is provided:

```
Countries
  NTI United States of America

United States of America
  NTP Alabama
  NTP Alaska
  NTP Arizona
```

This is mapped in RDF via the iso-thes-25964 model as follows using *turtle notation*:

```
@prefix ex: <http://example.com/skos/thes#> .
...
ex:id1 rdf:type skos:Concept;
skos:prefLabel "Countries"@en;
  it:narrowerInstantial ex:id2.
ex:id2 rdf:type skos:Concept;
  skos:prefLabel "United States of America"@en;
  it:broaderInstantial ex:id1;
  it:narrowerPartitive ex:id3, ex:id4, ex:id5.
ex:id3 rdf:type skos:Concept;
  skos:prefLabel "Alabama"@en;
  it:broaderPartitive ex:id2.
ex:id4 rdf:type skos:Concept;
  skos:prefLabel "Alaska"@en;
  it:broaderPartitive ex:id2.
ex:id5 rdf:type skos:Concept;
  skos:prefLabel "Arizona"@en;
it:broaderPartitive ex:id2.
```

Since ontologies necessarily distinct between classes and individuals, in order to enable reasoning and inferencing, rule *(5)* is adopted:

(5) "*Given two* `Concept` *nodes* <A> *and* <B> *such that* <A> `it:broaderInstantial` <B>*: if exists a* `Concept` *node* <C> *such that* <B> `skos:broader` <C>*, or* <B> `it:broaderGeneric` <C> *or* <B> `it:broaderPartitive` <C>*, then the same rule for NTP concepts is applied; otherwise,* <B> *becomes an instance of the OWL class defined for node* <A>".

According to rules (4) and (5), in the simplest case (NTP converted as individuals) the following OWL 2 DL is obtained:

```
<owl:Class rdf:ID="#United_States_of_America">
  <owl:subClassOf rdf:parseType="#Countries">
  </owl:subClassOf>
</owl:Class>
<United_States_of_America rdf:ID="Alabama">
...
</United_States_of_America>
<United_States_of_America rdf:ID="Alaska">
...
</United_States_of_America>
<United_States_of_America rdf:ID="Arizona">
...
</United_States_of_America>
```

Thesaurus Array. Another type of node of the RDF graph which is translated in our framework is `it:ThesaurusArray`, i.e. a node label used to represent a collection of siblings concepts with a common characteristic of division. As reported in ISO 25964, a node label is put between two concepts related by the NT/BT relation. A thesaurus array is not a concept itself. Moreover, following ISO 25964 definition for node labels, all the concepts under a node label represent disjoint classes. An example from the EARTh thesaurus [1] follows.

```
Forecasting
  [Forecasting by length]
    NT Long-term forecasting
    NT Short-term forecasting
  [Forecasting by target]
    NT Drought forecasting
    NT Earthquake forecasting
```

From this example, as stated in Sec. 11 of [12], it can be asserted that the concepts following the "by" clause in the node label represent a property of the concept "Forecasting". Thus, we map "length" as an object property between the OWL class for concept "Forecasting" and the two OWL classes for the concepts "Long-term forecasting" and "Short-term forecasting". Specifically, the following rules are applied:

(6) "*Given two* `Concept` *nodes* *<A> and <B> and a* `ThesaurusArray` *node <TA> such that <A>* `skos:broader` *<B> (or <A>* `it:broaderGeneric` *<B> or <A>* `it:broaderPartitive` *<B>), and <A>* `it:subordinateArray` *<TA> and <TA>* `skos:member` *<B>, then <B> is mapped as a subclass of the OWL class defined for node <A> and an* `owl:ObjectProperty` *is defined between <A> and <B> with* `rdf:ID="<TA>"`".

(7) "*Given two* `Concept` *nodes <B1> and <B2> and a* `ThesaurusArray` *node <TA>, such that <TA>* `skos:member` *<B1>, <B2>, then the constructor* `owl:disjointWith` *must be used between the resulting OWL classes for nodes <B1> and <B2>*".

According with these rules, the following OWL 2 DL conversion is obtained:

```
<owl:Class rdf:ID="#Forecasting">
  <owl:unionOf rdf:parseType="Collection">
    <owl:Class rdf:ID=
      "#Forecasting_by_length">
      <owl:oneOf rdf:parseType="Collection">
        <owl:Class rdf:ID="#Long-term_forecasting">
        </owl:Class>
        <owl:Class rdf:ID="#Short-term_forecasting">
        </owl:Class>
      </owl:oneOf>
    </owl:Class>
    <owl:Class rdf:ID="#Forecasting_by_target">
      <owl:oneOf rdf:parseType="Collection">
       <owl:Class rdf:ID="#Drought_forecasting">
        </owl:Class>
        …
      </owl:oneOf>
    </owl:Class>
  </owl:unionOf>
</owl:Class>
```

Since a `ConceptScheme` can be divided in `ConceptGroup` nodes, each one typically used to represent a self-contained thesaurus identified by a typical subject area (e.g. to represent a domain, theme, microthesaurus or simply a "group"). Thus, in our opinion, all the OWL classes defined from the `it:Concept` nodes belonging to an `it:ConceptGroup` are simply enriched with a datatype property via `owl:topDataProperty`, by considering the concept group as a literal that annotates all the concepts from that group.

Finally, observe that all the SKOS labels associated to a `skos:Concept` node in the graph are preserved in the ontology for completeness. Labels are used, in our framework, primarily to formalize a USE/UF relation between the terms associated to the same concept.

5 Conclusions and Future Work

In this paper we described an alternative application framework to convert thesauri into OWL 2 ontologies. The framework includes a translator module that, starting from an ISO 25964 compliant thesaurus, applies conversion rules to obtain the ontology. From the methodological point of view, along with the classical conversion rules focused on the hierarchical relations, the proposed framework introduced some additional conversion rules for thesaurus entities (e.g. facets, thesaurus arrays, partitive relations) that enable it to provide a stronger formalization.

As stated above, in this paper we focused only on some of the entities introduced or updated by the new ISO standard. In fact, the framework currently does not support some semantic relations, like USE+/UF+, and RT, nor it is designed to handle with interoperability between controlled vocabularies. These features are part of an extension of the framework currently under development.

Moreover, since at the moment very few example of ISO 25964 fully compliant thesauri exists (e.g. AAT), another interesting research stream is related to the necessity for new modules for the standardization of existing thesauri into ISO 25964 compliant ones. This will allow for a complete evaluation of the proposed framework.

Acknowledgements

This work is supported by the "Science & Technology Digital Library" (S&TDL) Project, funded by the Agency for Digital Italy (AgID).

References

1. R. Albertoni, M. De Martino, S. Di Franco, V. De Santis, and P. Plini. EARTh: an Environmental Application Reference Thesaurus in the Linked Open Data Cloud. In Semantic Web Journal, 2012, vol. 5(2): 165-171.
2. ANSI/NISO. Documentation – Guidelines for the construction, format, and management of monolingual controlled vocabularies. 2005. Report ANSINISO Z3919.
3. N. Aussenac-Gilles, S. Despres, and S.Azulman. The TERMINAE method and platform for ontology engineering from texts. In Proceedings of the Conference on Ontology Learning and Population: Bridging the gap between Text and Knowledge, Amsterdam, NL, 2008, pp. 199–223.
4. T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. In *The Scientific American*, 2001, vol. 284(5), pp. 34–43.
5. British Standards Institution (BSI). Documentation - Structured vocabularies for information retrieval. Guide. Exchange formats and protocols for interoperability. 2008. Report BS 8723-5.
6. S.G. Dexter Clarke and M. L. Zeng. From ISO 2788 to ISO 25964: The evolution of thesaurus standards towards interoperability and data modelling. In *Information Standards Quarterly (ISQ)*, 2012, vol. 24, n. 1.

7. T. R. Gruber. Toward principles for the design of ontologies used for knowledge sharing?. In International Journal of Human-Computer Studies, 1995, vol. 43(5–6), pp. 907–928.
8. M. Hepp and J. de Bruijn. GenTax: A Generic Methodology for deriving OWL and RSF-S Ontologies from Hierarchical Classifications, Thesauri and Inconsistent Taxonomies. In The Semantic Web: Research and Applications, E. Franconi, M. Kifer, M. Michael and W. May (eds.) Heidelberg, DE: Springer, 2007, pp. 129–144.
9. D. Kless, L. Jansen, J. Lindenthal, and J. Wiebensohn. A method for re-engineering a thesaurus into an ontology. In Proceedings of International Conference on Formal Ontology in Information Systems (FOIS 2012), Gray, AU, 2012, pp. 133-146.
10. D. Kless and S. Milton. Comparison of thesauri and ontologies from a semiotic perspective. In Proc. of the 6^{th} Australian Ontology Workshop (AOW 2010), T. Meyer, M. A. Orgun and K. Taylor (eds.) Adelaide, AU: Australian Computer Society, 2010, pp. 35–44.
11. A. Isaac and J. De Smedt. ISO 25964 SKOS extension. 2013. Available: http://purl.org/iso25964/skos-thes.
12. International Standard Organization (ISO). Documentation - Thesauri and interoperability with other vocabularies: Part 1, Thesauri for information retrieval, 2011. Report ISO 25964.
13. International Standard Organization (ISO). Documentation. Thesauri and interoperability with other vocabularies: Part 2, Interoperability with other vocabularies, 2013. Report ISO 25964.
14. International Standard Organization (ISO). Documentation – Guidelines for the establishment and development of multilingual thesauri, 1985. Report ISO 5964.
15. International Standard Organization (ISO). Documentation – Guidelines for the establishment and development of monolingual thesaurus, 1986. Report ISO 2788.
16. P. Li and Y. Li. On Transformation from the thesaurus into domain ontology. Advanced Materials Research, 2013, vol. 756 – 759, pp. 2698–2704.
17. S. Mouhim, T. Khalid, C. Cherkaoui, H. Douzi, and D. Mammas. A methodological approach for converting thesaurus into domain ontology: application to tourism. In International Journal of Engineering and Innovative Technology, 2013, vol. 3(5), pp. 315–319.
18. A. Shiri and C. Revie. Thesauri on the Web: current developments and trends. In Online In-formation Review, 2000, vol. 24, no. 4, pp. 273–280.
19. D. Soergel, B. Lauser, A Liang, F. Fisseha, J. Keizer, and S. Katz. Reegineering Thesauri for New Applications: The AGROVOC example. In Journal of Digital Information, 2004, vol. 4(4).
20. B. M. Villazón Terrazas. A method for Reusing and Re-engineering Non-ontological Resources for Building Ontologies. Ph.D. Disseertation, Departamento de Inteligencia Artificial, Facultad de Informática, Universidad Politécnica de Madrid, Madrid, ES, 2011.
21. B. M. Villazón Terrazas and Freddy Priyatna. Building Ontologies by using Reengineering Patterns and R2RML Mappings. In Eva Blomqvist, Aldo Gangemi, Karl Hammar and, Marìa del Carmen Suàrez-Figueroa (eds.). Proceedings of the 3rd Workshop on Ontology Patterns, Boston, USA, November 12, 2012, CEUR WS, vol. 929, pp.109-120, 2012.
22. B. J. Wielinga, A. T. Schreiber, J. Wielemaker, and J. A. C. Sandberg. From thesaurus to on-tology". In Proceedings of the 1st International Conference on Knowledge Capture (K-CAP '01), Victoria, CA, pp. 194–201, October 2001.
23. L. Will. The ISO 25964 data model for the structure of an information retrieval thesaurus. In Bulletin of the American Society for Information Science and Technology, 2012, vol. 38(4), pp. 48–51.