



HAL
open science

Multilingual Thesaurus: The Ontoterminology Approach

Christophe Roche, Luc Damas, Julien Roche

► **To cite this version:**

Christophe Roche, Luc Damas, Julien Roche. Multilingual Thesaurus: The Ontoterminology Approach. CIDOC 2014 (International Committee for Documentation) - Access and Understanding – Networking in the Digital Era, ICOM - International Council of Museums, Sep 2014, Dresden, Germany. hal-01356631

HAL Id: hal-01356631

<https://hal.science/hal-01356631>

Submitted on 26 Aug 2016

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

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

Multilingual Thesaurus: The Ontoterminology Approach.

Authors: **Christophe Roche,**

Luc Damas

Equipe Condillac – Listic

Université de Savoie

France

christophe.roche@univ-savoie.fr

luc.damas@univ-savoie.fr

www.condillac.org

Julien Roche

University of Liaocheng

China

zhuliann@foxmail.com

Abstract:

Making collections accessible to users speaking different languages raises several issues among which the multilingualism problem. Finding equivalent terms as well as mapping thesauri require a common conceptualisation or at least “compatible” conceptualisations.

In that context, ontology of the knowledge engineering, defined as a formal specification of a domain conceptualization, is one of the most promising perspectives for the concept system of thesaurus and terminology. By distinguishing the two conceptual and linguistic dimensions, such an approach enables a more comprehensive understanding of the domain, and a more precise description of collections in relation to the domain knowledge. It also enables to improve search functionalities by using both the linguistic relationships between terms and the logical properties of the relationships between concepts. At last, mapping thesauri can gain benefits from works on ontology alignment.

The ontoterminology approach of thesaurus – an ontoterminology is a terminology whose conceptual system is a formal ontology – has been deployed in the framework of a first European project for multilingual document management system (ASTECH). It is now involved into different European projects dedicated to cultural heritage (Linked Heritage, AthenaPlus), to multilingual knowledge sharing (SIERA) and to multilingual terminology (University of Liaocheng).

This approach is illustrated with *OTe-for-Thesaurus*, a thesaurus editor based on ontoterminology. In accordance with the ISO 25964-1 and 25964-2 Standards on Thesaurus, the *OTe-for-Thesaurus* provides functionalities for creation, editing, and mapping thesauri. Based on the SKOS interchange format, it also allows importing and exporting thesauri. The TMP2, stands for Thesaurus Management Platform version 2, is the version of *OTe-for-Thesaurus* developed for the AthenaPlus European project.

Résumé:

Rendre accessibles des collections à des utilisateurs parlant d'autres langues soulève entre autres le problème de la prise en compte du multilinguisme. La recherche d'équivalents linguistiques tout comme la mise en relation de thésaurus supposent une conceptualisation du domaine commune ou du moins des conceptualisations "compatibles".

Dans ce contexte, l'ontologie au sens de l'ingénierie des connaissances, définie comme une spécification formelle d'une conceptualisation, constitue une des voies les plus prometteuses pour la représentation du système conceptuel des thesaurus et des terminologies. En distinguant les dimensions conceptuelle et linguistique, elle permet une meilleure compréhension du domaine amenant à une description plus précise des contenus au regard des connaissances du domaine. Elle permet d'exploiter pour la recherche d'information aussi bien les relations linguistiques entre termes que les propriétés logiques des relations entre concepts. Enfin, elle fait bénéficier la mise en relation de thesaurus des résultats des travaux menés sur l'alignement d'ontologies.

L'approche ontoterminologique des thesaurus (une ontoterminologie est une terminologie dont le système notionnel est une ontologie formelle) a été mise en œuvre dans le cadre d'un 1^{er} projet européen de gestion de contenus multilingues (ASTECH). Elle est aujourd'hui utilisée au sein de projets liés au secteur culturel (Linked Heritage, Athena Plus) et au partage de connaissances multilingues (SIERA, Université de Liaocheng).

L'article est illustré avec l'environnement *OTe-for-Thesaurus* qui permet, sur la base du modèle ontoterminologique, l'import et l'export de thesaurus au format d'échange SKOS, la création, l'édition et l'alignement de thesaurus dans le respect des normes ISO 25964-1 et -2.

1. Introduction

Opening and linking data in a global and multilingual world are among the most important challenges of the new century [Linked Data] [Hyvonen 2012]. From a thesaurus point of view, they raise several issues among which multilingualism management and mapping thesauri. Making collections accessible to users speaking different languages require a shared conceptualization - or at least "compatible" conceptualisations – of the domain. The notion of concept is become the core element of thesaurus as the last versions of the ISO Standards highlight it [ISO 25964-1], [ISO 25964-2].

In that context, Ontology coming from Knowledge Engineering and defined as a "formal specification of a domain conceptualization" [Gruber 92], i.e. a set of concepts linked by logical relationships, is one of the most promising perspectives for Thesaurus. By distinguishing the two conceptual and linguistic dimensions, such an approach enables a more comprehensive understanding of the domain, and a more precise description of collections in relation to the domain knowledge. It also enables to improve search functionalities (precision and recall) by using both the linguistic relationships between terms and the logical properties of the formal relationships between concepts. At last, mapping thesauri can gain benefits from works done on ontology alignment.

But an ontology is not a thesaurus – neither a terminology – even if ontology can be used for information management and information retrieval [Kiryakov et al. 2005] [Fernandez et al. 2011]. Ontology does not include a linguistic dimension: a concept is an extra-linguistic knowledge. It means that a concept cannot be reduced to terms that designate it. As a matter of fact, the conceptual system does not match with the lexical system [Roche 2007].

It remains to combine Ontology and Terminology in one paradigm for thesaurus purposes. This leads to the notion of *Ontoterminology* (a terminology whose conceptual system is a formal ontology) whose application to Thesaurus opens new perspectives both for information retrieval and mapping thesauri.

The structure of the article is fourfold. The next section focuses on the notion of concept as it is specified by the ISO Standards on Thesaurus and Terminology. The second section is about ontology from the Knowledge Engineering point of view. The third one is dedicated to ontoterminology, its principles, and its application to thesaurus. The article ends with the presentation of *OTe-for-Thesaurus*, an ontoterminology-oriented thesaurus environment.

2. Concept in Thesaurus and in Terminology

2.1 Concept in Thesaurus

The “Overall objective” section of the ISO 25964-1 Standard on Thesaurus starts by these first sentences: “*The traditional aim of a thesaurus is to guide the indexer and the searcher to choose the same term for the same concept. In order to achieve this, a thesaurus should first list all the concepts that might be useful for retrieval purposes in a given domain. The concepts are represented by terms, and for each concept, one of the possible representations is selected as the preferred term.*” The concept is therefore the core element of thesaurus: “*a thesaurus should first list all the concepts [...] in a given domain*”.

Such an importance given to the concept finds its roots in its extra-linguistic nature. A concept is an abstract entity independent of terms that designate it. Indexing on concepts enables information retrieval on the meaning of terms rather than on the terms themselves; i.e. on concepts not on keywords. Hence, multilingualism can be easily managed insofar users share the same conceptualization. It means that different terms in different languages can be attached to the same concept (figure 1). Focusing is done on understanding of the domain, rather than on the way one speaks about it.

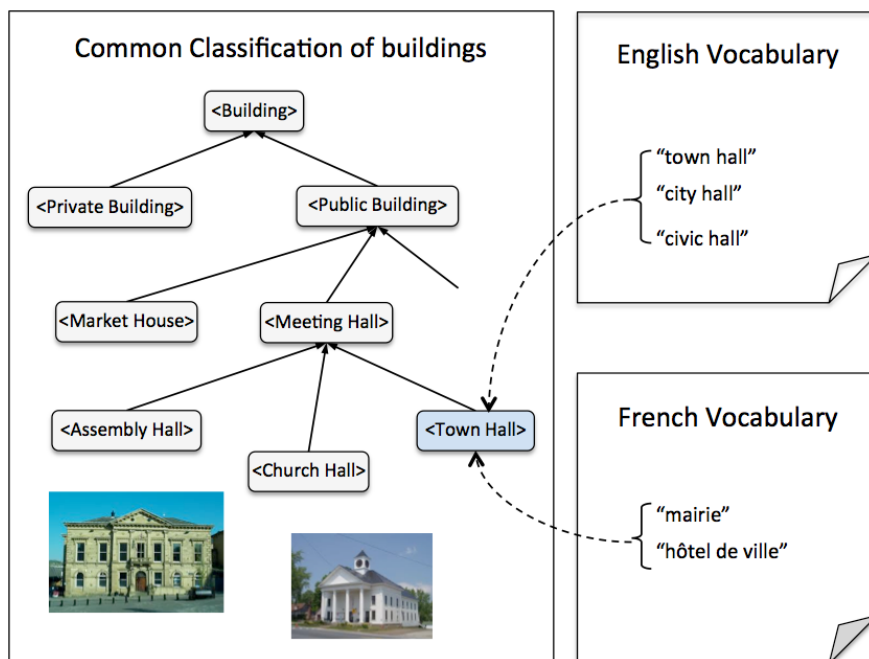


Figure 1. Different vocabularies sharing a common conceptualisation

Although the concept is the central element of Thesaurus, it is simply defined as a “unit of thought” without specifying its characteristics. The notion of “definition” in the literal sense¹ does not exist in ISO 25964-1 – the DEF (for definition) note is only a linguistic explanation.

In Thesaurus, a concept finds its meaning through its relationships with the other concepts. The hierarchical relationships (generic, whole-part, instance) allow to express a subordination level between concepts. The generic relationship links two concepts whose one is more specific than the other, e.g. <Wrist Watch> is a kind of <Watch>, when the whole-part relationship links two concepts whose one is a part of the other considered as a whole, for example <Mechanical Movement> is a part of <Mechanical Watch>. The instance relationship links a concept to one of its individual instances, like /Longines/ is an example of <Watch Brand>. These three kinds of hierarchical relationships are not always distinguished, gathered under the same BT/NT tags (Broader Term/Narrower Term), under the pretext that “The extra work and complexity required in doing so should be balanced against the benefits for the application anticipated” [ISO 25964-1]. Nevertheless, it is important to distinguish them using the BTG/NTG tags (Broader Term Generic/Narrower Term Generic) for the generic relationship, the BTP/NTP tags (Broader Term Partitive/Narrower Term Partitive) for the whole-part relationship, and at last BTI/NTI tags (Broader Term Instancial/Narrower Term Instancial) for the instance relationship. And this for two main reasons. The first one is for understanding the domain – a generic relationship is not a partitive relationship neither an instance relationship, and an instance is not a concept – the second one concerns the *precision* criterion² in information retrieval: for example, looking for a concept does not mean looking for its parts. At last, the associative relationship enables to express associations between concepts which are not related hierarchically.

There is not a clear distinction between concepts and terms in Thesaurus, when they belong to different semiotic systems³, which should not be confused. Even if “The traditional aim of a thesaurus is to guide the indexer and the searcher to choose the same term for the same concept” [ISO 25964-1], the information retrieval relies on concepts and not on terms designating them: “The prime application for a thesaurus is in information retrieval, where the aim is to search for concepts” [ISO 25964-1]. Agreement must be reached on the concepts whatever the languages and whatever the terms – a concept, as a unit of thought, is language independent⁴. Furthermore, some concepts are necessary only for understanding the domain and do not require any descriptors (or terms). For example, management contents about economic information by country require descriptors on concepts representing countries and cities, e.g. Germany and Berlin, but not on the Country concept itself.

¹ A definition of concept – which must not be confused with definition of word – specifies the characteristics, either essential or descriptive, of the concept which uniquely identify it.

² *Precision* in information retrieval is the ratio of the number of relevant retrieved contents to the total number of retrieved contents.

³ A concept, a “unit of thought”, is an extra-linguistic knowledge, when a term, a “word or phrase used to label a concept”, belongs to linguistics. Even if strictly speaking a descriptor is for indexing and retrieval purposes, it is not necessary a lexical unit belonging to the vocabulary of the domain.

⁴ It is clear that is not so simple. A concept depends on cultural background which can lead to different categorizations which are themselves not completely independent of languages – it is the famous Sapir-Whorf’s hypothesis which states that the category of thought depends on category of language [Kay and Kempton 1984].

The confusion between concepts and terms is maintained when the relations between concepts are named as being relations between terms⁵: BT (Broader Term) instead of a more appropriate BC (Broader Concept), BTI (Broader Term Instancial) instead of a more appropriate Instance-Of, etc.

2.2 Concept in Terminology

As in Thesaurus, the couple “concept-term” plays a central role in Terminology. Their definitions in ISO Standards looks quite similar: a concept is a “unit of thought” for Thesaurus when it is defined as a “unit of knowledge” in Terminology [ISO 1087-1], and a term is a “word or phrase used to label a concept” for Thesaurus when it is a “verbal designation of a general concept” in Terminology [ISO 1087-1]. In a similar way, concepts in Terminology are structured into a concept system linked by hierarchical relationships, either generic or partitive, and associative relations. It is the reason why thesaurus and terminology are sometimes confused. For examples, English Heritage⁶ defines a thesaurus as “a structured wordlist used to standardise terminology” and the software environment for thesaurus management developed during the Linked Heritage⁷ European project was called TMP for Terminology Management Platform. But, a terminology is not a thesaurus and vice versa, a thesaurus is not a terminology. The main goal of terminology is not indexing contents for retrieval information but to aim to a “clarification and standardization of concepts and terminology for communication between humans” [ISO 704]. The main activities of Terminology include identifying concepts and concept relations, analysing and modelling concept systems, establishing representations of concept systems, defining concepts, and at last attributing designations (predominantly terms) to each concept in one or more languages [ISO 704]. A concept is then defined as a “unit of knowledge created by a unique combination of characteristics” [ISO 1087-1], and a definition is defined as a “representation of a concept by a descriptive statement which serves to differentiate it from related concepts”. The concept definition relies on a typology of characteristics⁸ whose delimiting characteristics, which are essential characteristics used for distinguishing a concept from related concepts. To sum up, the concept is not defined in Thesaurus whereas the definition of concept is a central activity in Terminology.

2.3 Concept system *versus* Term systems

To conclude this section, it is important to bear in mind that there are two different semiotic systems, a concept system and one or more term systems (one per language). The concept system represents the domain knowledge; a language-independent representation shared by users whatever their mother language. It relies on a precise definition of concept, i.e. on a specification of the characteristics of the concept, either essential or descriptive, which uniquely identifies it and distinguishes it from the other concepts. The concept system is structured according to different conceptual relationships, mainly generic, partitive and

⁵ Relations between terms are linguistic relationships, e.g. hyponymy, when relations between concepts are formal relationships like the subsumption. It is important to bear in mind that these relations do not belong to the same system and that hyponymy cannot be reduced to a linguistic translation of the subsumption relationship.

⁶ <http://thesaurus.english-heritage.org.uk/newuser.htm>

⁷ <http://www.linkedheritage.org/>

⁸ A characteristic is an “abstraction of a property of an object or of a set of objects” [ISO 1087-1].

associative. On the other hand, a term system gathers, for a given language, the words designating the concepts. Even if a term can be considered as a concept label, the term systems do not necessary match with the concept system⁹ since there can be concepts without designations as well as some terms can be polysemic (figure 2). Furthermore, the linguistic relationships between terms cannot be reduced to a linguistic translation of conceptual relationships. For examples, hyperonymy allows multi-hierarchy when subsumption defined by specific difference only allows simple hierarchy.

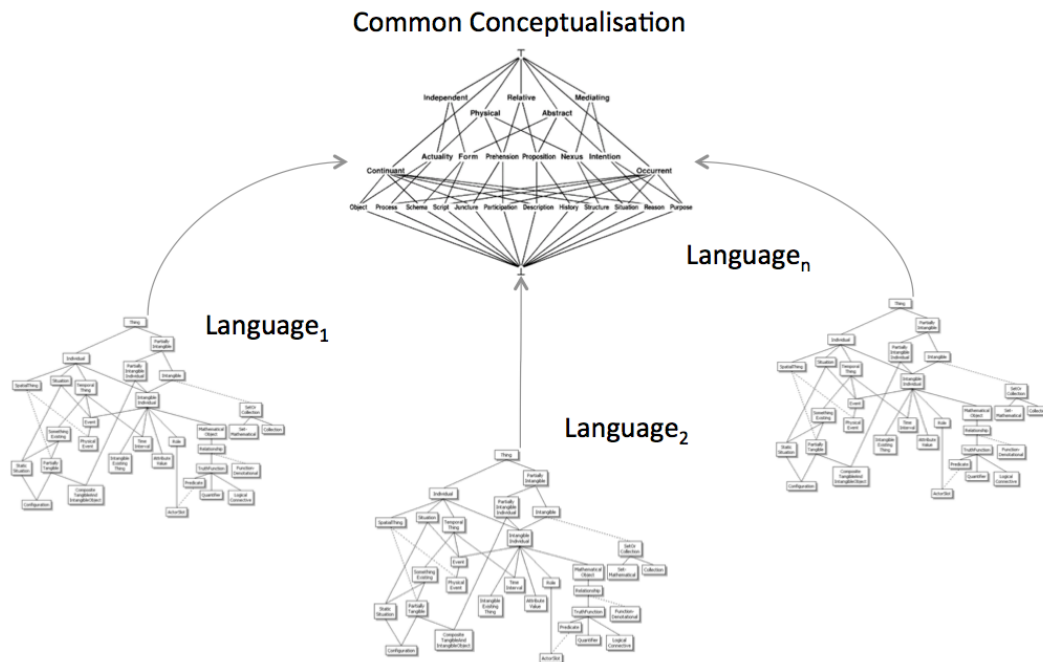


Figure 2: Concept system versus Term systems

3. Ontology

In its Introduction section the ISO 25964-1 Standard states that “today there is a demand [...] for vocabularies that enable inferencing by machines”. Operationalizing thesaurus for IT applications require a computational representation of the conceptual system. It is the definition of ontology from the Knowledge Engineering point of view [Gruber 1992] [Staab et al. 2004].

3.1 Definition

In the context of knowledge representation, an ontology is defined as “a specification of a conceptualization” [Gruber 1992], that is a shared description of concepts of a domain and their relationships expressed in a formal and computer readable language. In other words, an ontology is a system of concepts linked by relationships like “a kind of” (or “is-a”), “part of”, and “associative” relationships; and where each concept is defined in a formal language specifying its characteristics, either essential or descriptive. The two following figures are examples of ontologies. The first one is a domain ontology where concepts are defined by

⁹ Although it is not the purpose of this article, it is important to bear in mind that a term is not a concept name as well as the signified (meaning of a term) is not a concept [Roche 2007].

specific differences¹⁰, e.g. a <Watch> is a ‘portable’ <Timepiece> (‘portable’ is a specific difference).

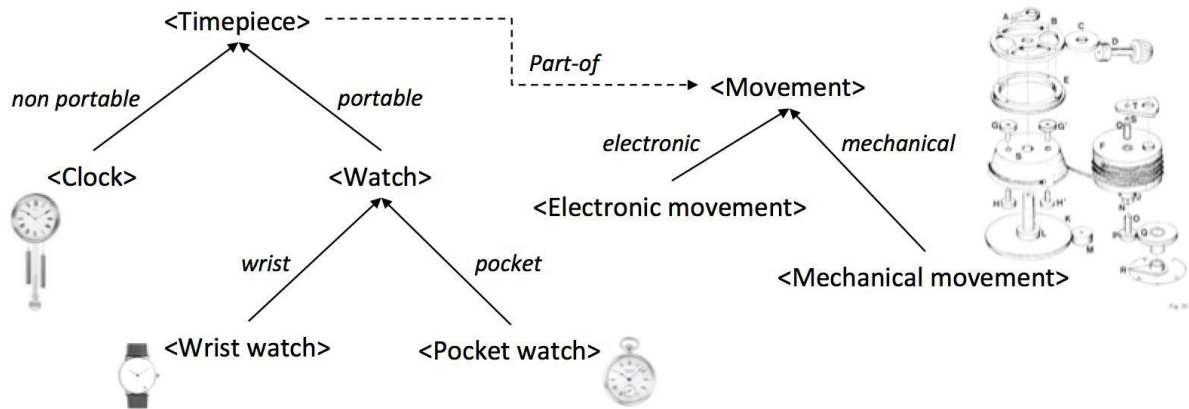


Figure 3. A domain ontology

The second one, Mikrokosmos [Mahesh and Nirenburg 1995], is a top ontology, which aims to describe everything. The representation language is a frame-based language¹¹ where every concept (more often called *class*) is defined by a set of attributes common to all instances of the class. The classes are structured according to the hierarchical subclass relationship (the subclasses inherit the structure (the set of attributes) of its superclasses).

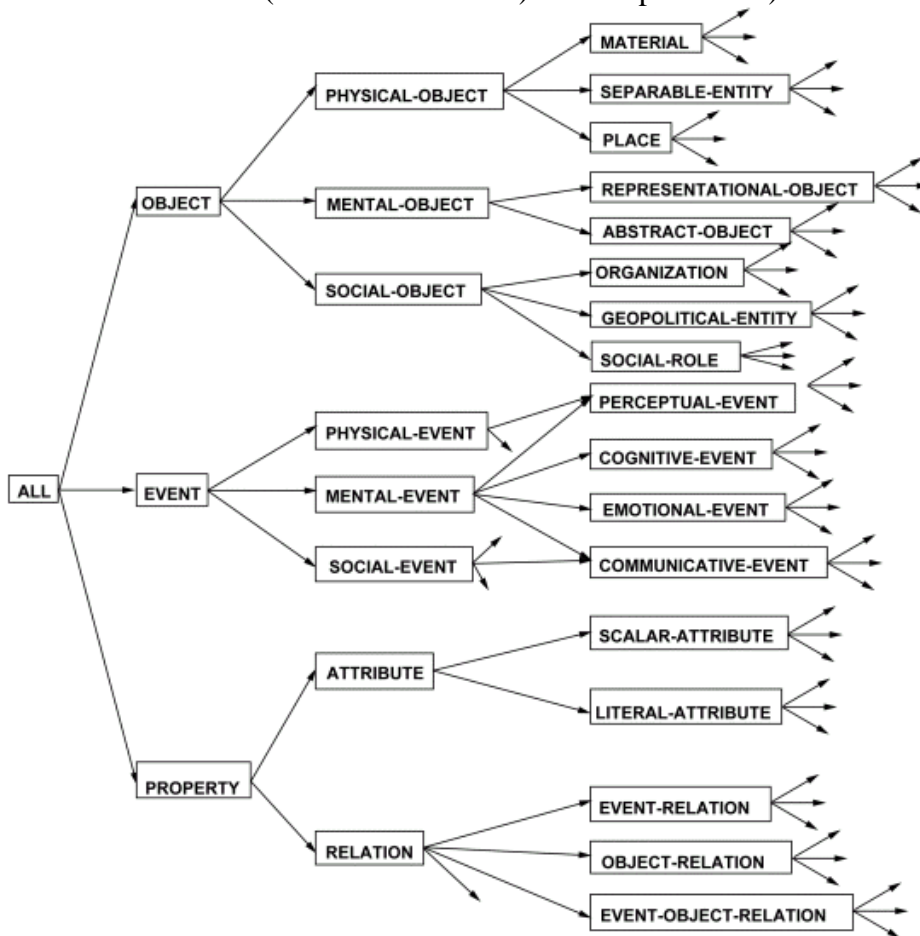


Figure 4. The Mikrokosmos ontology

¹⁰ Corresponding to the Aristotelian genus-differentia definition.

¹¹ Also called schema-oriented language, coming from Artificial Intelligence

3.2 Formal languages

Building an ontology requires a formal and computational language for concept definition. Not all formal languages are of equal merit. They offer neither the same functionalities nor the same guarantees. Languages coming from Artificial Intelligence (see the previous example of Mikrokosmos), in a human readable form, offer interesting characteristics both from an epistemological¹² and computational point of view [Minsky 1974], [Wright et al. 1984], [Brachman et al. 1985]. On the other hand Logic-based languages occupy a special position. More than anything else, it is the definition of concept and relation that makes logic an important language. A concept is a unary predicate, like ‘WristWatch (x)’¹³, and relations are many-place predicates. With their clear, precise syntax and semantics, logic-based languages guarantee coherent and objective definitions as well as sound inference mechanisms. Furthermore, their universal formalism makes them an ideal exchange format. “Description Logic” [Baader et al. 2003], based on the notions of individual, concept (a set of individuals) and role (binary relation between individuals) is an example of such languages, is at the core of the Protégé ontology editor [Protégé].

4. Ontoterminology

An ontology is neither a thesaurus nor a terminology. Nevertheless, ontology constitutes one of the more promising approaches for the representation of the conceptual system of thesaurus and terminology [Roche 2005].

4.1 Definition

Combining ontology and terminology into a single paradigm leads to the notion of *ontoterminology* [Roche 2012], a terminology whose conceptual system is a formal ontology. A double semantic triangle specifies the different involved notions and their relations. So, terms are separated from concepts as well as term definitions written in natural language are separated from concept definitions written in a formal language. It becomes possible to manage the two dimensions, conceptual and terminological, which compose every terminology and thesaurus.

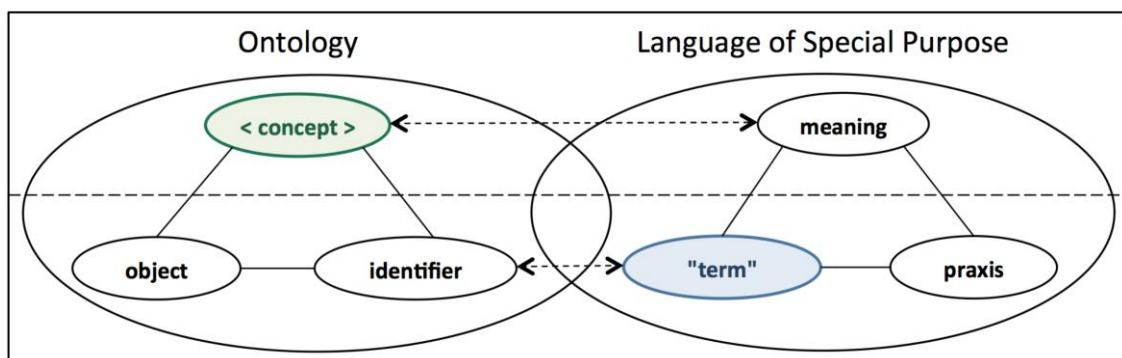


Figure 5. The double semantic triangle

¹² Where objects are described by a set of valued attributes.

¹³ Concepts are well-formed formulas (wff) defined using the logical operators. For example the WristWatch predicate (concept) is defined by the following wff:

$$\text{WristWatch}(x) \stackrel{\text{def}}{=} \text{Watch}(x) \wedge \text{Wrist}(x).$$

4.2 Ontoterminology for indexing and information retrieval

By separating the conceptual dimension from the linguistic dimension, ontoterminology allows a new approach for indexing and retrieval information. Terms are used to describe the contents which will be classified (indexed) under the corresponding concepts. Insofar concepts are extra-linguistic knowledge, i.e. language-independent, it becomes possible to manage multilingualism. The search in a given language returns all the contents that have been classified under the concepts corresponding to the request, whatever the language used for indexing these contents.

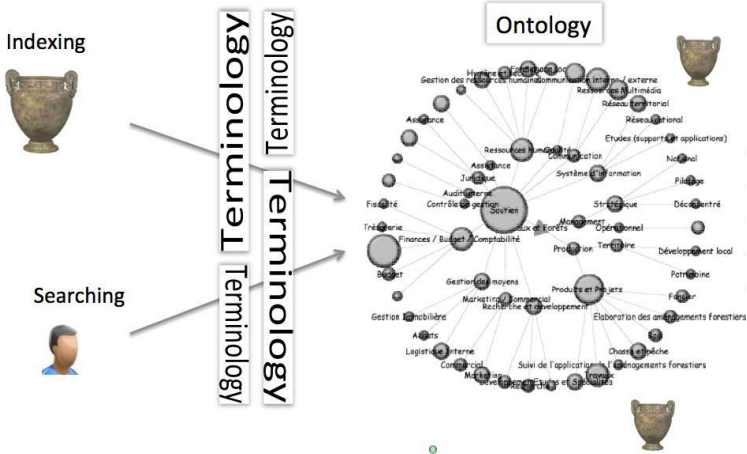


Figure 6. Ontoterminology for indexing and searching

Furthermore, taking into account the logical properties of the relationships between concepts enables to improve the recall and precision¹⁴ criteria.

4.3 Ontoterminology for mapping thesauri

Looking for information among vast collections is one of the big challenges today (especially in a linked and opened data world), hence the need for semantic interoperability between thesauri. It is the scope of the second part of the 25964-2 ISO Standard on ‘‘Interoperability with other vocabularies’’[ISO 25964-2]. The approach remains vocabulary- and language-dependent: ‘‘the principal aim of interoperability between vocabularies is to enable an expression formulated using one vocabulary to be converted to (or supplemented by) a corresponding expression in one or more other vocabularies’’ [ISO 25964-2]. Ontoterminology mapping enables to take into account the two dimensions, linguistic and conceptual, and to benefit from results on ontology alignment. Regarding ontology matching, two kinds of methods are distinguished. The linguistic methods compare concept names, based on the principle that the more the concept names are similar, the more they denote the same concept¹⁵. The semantic methods relies on either an extensional approach, i.e. taking

¹⁴ In information retrieval, *precision* is the ratio of the number of relevant retrieved contents to the total number of retrieved contents; and *recall* is the ratio of the number of relevant retrieved contents to the total number of relevant contents.

¹⁵ Using, for example, string-based techniques. One of the most popular measures of similarity between strings of characters is the Levenshtein distance defined as the minimum

into account the sets of items subsumed by concepts (extensions of concept), or an intensional approach relying on the definition of concepts (attributes) and/or their relationships (Figure 7).

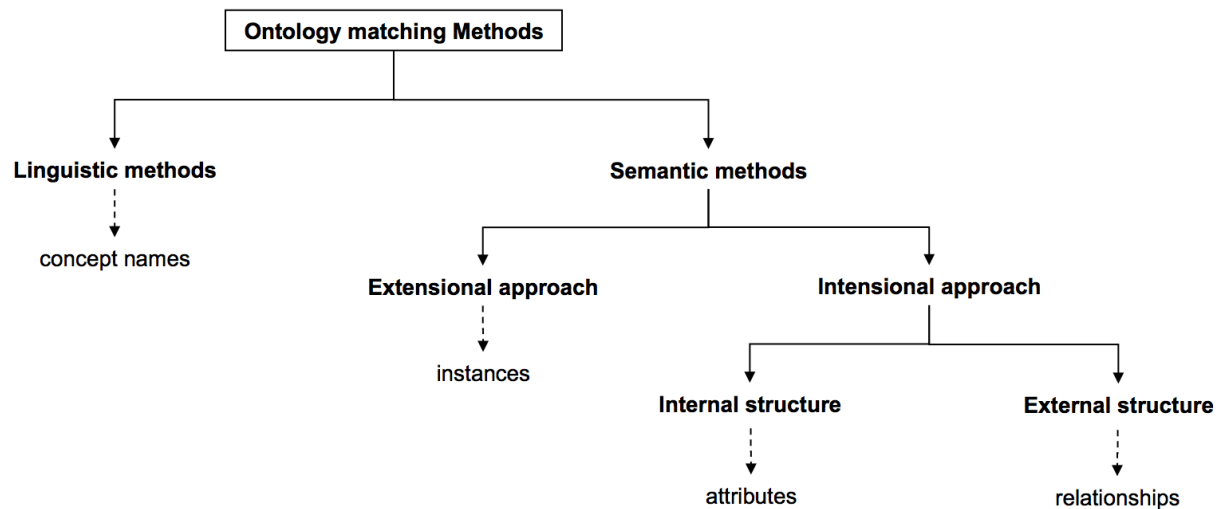


Figure 7. Ontology matching methods

5. OTe-for-Thesaurus

The ontoterminology approach has been used in different projects in Knowledge Management, Multilingual Encyclopaedias and Information Retrieval. *OTe-for-Thesaurus* is one of them.

5.1 Principles

OTe-for-Thesaurus is a Thesaurus editor environment based on *OTe*, the *OntoTerminology Engine* developed by the Condillac Research Group on “Ontology and Terminology” of the University of Savoie. Concepts and individuals, as well as terms and proper names, are explicitly represented in thesauri built with *OTe-for-Thesaurus*. The concepts are structured according to the “is-a”, “part-of” and associative *OTe* relationships whose formal properties¹⁶ guarantee the logical consistency of the conceptual system of thesauri.

In order to be in accordance with the ISO Standards, the *OTe-for-Thesaurus* interfaces use the terminology of Thesaurus. So, the “is-a” relation is called BTG (*versus* NTG), “part-of” is named BTP (*versus* NTP), and the instantiation relationship named BTI (*versus* NTI) (figure 8).

The display layout takes up the double dimension of the ontoterminology. The concept pane displays information about the selected concept (broaders, narrowers, and related concepts); when the term pane displays the preferred term and non-preferred terms for the selected

number of necessary single-character edits to change one word into the other (insert, delete, substitute).

¹⁶ For examples, the *is-a* relationship is a strict-order binary relation (i.e. irreflexive, asymmetric, and transitive), and the *is-a* and *part-of* are acyclic relations.

language (figure 9). The bottom pane displays the notes and external resources linked to the concept (e.g. information from Wikipedia).

The screenshot displays the OTe-for-Thesaurus interface. At the top left is the logo for 'TMP² OTe for THESAURUS' featuring a stylized owl. In the top right corner, the user 'Christophe Roche' is logged in, with a 'disconnect' button. The main content area shows a concept page for 'Time Device' by Christophe Roche. The page includes a description: 'Watch Clock Foliot Automatic Pocket Wrist Electronic Rolex Longines Cartier Tissot Breitling', and creation/modification dates. A 'modify' button is visible. To the right of the concept page is a 'Start mapping' button with a network diagram icon. On the left, a vertical sidebar labeled 'All thesauri' contains a navigation menu with categories like 'Brand' (Breitling, Cartier, Longines, Rolex, Tissot), 'Movement', 'Time Device', 'Movement', 'Clock', 'Watch', 'Pocket Watch', 'Electronic Pocket Watch', 'Mechanical Pocket Watch', and 'Wrist Watch'. Below the sidebar, there are sections for 'Concept' (showing broader and narrower terms), 'Term' (preferred and non-preferred terms), and 'Notes' (including a DBpedia abstract).

Figure 8. OTe-for-Thesaurus

5.3 Interchange Format

The internal representation of a thesaurus in *OTe-for-Thesaurus* is an ontoterminology, a more general representation than thesaurus – a thesaurus is a particular case of ontoterminology.

The external representation of thesaurus in *OTe-for-Thesaurus* (for importing and exporting) relies on a web standard format, in this case SKOS. SKOS (Simple Knowledge Organization System) provides a model for expressing in RDF (Resource Description Framework) the basic structure and content of concept schemes such as thesauri, classification schemes, taxonomies, folksonomies, and other similar types of “controlled vocabulary”. Concepts are labeled with strings in one or more natural languages. It thereby enables a simple form of multilingual labelling.

It is important to notice that *OTe-for-Thesaurus* is independent of SKOS, which is only used as an interchange format. As a matter of fact, SKOS is not a modelling language neither a formal knowledge representation language. It does not guarantee any logical properties about relations. So, the *skos:broader* is not defined as irreflexive in order to enable import ontologies written in OWL¹⁷. Furthermore the *skos:broader* relationship does not distinguish between the *is-a* relationship and the *part-of* relationship, two fundamental relations in ontology building.

Conclusion

The notion of concept has become the central element in Thesaurus as the ISO Standards highlight it. Multilingualism management and mapping thesauri are examples that rely on sharing conceptualization. In this context, ontology of Knowledge Engineering forms one of the most promising pathways for Thesaurus.

By distinguishing and linking the two conceptual and linguistic dimensions, ontoterminology – a terminology whose conceptual system is a formal ontology – enables to take into account multilingualism and allows a concept-oriented indexing and search. It also enables to improve the information retrieval criteria of precision and recall, based on the logical properties of the formal relations between concepts. At last, mapping thesauri gain benefits from works done on ontology alignment.

Ontoterminology is the heart of the *OTe-for-Thesaurus* software, a Thesaurus editor environment, deployed in different European projects among which AthenaPlus for Culture Heritage and SIERA for Multilingual Knowledge Sharing.

Bibliography

Baader Franz, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter Patel-Schneider. (2003). *The Description Logic Handbook*. Cambridge: Cambridge University Press.

Brachman, Ronald J. and Hector J. Levesque. (1985). *Readings in Knowledge Representation*. Los Altos, CA: Morgan Kaufmann Publishers, Inc.

Miriam Fernández, Iván Cantador, Vanesa López, David Vallet, Pablo Castells, Enrico Motta (2011). “Semantically enhanced Information Retrieval: An ontology-based approach”. Web

¹⁷ The *rdfs:subClassOf* statement of OWL is translated in the *skos:broader* relationship. The reflexive *subClassOf* relationship has a set-inclusion meaning.

Semantics: Science, Services and Agents on the World Wide Web. Volume 9, Issue 4, December 2011, Pages 434–452. JWS special issue on Semantic Search.

Gruber, Thomas R. (1992). “A Translation Approach to Portable Ontology Specifications.” *Knowledge Acquisition* 5(2): 199–220. DOI: 10.1006/knac.1993.1008

Eero Hyvonen (2012). *Publishing and Using Cultural Heritage Linked Data on the Semantic Web*. Morgan & Claypool

ISO 704. (2009). *Terminology work – Principles and methods*. Geneva: International Standards Organisation.

ISO 1087-1. (2000). *Terminology work – Vocabulary – Part 1: Theory and application*. Geneva: International Standards Organisation.

ISO 25964-1. (2011). *Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval*. Geneva: International Standards Organisation.

ISO 25964-2. (2013). *Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies*. Geneva: International Standards Organisation.

Paul Kay und Willett Kempton (1984): “What is the Sapir-Whorf hypothesis? *American Anthropologist* 86, 65-79.

Kiryakov, Atanas, Borislav Popov, Ivan Terziev, Dimitar Manov, and Damyan Ognyanoff. (2005). “Semantic Annotation, Indexing, and Retrieval.” *Elsevier’s Journal of Web Semantics* 2(1).

Linked data. <http://www.w3.org/standards/semanticweb/data>

Mahesh, Kavi, and Nirenburg, Sergei. (1995). “A situated ontology for practical NLP”. *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, International Joint Conference on Artificial Intelligence (IJCAI-95), Montreal*.

Minsky, Marvin. (1974). *A Framework for Representing Knowledge*. Memo 306, Massachusetts Institute of Technology, AI Laboratory, June 1974.

OWL. (2004). *Web Ontology Language*. Overview. W3C Recommendation. <http://www.w3.org/TR/owl-features/>

Protégé. *A free, open-source ontology editor and framework for building intelligent systems*. <http://protege.stanford.edu/>.

RDF Schema. (2004). *Resource Description Framework*. W3C Recommendation, <http://www.W3.org/TR/rdf-schema/>

Roche, Christophe. (2005). “Terminologie et ontologie.” *Revue Langages*, 157, March 2005, 48–62. Paris: Éditions Larousse. DOI: 10.3917/lang.157.0048

Roche, Christophe. (2007). "Saying is not Modelling." *Natural Language Processing and Cognitive Science (NLPCS) 2007*, 47–56. ICEIS 2007, Funchal, Portugal, June 2007.

Roche, Christophe. (2012). "Ontoterminology: How to unify terminology and ontology into a single paradigm". Eighth International Conference on Language Resources and Evaluation (LREC 2012). Istanbul, Turkey, May 21–27, 2012.

SKOS (2009). *Simple Knowledge Organization System*. W3C Recommendation.
<http://www.w3.org/TR/skos-reference/>

Staab, Steffen and Rudi Studer (editor). (2004). *Handbook on Ontologies*. Heidelberg: Springer. DOI: 10.1007/978-3-540-24750-0.

Wright, J.Mark, Mark S. Fox, and David Adam. (1984). *SRL/1.5 Users Manual*. Technical report. Robotics Institute, Carnegie-Mellon University.