Anchoring Dutch Cultural Heritage Thesauri to WordNet: two case studies

Véronique Malaisé and **Antoine Isaac**
Vrije Universiteit
Amsterdam
The Netherlands
{vmalaise, aisaac}@few.vu.nl

Luit Gazendam
Telematica Instituut
Enschede
The Netherlands
Luit.Gazendam@telin.nl

Hennie Brugman
Max Planck Institute
for Psycholinguistics, Nijmegen
The Netherlands
Hennie.Brugman@mpi.nl

Abstract

In this paper, we argue on the interest of anchoring Dutch Cultural Heritage-specific controlled vocabularies to WordNet and demonstrate a reusable methodology for achieving this anchoring. We test it on two controlled vocabularies, namely the GTAA thesaurus, used at the Netherlands Institute for Sound and Vision (the Dutch radio and television archives), and the GTT thesaurus, used to index books of the Dutch National Library. We evaluate the two anchorings in terms of number of successful equivalents found and in terms of a concrete use case: we demonstrate how such a mapping can be used in generic alignment scenarios, where concepts from one thesaurus must be aligned to concepts from another vocabulary.

1 Introduction

Mapping Cultural Heritage controlled vocabularies in Dutch to WordNet can be beneficial on different points of views. Cultural Heritage Institutions are the keepers of large collections of data. To optimize the core tasks of indexing and searching through these collections, controlled vocabularies, such as thesauri are usually used. These vocabularies, coming as structured descriptor networks[1], help indexers to select proper concepts for description, and users to formulate queries or to browse collections using the subjects that are effectively mentioned in the collection.

Additionally to these traditional uses, thesauri can also enable a certain number of (semi-)automatic optimization of search process, such as query expansion based on their hierarchical structure. But the available structure might not be rich and regular enough. In fact, it has been shown that a mapping to a richer and sounder terminology, like the English WordNet (Fellbaum, 1998), would enable more sophisticated query expansion or other inferencing possibilities, see (Voorhees, 2006; Hollink, 2006). This will become especially true now that WordNet exists in the form of an RDF ontology (van Assem et al., 2006).

Another aim of the Dutch Cultural Heritage Institutions for which a mapping to WordNet can be useful is the opening of their collections to a wider public: linking their vocabularies to an English resource would open search possibilities to non-Dutch speakers.

The sharing of data across Institutions is also difficult when the metadata attached to the different documents come from different thesauri. This issue can be solved by building equivalence links between the elements from the different vocabularies that are used. With such links, access can be granted to different collections, using either one vocabulary or another (van Gendt et al., 2006). This *vocabulary alignment* problem is comparable to the *ontol-*

---

[1]The typical semantic relationships found between elements from thesauri are Broader Term linking a specialized term to a more general one, Narrower Term, its inverse relationship, and Related to, which denotes a general associative link. Thesauri can also contain lexical information, where the preferred terms used for description are given synonyms (Use and Used for links), as well as general scope notes, giving indexers instructions regarding the use of a term.

ogy matching one, and one could hope to apply here the techniques that were developed by the Semantic Web research community. As found in e.g. (Euzenat, 2006), these alignment methods are quite diverse, and proposed strategies often mix several individual techniques. However, they can be roughly categorized as:

- purely lexical techniques, trying to compare directly the labels of the vocabularies to align;

- structural techniques, evaluating the similarities between the vocabulary elements using the structures these elements are embedded into (e.g. hierarchical links);

- instance-based techniques, looking at the elements that are actually populating the ontologies to infer from the similarities between them correspondences between the concepts they instantiate.

- techniques making use of semantic links contained in some structured background knowledge source, by trying to derive from these links relations between the elements from the original vocabularies.

Amongst these, we are mainly interested in the last ones. Typically, in such approaches, as (Aleksovski, 2006), elements from the vocabularies that are to be aligned are first attached – "anchored" – to the ones of another vocabulary. Then, the corresponding elements in this background vocabulary are compared one another. If a relation is found between these elements[2] then a similar relation can be inferred between the elements from the vocabularies to be aligned. This approach is especially interesting when the lexical coverage of the elements in the vocabularies is low (different Terms for synonyms) or when the vocabularies are quite poorly structured. In such situations – and our cases fit the second category – it is expected the background knowledge will alleviate these shortcomings. The choice of this knowledge is therefore crucial. WordNet, having a rich structure and broad coverage, has been proposed in many existing methods (Giunchiglia et al., 2005; Castano et al., 2005).

---

[2]The reader can turn to (Budanitsky and Hirst, 2006) for an overview of the different methods that have been proposed in this field.

For these three reasons, namely improved Information Retrieval possibilities, multilinguality and vocabulary alignment, we experimented the anchoring of two Dutch thesauri to WordNet. In this paper, we will however focus only on the last usecase: the vocabulary enrichment. Litterature about linking thesauri to WordNet often deals with English controlled vocabularies. We propose in this paper an anchoring method for thesauri in other languages, and experiment it on two thesauri in Dutch, and test its usefullness in terms of possibilities of vocabulary alignment. The remainder of the paper is organised as follows: in the following section (section 2), we present the general anchoring methodology, and the subsequent vocabulary alignment process itself. The anchoring experiment with the two thesauri is described in section 3: first the GTAA usecase (section 3.2), and then the GTT one (section 3.4), as a reusability test. We evaluate the two anchoring processes in section 3.5 and conclude on general reflexions about this method. Then, we show examples of the usefullness of such anchorings in the context of a possible alignment between GTAA and GTT in section 4. We conclude on perspectives to this research (section 5).

2 Anchoring methodology and alignment process

2.1 Anchoring methodology

This anchoring experiment is based on a comparison of lexical descriptions of the thesaurus Terms with the ones of WordNet Synsets, the glosses: WordNet is a lexical database of English, which entries are grouped "into sets of cognitive synonyms (synsets), each expressing a distinct concept."[3] In contrast to (Khan and Hovy, 1997) anchoring experiment, but as the papers they cite (see (Knight and Luk, 1994), the foundational experiment of the kind, for example), we do not compare the Terms from our thesauri to the labels of Synsets, but compare the lexical overlapping in their descriptions.

As the thesauri we focus on in this paper are in Dutch we first need to map their Terms to English descriptions, and ideally translations, to make a comparison with the English glosses. Given the fact

---

[3]http://wordnet.princeton.edu/

that these thesauri are not focused on a specific domain but cover a broad range of topics, we hypothesised that using a general language bilingual dictionary could lead to a good coverage of their content, and give, on top of the definitions – natural language descriptions of the Term's meaning –, terms translations and Part Of Speech tags (POS tags, their grammatical category in the sentence: Noun, Verb, etc.). For each thesaurus term sucessfully defined, the anchoring procedure consists in checking the overlap between the lexical content of the definitions and the one of the different WordNet glosses. The hypothesis is that the closest gloss should give us the identifier of a Synset sematically equivalent to the intended meaning of a GTAA term.

2.2 Vocabulary alignment process

Once the anchoring is performed, the Synsets corresponding to the Terms from the different thesauri are compared, in order to infer from them equivalences between the original sets of Terms. Using this method, alignments that could not be done on the sole lexical equivalence between the two thesauri's terms would hopefully be achieved.

3 Anchoring experiment

3.1 Material: the two thesauri

The thesaurus that constitutes our first case study is the GTAA, a Dutch abbreviation for "Common Thesaurus [for] Audio-visual Archives". It is the controlled vocabulary used for indexing the public radio and TV programs, at the Netherlands Institute for Sound and Vision[4].

The GTAA contains approximatly 160.000 terms, organized in six facets: four facets to describe the program's main topic(s) (*Subject*, *Person*, *Name*[5], *Location*) and two for cataloging meta-information (*Genre* and *Maker name*). The Subject and Genre facets are organised hierarchically. In this experiment, we focus on the anchoring of the Subject facet to WordNet, given the fact that WordNet is a general lexicon and is not likely to have much overlap with the other facets. The Subject facet contains about 3800 PreferredTerms and 2000 Non Preferred

terms. Following the recommendations of the ISO standard, most of the thesaurus' Terms are in plural form.

3.1.1 The GTT

The GTT thesaurus is used to index and retrieve books from the Dutch National Library[6] *scientific collection*. Its Terms are organized into eight categories like *genres*, *geographic descriptors*, *corporate bodies*, etc. We chose to focus on the *"general" descriptors* for this experiment because it is the larger category (35000 Terms out of 65000) and also the closest to the GTAA Subject's themes.

Descriptors from GTT have a preferred form and synonyms – 50000 Dutch terms for the general category. They are also organised hierarchically, but some 20000 descriptors have no parents at all. As in GTAA case, GTT Terms are generally in plural form. Finally, more than 70% of the descriptors were given an English translation, which we will use for evaluation purposes.

3.2 Anchoring the GTAA terms: experiment and evaluation

First step: Finding English definitions for GTAA terms The first step in mapping Dutch terms from the GTAA to WordNet was to select an online dictionary that would cover a significant part of the thesaurus' entries and that would allow automatic queries for these terms. We tested our query set (see below) on the bilingual dictionary LookWAYup [7] and found 2222 results.

The query set consisted in the list of GTAA Preferred terms (3800), Non preferred terms (2000) and their singular forms (3200). These singular forms were computed in the context of a MultimediaN project[8], on the basis of linguistic derivational rules and a manual correction .

Given the fact that most of the thesaurus terms

---

[4] http://www.beeldengeluid.nl

[5] This facet contains names of Companies, groups (like music bands, for example) or organisations.

[6] GTT is a Dutch abbreviation for "GOO keyword thesaurus", GOO refering to the so-called Joint Subject Indexing system used by many libraries throughout the Netherlands.

[7] Built by RES Inc., Canada, online at the URL: http://lookwayup.com/free/.

[8] MultimediaN Project 5 – Semantic Multimedia Access, http://monetdb.cwi.nl/projects/trecvid/MN5/index.php/Main_Page, transformation done by Gijs Geleijnse, from the Philips Research group.

are in plural form, but not all of them[9], and knowing that the dictionary's entries are only standard lemma forms, we first assumed that queries on the dictionary with a plural form would not generate a result, and did not bother separating the singular forms from plural ones in the original thesaurus, but simply added the singular forms to the query set. It turned out that some plural forms returned a result, creating a subsequent noise: some plural forms corresponded to the lemma of verbs and a spelling correction facility provided definitions for plural forms of Terms. We cleaned automatically the last set of errors, and manually the first one, based on POS tag information. In the future, we will avoid duplicates in our the query set.

After cleaning, 1748 Terms had one or more translation in English, POS tag(s) and defintion(s). This low number, compared with the original set of 5800 distinct thesaurus Terms can be explained by the fact that our vocabulary contains numerous mulriwords Terms and also compound entries, both of which are rarely dictionarie's entries. We discuss possible solutions to this shortcoming in section 3.5.

We did then a rough manual evaluation of these candidate definitions. It turned out that some of the definitions were irrelevant for our task: the Dutch *Bij* was associated with the English *Bee* and *Honey bee*, but also with the preposition *by*. We used again the information given by the POS tag to remove these irrelevant definitions: we kept only definitions of Nouns and (relevant) Verbs. Among this last set of definitions, some different ones were still associated with the same thesaurus Term[10]. We used the hierarchical relationship in the thesaurus to check for the intended meaning of these Terms and disambiguate them: for example, *Universiteit* (*University*) had a BroaderTerm relationship with *Wetenschappelijk onderwijs* (*Scientific education*), so its meaning is restricted to the "Educational aspect", and it should not be used to describe TV programs about University buildings for instance. We used this information to restrict the number of valid candidate defintions associated with every GTAA Term, but sometimes the distinction was

hard to make between the different definitions, or no clue was provided in the thesaurus to get an insight of the meaning of the Term: sometimes it did not have any relationship to others, and no explanatory text (Scope Note).

Conclusion of the first step   As a final result, 1655 GTAA terms had one or more English equivalent and one or more candidate definitions. We decided to postpone a more thouroughful validation to the evaluation of anchoring results with WordNet: we decided to keep all candidate definitions and translations that were not obviously incorrect, and check the WordNet anchoring result to see if some further refinment had to be done. The idea was that this process would only work for parts of the definitions, so we wanted to keep as many data as possible.

Second step: Anchoring in WordNet glosses   We stemmed the candidate definitions of GTAA terms and the glosses from WordNet with the Porter stemmer to augment mapping possibilities. Stemming is the operation of reducing words to a root, by removing the final "s" character at the end of a word in plural form, for example. This process can reduce different unrelated words to the same root, so it should be handeled with care, but it requires less resources then a full fledged lemmatizing and helps comparing a larger number of words then on the basis of the graphical forms only. In order to try to map Synset identifiers to GTAA terms, we compared their lexical content. We found out that the definitions of the online dictionary were exact matches with WordNet glosses, thus all defined Terms could be anchored to one or more Synsets.

3.3   Evaluation of the results

We evaluated the number of semantically relevant anchorings for a representative part of the the 1655 GTAA terms that had one or more WordNet anchor: we evaluated 1789 mappings out of 7530. On these 1789 mappings, 85 were not directly usefull for our alignment perspective: 5 out of the 85 non equivalent links were relating Related Terms, 17 pointed to Broader Terms, and the others were mapping a Term with a correct translation, but the translation did not correspond to the intended meaning of the term in GTAA. For example, two mappings were proposed

---

[9]For example, the term corresponding to *Baptism* is in singular form.

[10]In the first result set, 1299 terms had more then one definition.

for *Vrouwen*: *married_woman* and *female_person*, the latter one being the only valid anchoring for our thesaurus. The first cases still provide usefull information for aligning vocabularies, but we took only equivalence relationships into account in this experiment.

An additional evaluation that was also performed on a sample set was to check that Non preferred terms that had a definition pointed to the same Synset as their related Preferred terms. It turned to be correct for the evaluated pairs.

On a qualitative perspective, we found different types of mappings:

- some GTAA terms had more then one translation, all of them pointing to the same synset: this was the confirmation that the mapping from the term to the synset was correct;

- somme GTAA terms had more then one translation, pointing to different, close, synsets: nothing in the thesaurus content could help us making a distinction between the different sysnsets, thus we kept the different possibilities;

- some different GTAA terms pointed to the same Synset and, although they had no relationships in the thesaurus, they had a semantic relationship. This information can be used to enrich the structure of the GTAA.

We can conclude that the anchoring was quite successfull: only 4.7% of the anchorings were incorrect in the test sample, due to multiple senses possibly linked with the terms. Moreover, this process can bring an additional value to the thesaurus structure itself, on top of the possible applications mentionned in the introduction.

3.4 Anchoring the GTT terms: experiment and evaluation

Setting. We carried out the same experiment with the GTT as for the GTAA. However, because GTT had some descriptors translated to English by human experts, we decided that we would assess the performances of our method based on two evaluations:

- a *global evaluation* of the anchoring process, measuring manually the relevance of the produced anchorings from GTT to WordNet;

- an *evaluation of translations*, measuring the relevance of the translations found during the first step of the method by comparing them to the expert ones already contained in GTT.

We have to mention that there was no morphological pre-processing involved in the GTT case, nor did we compute singular forms of the plural-form descriptors found in this vocabulary before trying to translate them.

Results. Out of the 35194 descriptors from the GTT general category, only 2458 were found matching english definitions and translation by the dictionary service we used. If we restrict the evaluation on the set of descriptors for which there was already a translation (25775) we find 2279 descriptors for which we obtained some translation, that is slightly less than 9%. This very low performance is largely due to three problems for which the dictionary did not give any translation of GTT Terms:

- an encoding problem linked to some special characters in Dutch that the dictionary could not deal with and which account for more than 2000 descriptors;

- specialized scientific terms were not defined;

- a problem linked with compound terms: GTT contains complex notions, rendered in Dutch by compound words (e.g. *gebruikersinterfaces* for *user interfaces*), multi words (*Algemene kosten* for *general costs*) or a mixture of the two (*Grafische gebruikersinterfaces* for *graphic user interfaces*).

Whereas the two last problems are clear drawbacks of the methodology and were also encountered in the GTAA usecase, namely the fact that a general language dictionary does not contain domain specific term's definitions and very few multiword entries, the encoding problem is trivial. We discuss the two last problems in section 3.5.

We then investigated the accuracy of the translations found. For that, we first selected the subset of the 179 descriptors that were translated but had not previously been assigned English labels by experts. For this subset, 441 glosses were assigned.

Of these, 172 were correct, concerning 138 descriptors. That is, in terms of information retrieval precision and recall measures, a raw precision of 80% and a recall of 77% in terms of translated descriptors. However, this figures have to be considered in relationship with the number of obtained glosses, where precision falls to 39%[11]. To filter out the wrong glosses, we first applied with significant success the heuristics mentioned in the previous section (use of Broader Terms, POS tags and checking of the singular form of the descriptor), which gave 109 obviously wrong results. The other wrong results were mainly determined using the remaining information found in the thesaurus: the synonyms assigned to the descriptors and their scope notes. In 14 cases, it was necessary to check the books which have been indexed by a descriptor to find out its precise meaning.

In the second part of the evaluation, we compared the translations we found with the ones manually assigned by the experts. First we tested the number of descriptors for which there was an English translation validating the one found by our method, assuming such overlap warranted correctness[12]. We found that 1479 of the 2279 descriptors having both a expert translation and an automatic one had the expert translation confirming one of the found translations, that is a precision of 65% in terms of translated descriptors. When measuring accuracy at the level of found english definitions, we obtained a similar precision, that is 69% (2626 out of 3813 english definitions found had an associated translation which was confirmed by expert translation). Finally, regarding the translations themselves, we found a good recall of 96.4% (our method retrieved 1479 out of 1534 expert translations) while (partial) precision fell at 25.7% (out of 5747 translations found, 1479 were

confirmed by expert translation[13]).

The second problem we faced was of disambiguation: when several English definitions are found, how can we select the proper one? It turned out that both GTT and WordNet have a high level of precision, but they are focused on different matters. It was not a surprise to see that for a same GTT term the dictionary pointed at several meanings that were very close but considered as different Synsets in WordNet. A typical example was the distinction made between the gloss attached to moderation and temperance, "the trait of avoiding excesses", and the one attached to moderateness and moderation, "quality of being moderate and avoiding extremes". Looking at the books indexed by the descriptors which these glosses were attached to, it was not clear whether the indexers systematically considered such a distinction.

However, there were lots of case where, as in the GTAA case, the found definitions were trivially false. To discard these, we could count on the additional information attached to the examined concept, such as the broader and related descriptors. Out of our sample of 179 previously untranslated GTT descriptors for which we found a English definition (441 definitions in total), we found 68 cases for which the Broader Term information helped to discriminate, compared with 6 for Related Term, 6 for synonyms and 15 for scope notes. It is however still unsure whether these different kinds of information can help beyond the situation of manual checking. If translation of broader and related terms is done by the process we have applied, taking into account scope notes would require more efforts. And the poor structure of thesauri such as GTT makes such validations difficult.

As another result, the anchoring method also gave us semantic links that were not present in the original vocabulary, as in the GTAA case. Amongst the translated descriptors from GTT, we found 689 concepts sharing at least one Synset and that were not previously connected by a BroaderTerm or RelatedTerm link. Amongst these, we find interesting matches, such as *gratie* (pardon) and *absolutie*

---

[11]We have no recall in terms of obtained glosses: it is almost impossible to anticipate all the synsets a descriptor could be mapped to, given the difference between the lexical precisions of GTT and WordNet. Furthermore, as will be evoked in our conclusion, in a alignment task such as the one we target, precision is assumed to be more important than recall, which is assumed to be easily highered by turning to several different techniques in a same process.

[12]A manual checking on the first 150 descriptors matching this criteria indeed demonstrated an error rate of 4%, that is 4% of the matching descriptors having no correct definition in their associated glosses.

[13]This figure might actually be caused by the fact that the expert translation, as opposed to the automatic one, provided with a far less important number of synonyms, 90 on our sample.

(absolution) or between *honger* (hunger) and *dorst* (thirst). This potential for enriching thesauri could actually be used to start some kind of positive feedback loop for the anchoring process itself: a richer vocabulary enables for example to use with greater profit the translation selection strategies based on thesaurus structure.

3.5 Conclusion on the anchoring methodology

As mentionned before, a drawback of this method is the fact that there are very few multi-word entries in dictionaries but they compose a large part of thesauri. Previous work about assigning a semantic relationship between a multiword term and its components (see (Ibekwe, 2005)) could be used in order to give elements of solutions to this problem. Using this pre-processing, we could apply our method to the single-word part that corresponds to the generic meaning of the original multiword Term.

From a more conceptual point of view, however, further effort would be needed to adapt our anchoring method – and the subsequent alignment of one vocabulary with the other – to the cases were a descriptor from one vocabulary should be anchored to more than one element from WordNet. We could keep to a simple solution, where we would only try to align concepts based on the links between their simpler components and the fact that these simpler components are often more general than them. More complex heuristics come closer to traditional anchoring problems cases – without translation – and could be solved using existing solutions, as proposed by (Giunchiglia et al., 2005; Castano et al., 2005).

The last problem encountered in the anchoring process was the fact that specialised notions, that also appear in general purpose thesauri, have usually no definition in a general language dictionary. Specialised dictionaries should be used as a complementary ressource.

These different shortcomings reduced the coverage of the anchoring, but our method has are still positive points: the number of obviously wrong anchors was rather low for the found pairs and additional links could be provided for both of the source thesauri.

4 GTAA and GTT alignment via the WordNet anchoring: qualitative evaluation

In this section, we present some examples illustrating the kind of alignment result one can expect from a proper anchoring of our Dutch controlled vocabularies to WordNet. First we can confirm alignments that would be obtained based on the equality of Dutch labels, as in: *gtaa:arbeiders*, aligned to *gtt:arbeiders*. In some cases, though, a first stemming or lemmatizing process would be needed to achieve alignment, as in the example of: *gtaa:bekeringen* and *gtt:bekering* (*Conversion*, respectively in plural and singular form), or *gtaa:biljart* and *gtt:biljartspel*[14] (*Billiard* and *Billiard game*). The more interesting cases are the ones of large semantic overlap but few lexical one, as in the alignment of *gtaa:plant* and *gtt:begroeiing*, via the WordNet *flora* Synset. *Begroeiing* is actually semantically related in the GTT to the descriptor *Planting*. Here the translation process compensates for the lack of lexical coverage in the respective vocabulary, which precisely correspond to one of the traditional features the background knowledge based techniques usually boasts. We can also derive general conceptual similarity relationship based on the overlap between glosses, such as the one between *gtaa:drank* and *gtt:alcohol*, which are not direct matches but for which our method has found some common glosses like "an alcoholic beverage that is distilled rather than fermented".

5 Conclusion and perspectives

Our experiments showed that the partial anchoring of some Dutch controlled vocabularies can be done via a bilingual dictionary, even though there is an obvious loss in information: not 100% of a thesaurus content can be mapped to a general language bilingual dictionary, and a preprocessing of multiword and coumpound thesaurus entries has to be done. Nevertheless, a significant part of the GTAA thesaurus could be anchored, and with an extension to the method this could be true for GTT too. The extension would consist in, besides the multiword and

---

[14]Notice that substring-based matching could also give these results, but this method is usually very noisy during alignment process and therefore to be used cautiously.

compound words processing, the fact to take into account also specialised dictionaries and to have a closer look at methodologies for anchoring a thesaurus Term to multiple WordNet Synsets. We plan to test these extensions in future experiments, and hope to obtain a better coverage of the thesauri.

We evaluated these anchorings against the use case of vocabulary alignment and underlined the potential gains on test examples. Even if the number of results given by the current implementation of our method is quite low, the reader should nevertheless be aware that the process can already prove practically useful as is. Indeed, as already mentioned, it suggests new relationships between Terms of the source thesauri. Moreover, proposed strategies in the alignment field often advocate using several methods at the same time. Contributions from different techniques can be used to proceed with some crossvalidation if they overlap, or to provide with larger number of candidate for further (semi-)automatic selection. In such a setting, every contribution of candidate links is welcome. In this respect what is useful here is the ability of a WordNet-based method to provide with results that could not be obtained with other techniques because of the lack of explicit semantic information, such as hierarchical structure, in the original vocabularies.

Finally, as mentioned in the introduction, there are two other motivating use cases that we plan to experiment with. The first one is the way a mapping with WordNet can enhance the existing access to document collections of the Dutch Cultural Heritage Institutes, e.g. by providing with query refinement services. The second one, more prospective, concerns the interest of the bilingual access enabled by this mapping: to which extent such a feature could be used to effectively widen the audience of Dutch collections?

Acknowledgements

References

Fellbaum C. 1998. *WordNet An Electronic Lexical Database*. MIT Press.

van Assem M., Gangemi A. and Schreiber G. 2006. *RDF/OWL Representation ofWordNet*. W3C Working Draft, 19 June 2006. http://www.w3.org/TR/wordnet-rdf/

Hollink L. 2006. *Semantic annotation for retrieval of visual resources*. PHD Thesis, Vrije Universiteit Amsterdam.

Voorhees E. 2004. *Query expansion using lexical-semantic relations*. 17th International ACM/SIGIR Conference on Research and Development in Information Retrieval, 61–69.

Euzenat J., coordinator. 2004. *State of the art on ontology alignment*. KnowledgeWeb Deliverable 2.2.3.

Castano S., Ferrara A. and Montanelli S. 2005. *Matching Ontologies in Open Networked Systems: Techniques and Applications*, volume 5. Journal on Data Semantics (JoDS).

Giunchiglia F., Shvaiko P., and Yatskevich M. 2005. *Semantic Schema Matching*. 13th International Conference on Cooperative Information Systems (CoopIS 2005).

Aleksovski Z. 2006. *Matching Unstructured Vocabularies using a Background Ontology*. 15th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2006).

Budanitsky A. and Hirst G. 2006. *EvaluatingWordNet-based Measures of Lexical Semantic Relatedness*, volume 32(1). Computational Linguistics, 13–47.

Ibekwe-SanJuan F. 2005. *Clustering semantic relations for constructing and maintaining knowledge organization tools*. volume 62 (2). Journal of Documentation, Emerald Publishing Group, 229–250.

Khan L. R. and Hovy E. 1997. *Improving the Precision of Lexicon-to-Ontology Alignment Algorithm*. AMTA/SIG-IL First Workshop on Interlinguas, San Diego, CA, October 28.

Knight K. and Luk S. 1994. *Building a Large-Scale Knowledge Base for Machine Translation*. In Proceedings of the AAAI-94 Conference.

van Gendt M., Isaac A., van der Meij L. and Schlobach S. 2006. *Semantic Web Techniques for Multiple Views on Heterogeneous Collections: a Case Study*. 10th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2006), 426–437.