Automatically Annotating Text with Linked Open Data

Delia Rusu, Blaž Fortuna, Dunja Mladenić
Artificial Intelligence Laboratory, Jožef Stefan Institute, Ljubljana, Slovenia
name.surname@ijs.si

ABSTRACT

This paper presents and evaluates two existing word sense disambiguation approaches which are adapted to annotate text with several popular Linked Open Data datasets. One of the algorithms is based on relationships between resources, while the other one takes advantage of resource definitions provided by the datasets. The aim is to test their applicability when annotating text with resources from WordNet, OpenCyc and DBpedia. The experiments expose several shortcomings related to the current approaches, which are mostly connected to overfitting the datasets. Based on the findings, we indicate future work directions regarding text annotation with Linked Open Data resources, which can bridge these shortcomings.

Categories and Subject Descriptors

I.2.7 [**Natural Language Processing**]: *Text analysis*.

General Terms

Algorithms, Design.

Keywords

Linked Open Data, text annotation, word sense disambiguation.

1. INTRODUCTION

The Linked Open Data (LOD) numbers over 200 datasets, more than double compared to a year ago, spanning domains such as media, geography, publications, life sciences, etc, incorporating several cross-domain datasets. This is an important source of structured data, which so far has been employed for building Linked Data browsers, search-engines, or other domain-specific applications such as semantic tagging and rating ones [3]. In this paper, we are looking at ways to link structured and textual information on the Web, by annotating text with LOD resources, and as such moving closer to better machine text understanding. As a scenario, one can consider news articles as a source of unstructured, yet up-to-date knowledge, which can be linked with the LOD, providing additional context for entities (e.g. people, organizations) and events described in the articles.

The task of annotating text with LOD resources is closely related to word sense disambiguation (WSD), defined in natural language processing as identifying the meaning of words in a given context. Three main approaches have emerged for determining word senses [14]: *supervised* – employing machine learning methods for training a classifier on sense-annotated data, *unsupervised* – relying on clustering of word occurrences and *knowledge-based* which exploits knowledge resources like dictionaries, ontologies or thesauri. Up to now, most of the effort has been directed to identifying WordNet senses for ambiguous words, using Wikipedia for building sense-tagged corpora or extending WordNet with relationships extracted from Wikipedia.

LDOW'11, March 29, 2011, Hyderabad, India.

Another area related to annotating text with LOD resources is ontology matching (establishing mappings between two ontologies). The task is addressed by using the ontology content (e.g., concept labels and descriptions) [9] or, more recently, ontology structure (relations between the ontology concepts) [7]. In text annotation, as addressed in this paper, instead of matching two ontologies, we are dealing with the text to be annotated on one side, and the ontology, on the other side. Two main differences reside in the lack of ontology-like structure in text together with the ambiguity of words, which often depends on the context.

Up to now, word sense disambiguation approaches have been mostly developed and tested using WordNet. In this work we investigate the applicability of two existing WSD approaches to annotating text with several other LOD datasets. Due to the scarcity of sense-annotated corpora (mainly available for WordNet), we consider two knowledge-based algorithms, which we adapt to the task of automatically annotating text with LOD. The first approach, based on the Page Rank algorithm relies on the relations between resources defined within a dataset, while the second, called Context Similarity, takes advantage of the human-readable description of a resource coupled with local relationship structure. Our experiments are conducted on three LOD datasets: WordNet, OpenCyc and DBpedia, containing common features required by the above mentioned WSD approaches: resource descriptions and a rich set of relationships between resources. The experimental results reveal the obstacles when attempting to generalize the current state-of-the-art knowledge-based WSD approaches across LOD datasets.

The paper is structured as follows: Section 2 briefly lists related work, while Section 3 describes the algorithms for annotating text with Linked Open Data used in the paper. Section 4 elaborates on the specifics of the datasets that we used in our experiments, and Section 5 discusses the results obtained thus far. The last part of the paper is dedicated to conclusions and future work.

2. RELATED WORK

Supervised approaches to WSD have recorded better results in the past evaluation workshops, compared to their unsupervised and knowledge-based counterparts. The best performing system [5] in the SemEval 2007 course grained English all-words task [13] was a supervised approach based on the Support Vector Machine algorithm and was trained on several corpora: an English-Chinese parallel corpora, SemCor (a subset of the Brown corpus, where words are syntactically and semantically tagged), and the Defence Science Organisation corpus.

Among the knowledge-based approaches, attempts have been made to exploit the structure of the concept network, devising either similarity measures (most of them implemented in the WordNet::Similarity package [16]) or taking advantage of the graph structure of the knowledge base. Patwardhan et al [15] took the context of the word into account and performed disambiguation based on various similarity measures. Medelyan and Legg [10] used similarity and disjointness and disambiguated Wikipedia articles into concepts from the Cyc ontology. Previous attempts to annotate text with Cyc used its taxonomic knowledge as well as limited context (at the sentence level) or tried to formally describe the structure of the document, build hypothesis (interpretations) and try to reason based on them [6]. Mihalcea et al. [11] and Agirre and Soroa [1] adapt the PageRank algorithm to disambiguate words based the structure of the WordNet graph. Navigli and Velardi [12] propose the Structural Semantic Interconnections (SSI) algorithm, which further develops lexical chains (sequences of semantically related words) by encoding a context free grammar of valid semantic interconnection patterns. Ponzetto and Navigli [17] have shown that, when having a high-quality knowledge base (they enriched WordNet with relations from Wikipedia) simple knowledge-based approaches compete with state-of-the-art supervised approaches.

3. ALGORITHMS FOR ANNOTATING TEXT WITH LINKED OPEN DATA

The task of annotating text with Linked Open Data is defined as follows: given a word or an n-gram w_a from a text fragment which is to be annotated, the aim is to identify the corresponding LOD resource from a set of candidate resources that best matches the word/n-gram in its context (text fragment). In our implementations we consider as candidate resources for w_a all the resources defined by the dataset with w_a as their *rdfs:label*.

In this paper we are focusing on three LOD datasets with cross-domain coverage: WordNet, OpenCyc and DBpedia. We exploit two characteristics of these datasets. Firstly, we take into account their structure, which is based on the relations between the resources. Secondly, where available, we consider the human-readable description of a resource generally found under *rdfs:comment*. Based on these characteristics, we have adapted two text annotation algorithms: a structure based one – *Page Rank*, and a content-based one: *Context Similarity*.

3.1 Page Rank

PageRank [4] is a well-known algorithm used for ranking the vertices in a graph representing structure of Web pages. It has been previously applied to word sense disambiguation into WordNet [11, 1] by building a graph representing the text to disambiguate and identifying relationships between the vertices (which describe the words in the text fragment).

The LOD datasets also exhibit a graphical structure based on the relationships between resources, for e.g., between an instance and a class described by *rdf:type*, between a class and its superclass described by *rdfs:subClassOf*, and other dataset-specific relations such as, antonym, meronym and holonym in WordNet and broader-term relation in OpenCyc. In order to apply the PageRank algorithm in the case of a LOD dataset, we first build a graph of the dataset $G(V,E)$ where the vertices represent all the dataset resources and the edges are the relationships between these resources. As a next step, we identify all the candidate resources for all the words belonging to the text fragment, which are to be annotated. The main difference is the initialization step, which consists of setting the graph vertices to either of the values 0, if the vertex does not represent a candidate resource, or $1/R$, with R being the total number of candidate resources. The PageRank value for each vertex i ($PR[V_i]$) is computed using the formula:

$$PR[V_i] = \frac{1-D}{N} + D \cdot \sum_{V_j \in InEdges(V_i)} \frac{PR[V_j]}{OutEdges(V_j)}$$

where N stands for the total number of vertices in the graph and the damping factor D is set to 0.85. The algorithm converges when the difference between the previous and current PageRank values for a vertex is below 10^{-15} (the numerical error for double precision). Finally we select the candidate resource with the highest PageRank score for each word w_a.

3.2 Context Similarity

Many of the LOD datasets have textual definitions attached to resources: DBpedia, Freebase, OpenCyc, WordNet, etc. For example, in DBpedia, this definition is the human-readable description of a resource found under *rdfs:comment*. Based on this remark, we adapted the Extended Gloss Overlaps method used in word sense disambiguation and introduced for WordNet in [2] by Banerjee and Pedersen. The method scores each candidate resource based on the word overlap between the context around word w_a and the human-readable descriptions for a candidate resource, together with its neighbouring resources (directly connected to the resource under consideration). The candidate resource with the highest score for each word/n-gram will be selected as the annotation. The context of w_a is represented by the surrounding words in the text fragment, for e.g. all words from the same sentence or paragraph. The overlap between a candidate resource and the word context is computed using the measure of cosine similarity (sim_{cos}), a standard text mining approach to compute the similarity between documents. It is defined between two bag-of-words vectors A and B as

$$sim_{cos}(A,B) = \frac{A \cdot B}{\|A\|\|B\|}.$$

We summarize the algorithm for computing the similarity between a word/n-gram w_a and a resource in Figure 1. We start by determining the neighbourhood resources, and the context of w_a. Then, for each resource from the neighbourhood resources, we compute the cosine similarity between the bag-of-words representation of the resource definition ($NR[i]$) and the context of w_a.

```
ContextSimilarity(resource, w_a) returns Similarity
  Similarity = 0
  NR = GetNeighborhoodResources(resource)
  CW = GetContext(w_a)
  for i = 1 to Size(NR) do
     CS = sim_cos(NR[i], CW)
     Similarity = Similarity + CS
  end for
return Similarity
```

Figure 1 The *Context Similarity* algorithm. Given a candidate resource and a word (w_a), the algorithm computes the similarity between the resource and word's context.

4. DATASETS

4.1 WordNet

WordNet[1] is a lexical database of English, comprising nouns, verbs, adjectives and adverbs grouped into synsets. There are two RDF/OWL representations of WordNet in the LOD: WordNet

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

(W3C) models WordNet version 2.0, while WordNet (VUA) models the latest version, 3.0.

In order to evaluate the aforementioned approaches, we considered two scenarios: in the first one we opted for a dataset used in the word sense disambiguation tasks – SemEval 2007 Task 7: Course Grained English All Words. In our second scenario we used crowdsourcing for evaluating the Context Similarity measure. The SemEval series of evaluation workshops are regarded as a framework for comparing state-of-the-art WSD systems. In Task 7 of SemEval 2007, the participating systems are provided with a corpus consisting of 5 texts annotated with WordNet 2.1 senses from sources like the Wall Street Journal (D_1 to D_3), a Wikipedia entry on computer programming (D_4) and an excerpt from a fiction book (D_5). Out of the 2269 annotated words, 1591 are polysemous (have more than one meaning). The results of our experiments are summarized in Table 1. The precision, recall and F measure are the same in this case, as the system annotates all words and only one annotation/word is yielded. For WordNet, there is a most-frequent-sense baseline (which few knowledge-based systems outperform) obtained by always choosing the first sense of a word in WordNet. Notice that this first sense in WordNet is also the most frequent sense of that word, as measured in a corpus called SemCor. In all but D_4, which is domain specific, the baseline performs better than the proposed approaches; the F measure of the baseline is as high as 85.60% for D_1, and the lowest for D_5 – 74.20%; for D_4 the F measure is 75.19%. The best supervised system obtained an F measure of 82.50%, while the best unsupervised system recorded 77.04%.

Table 1 WordNet evaluation results (F measure, in %) for the SemEval 2007 Task 7 corpus. CS is the Context Similarity algorithm, while PR is the PageRank algorithm.

	D_1	D_2	D_3	D_4	D_5	*All*
CS	75.27	77.84	72.80	77.84	72.46	75.50
PR	74.19	74.67	73.60	73.71	71.01	73.51

In the second scenario, we used CrowdFlower[2], a labor-on-demand platform where one can assign tasks to a number of non-expert workers. Workers were assigned the task to select the correct human-readable description of a resource (annotation) from a list of possible descriptions. The example below shows a marked word for annotation (*painter*) together with its possible WordNet resource descriptions, from which the second one represents a correct annotation. The 4th option *(none of the above*) can be chosen in case none of the candidate resources represents a correct annotation.

It must have been about the same time when Fra Angelico was covering the walls of San Marco with his angel pictures, that a very different kind of ***{painter}*** *was working the Carmine church in Florence.*

1. *painter : a worker who is employed to cover objects with paint*
2. ***painter : an artist who paints***
3. *catamount, cougar, felis_concolor, mountain_lion, painter: large American feline resembling a lion*
4. *none of the above*

This scenario serves as a performance baseline for WSD methods developed for WordNet, when evaluated using crowdsourcing. The evaluation was done on a corpus composed of 325 words from the SemEval 2007 Task 7, which is shown in Table 2. We used some of the words as control words, that is, the correct annotations were reported to the CrowedFlower system which used them to filter out the workers with bad performance. Annotations from at least 5 workers were obtained for each word and as such we distinguish two cases. In the aggregate case the predominant annotation for a word was selected from the set of all annotations. In the non-aggregate case each obtained annotation was considered separately.

The inter-annotator agreement is usually considered an upper bound for WSD systems [8]; in our experiments it ranges between 59.61% for D_1 to 66.89% for D_5. The results when using the Context Similarity measure are reported in Table 3. We observe an expected decrease in the F-measure, as we are dealing with non-expert annotators, while the Context Similarity method results are close to the inter-annotator agreement.

Table 2 A subset of 325 words from the SemEval 2007 Task 7 corpus composed of 95 control words, and 230 test words, which was employed in the WordNet annotation task, using CrowdFlower.

	D_1	D_2	D_3	D_4	D_5	*Total*
Control	11	14	20	34	16	95
Test	30	41	56	71	32	230

Table 3 WordNet evaluation results (F measure, in %) when using CrowdFlower, a labor on demand platform. As annotation algorithm, we used Context Similarity.

	D_1	D_2	D_3	D_4	D_5
Aggregate	56.66	60.98	51.79	56.34	68.75
Non-aggregate	67.57	61.72	57.80	59.78	69.35

4.2 OpenCyc

OpenCyc[3] is the open source version of Cyc, a common-sense knowledge base, covering the top 40% of the complete Cyc knowledge base. OpenCyc is also available as a downloadable OWL ontology, and in this paper we refer to the 2008 version.

For the first experiment on OpenCyc we extracted a subset of 177 words from the SemEval 2007 Task 7 corpus, as shown in Table 4. The correct annotations were provided by CrowdFlower workers, and the workers were selected using the WordNet control words, as described in the previous sub-section.

Table 4 A corpus used to evaluate OpenCyc annotations, comprised of 177 words. The table shows the distribution of the extracted words over the five documents.

	D_1	D_2	D_3	D_4	D_5	*Total*
OpenCyc	25	30	43	50	29	177

Table 5 OpenCyc evaluation results (F measure, in %) when using CrowdFlower. As annotation algorithm, we used Context Similarity.

	D_1	D_2	D_3	D_4	D_5
Aggregate	24.00	36.67	27.91	42.00	34.48
Non-aggregate	41.29	37.42	29.95	37.01	42.08

The obtained results are listed in Table 5. After looking at the results, our first assumption was that non-expert annotators found

[2] http://crowdflower.com/

[3] http://sw.opencyc.org/

it difficult to identify the correct OpenCyc annotations based on the resource human-readable definition. For example, the word *boy* can be resolved in OpenCyc to the following three resources:

1. *The collection of all boys (juvenile male humans). A type of young animal and male person.*
2. *The collection of male children, male kids about 12 years of age or less.*
3. *(son PAR MALE) means that MALE is one of the sons (male children) of PAR. MALE could be a child of PAR by birth, by adoption, by marriage (e.g., if PAR had married a biological parent of MALE), or by some other social arrangement.*

Therefore, in our second OpenCyc experiment (see Table 6), we extracted a similar subset comprised of 50 words from D_3 with more than one candidate resource, which were manually annotated with OpenCyc resources by 2 expert annotators (A_1 and A_2). The inter-annotator agreement was 74.00%. The results turned out to be more or less the same as the ones obtained via crowdsourcing and reflect the difficulty of directly transferring WSD algorithms from WordNet to OpenCyc.

Table 6 OpenCyc evaluation results (F measure, in %) based on manual annotations provided by A_1 and A_2.

	A_1	A_2
Context Similarity	24.00	32.00
PageRank	32.00	48.00
Random	22.00	26.00

4.3 DBpedia

The DBpedia[4] dataset is based on a cross-domain ontology with most concepts representing places, persons, work, species, and organizations. The ontology was mostly extracted from infoboxes in Wikipedia. Each DBpedia resource is described by a label, a short and long English abstract, a link to corresponding Wikipedia page and a link to the image representation of the resource, when available.

Table 7 DBpedia evaluation results (F measure, in %) based on manual annotations provided by A_1.

	D_3
Context Similarity	17.86
PageRank	21.43
Random	14.28

For evaluation we randomly extracted a subset of words from D_3 with more than one candidate resource, similar to the OpenCyc experiments. These words were manually annotated with DBpedia resources by one annotator and the results are shown in Table 7. In the subset of D_3 all but two of the words to disambiguate are general, non-entity words. However, due to high emphasis of entities in DBpedia, often none of the candidate resources was correct. For example, in the sentence:

In France Americans it seems have followed Malcolm Forbes's hot-air ***lead*** *and taken to ballooning heady way.*

the correct annotation for the word *lead* would be:

With this pronunciation, 'lead' generally means 'first', 'ahead', or 'guide'.

[4] http://dbpedia.org/About

However, this meaning only appears in the disambiguation page of *lead* in Wikipedia and does not have a corresponding resource in DBpedia. On the other hand, DBpedia lists several other possible resource candidates: *lead* as the chemical element, *Lead* as a Japanese Hip-Hop Group, etc., all in all 32 candidate resources for *lead*.

5. DISCUSSION

The LOD datasets that we considered in this paper exhibit common features: the existence of human-readable resource descriptions as well as a relationship structure between resources. However, when applying text annotation approaches (adapted from state-of-the-art WSD algorithms) which are based on these common characteristics, the obtained results are not comparable. In this section we consider several characteristics of the employed datasets which serve to better understand the obtained results.

Firstly, each of the three datasets under consideration was developed for a different purpose, which has to be taken into account when developing annotation algorithms for these datasets. WordNet is a dictionary-based taxonomy with a good coverage of the common English lexicon and good, dictionary like, descriptions. OpenCyc is a common-sense knowledge base primarily developed for modeling and reasoning about the world. DBpedia is an effort to extract structured information from Wikipedia, has only a small upper level layer and, unlike WordNet and OpenCyc, has a rich set of instances (named entities such as places, people, and organizations).

As such, WordNet has the highest ratio of covered words, given the texts used in experiments. This is due to its dictionary-like nature and the fact that candidate resources correspond directly with the possible word meanings. On the other hand, OpenCyc contains many resources or distinctions between resources which are important from the reasoning perspective (e.g. the three candidate resources for the word *boy*) but are hard to disambiguate by looking at the word and text alone. The differences between the three candidates would only become apparent when faced with distinct reasoning tasks, requiring various representations of the sentence at hand. This aspect alone can explain a large portion of the performance gap between WordNet and OpenCyc annotations. One possible solution is relaxing the evaluation measures and allowing for more than one possible annotation to be correct. Moreover, the annotation algorithms need to assume that there is not always one correct annotation; there can be more correct annotations or, as it is often the case with OpenCyc and DBpedia, none.

Secondly, although all three datasets share common features, these features are actually quite different due to dataset design. For example, human-readable descriptions in all three cases are written in very different genres and target different users. WordNet descriptions are written similar to dictionary entries, DBpedia descriptions are, by definition, written like encyclopedia entries and OpenCyc descriptions are meant as documentation to the ontology engineer using it to model some world phenomena. Similarly, relations in all three datasets have very distinct semantics, and the annotation methods developed or focused on so far either pay little attention to this or are largely overfitted to the few relations used in WordNet. Each of the datasets has its own vocabulary for determining the closeness of concepts. For example, OpenCyc uses relations such as *nearestIsa*, *nearestIsaOfType* or *conceptuallyRelated.* WordNet largely focuses on the closeness of concepts within one part of speech (e.g. nouns) having less relation types defined between different parts of speech. Both OpenCyc and DBpedia contain relations which mostly regard their infrastructure, (*wikiPageUsesTemplate* is the most common relation in infobox triplets) and, when naively used, are not a good indicator of concept closeness (e.g. PageRank approach from Section 3). To overcome this drawback, the annotation methods have to better take advantage of the rich relationship structure of LOD datasets and to allow for an easy addition of new relations and datasets.

With the future evolution of LOD, it would also be beneficial to introduce a model for defining lexical resources, which would be attached to the LOD resources. Currently, each resource can contain a label (*rdfs:label*) in one or more languages. It would be useful to assign more linguistic meta-data to these labels, such as part-of-speech, inflected forms (e.g. *go, goes, going, went,* and *gone*), etc. Since this is generally expensive to build, tools for doing this (semi-)automatically would also be of great benefit.

6. CONCLUSION AND FUTURE WORK

In this paper we investigated the applicability of two common approaches, taken from the word sense disambiguation community, for annotating text with LOD datasets. One of the approaches relies on the dataset relationship structure and is based on the Page Rank algorithm; the second one, called Context Similarity, takes advantage of the human-readable description of a resource as well as neighbourhood relationships defined for that resource. These approaches were chosen based on the common characteristics of three datasets: WordNet, DBpedia and OpenCyc. The experimental findings revealed the shortcomings of the current state-of-the-art word sense disambiguation methods when applied to different LOD datasets. In the discussion section we provided several possible explanations for these shortcomings together with alternatives and solutions.

As far as future work is concerned, we plan to use the lessons learned in the experiments we presented in order to further develop text annotations methods which can offer better performance on datasets, such as OpenCyc and DBpedia, and can be, with a reasonable and predictable amount of effort, transferred to other LOD datasets.

7. ACKNOWLEDGMENTS

The research leading to these results has received funding from the Slovenian Research Agency and the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°257790.

8. REFERENCES

[1] Agirre, E. and Soroa, A. 2009. *Personalizing PageRank for word sense disambiguation*. In Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL). Athens, Greece.

[2] Banerjee, S. and Pedersen, T. 2002. *An Adapted Lesk Algorithm for Word Sense Disambiguation using WordNet*. In Proceedings of the 3rd International Conference on Intelligent Text Processing and Computational Linguistics, pp. 136--145, Mexico City, Mexico.

[3] Bizer, C., Heath, T. and Berners-Lee, T. 2009. *Linked Data – The Story So Far*. International Journal on Semantic Web and Information Systems.

[4] Brin, S. and Page, M. 1998. *Anatomy of a large-scale hypertextual Web search engine*. In Proceedings of the 7th Conference on World Wide Web (WWW). Brisbane, Australia.

[5] Chan, Y. S., NG, H. T. and Zhong, Z. 2007. *NUS-PT: Exploiting parallel texts for word sense disambiguation in the English all-words tasks*. In Proceedings of the 4th International Workshop on Semantic Evaluations, Prague, Czech Republic.

[6] Curtis, J., Cabral, J. and Baxter, D. 2006. *On the application of the Cyc ontology to word sense disambiguation*. In Proceedings of the 19th International Florida Artificial Intelligence Research Society Conference.

[7] David, J., Euzenat, J. and Zamazal, O. S. 2010. *Ontology similarity in the alignment space*. In Proceedings of the 9th International Semantic Web Conference (ISWC), Shanghai, China.

[8] Edmonds, P. and Kilgariff, A. 1998. *Introduction to the special issue on evaluating word sense disambiguation systems*. In Journal of Natural Language Engineering, 8(4). Cambridge University.

[9] Euzenat, J. and Shvaiko, P. 2007. *Ontology matching*. Springer, Heidelberg, Germany.

[10] Medelyan, O. and Legg, C. 2008. *Integrating Cyc and Wikipedia: Folksonomy meets rigorously defined common-sense*. In Proceedings of Wikipedia and AI workshop at the AAAI 2008 Conference. Chicago, USA.

[11] Mihalcea, R., Tarau, P. and Figa, E. 2004. *PageRank on semantic networks, with application to word sense disambiguation*. In Proceedings of the 20th International Conference on Computational Linguistics (COLING). Geneva, Swizerland.

[12] Navigli, R and Velardi, P. 2005. *Structural semantic interconnections: A knowledge-based approach to word sense disambiguation*. IEEE Transactions on Pattern Analysis and Machine Intelligence Volume 27, Issue 7, pp. 1075—1088.

[13] Navigli, R., Litkowski, K. C. and Hargraves, O. 2007. *Semeval-2007 task 07: Coarse-grained English allwords task*. In Proceedings of the 4th International Workshop on Semantic Evaluations, Prague, Czech Republic.

[14] Navigli, R. 2009. *Word sense disambiguation: A survey*. ACM Computational Surveys, 41(2).

[15] Patwardhan, S., Banerjee, S. and Pedersen, T. 2007. *UMND1: Unsupervised Word Sense Disambiguation Using Contextual Semantic Relatedness*. In proceedings of the 4th International Workshop on Semantic Evaluations (SemEval). Prague, Czech Republic.

[16] Pedersen, T., Patwardhan and S. Michelizzi, J. 2004. *WordNet::Similarity - Measuring the Relatedness of Concepts*. In Proceedings of Fifth Annual Meeting of the North American Chapter of the Association for Computational Linguistics (ACL). Boston, USA

[17] Ponzetto, S.P. and Navigli, R. 2010. *Knowledge-rich Word Sense Disambiguation Rivaling Supervised Systems*. In Proceedings of the 48th Conference of the Association for Computational Linguistics. pp 1522--1531. Uppsala, Sweden.