FinMatcher at FinSim-2: Hypernym Detection in the Financial Services Domain using Knowledge Graphs

Jan Portisch
University of Mannheim, Data and
Web Science Group
Mannheim, Germany
jan@informatik.uni-mannheim.de

Michael Hladik
SAP SE Product Engineering
Financial Services
Walldorf, Germany
michael.hladik@sap.com

Heiko Paulheim
University of Mannheim, Data and
Web Science Group
Mannheim, Germany
heiko@informatik.uni-mannheim.de

ABSTRACT

This paper presents the *FinMatcher* system and its results for the FinSim 2021 shared task which is co-located with the Workshop on Financial Technology on the Web (FinWeb) in conjunction with The Web Conference. The FinSim-2 shared task consists of a set of concept labels from the financial services domain. The goal is to find the most relevant top-level concept from a given set of concepts. The FinMatcher system exploits three publicly available knowledge graphs, namely WordNet, Wikidata, and WebIsALOD. The graphs are used to generate explicit features as well as latent features which are fed into a neural classifier to predict the closest hypernym.

CCS CONCEPTS

• **Theory of computation** → *Semantics and reasoning*; • **Information systems** → **Web mining**; **Language models**.

KEYWORDS

financial services, knowledge graphs, wikidata, knowledge graph embeddings, RDF2vec, hypernymy detection

1 INTRODUCTION

A *hypernym* or *hyperonym* is a concept which is superordinate to another one. In computer science, it is often represented as an *IS-A* relationship. For example, *animal* is a hypernym of *cat* and *equity index* is a hypernym of *S&P 500 Index*. A *hyponym*, on the other hand, is a concept which is subordinate to another one. For example, *cat* is a hyponym of *animal* and *S&P 500 Index* is a hyponym of *equity index*. [15] Hypernymy detection can be broadly applied in real-world applications. The detection of hypernyms in the financial services domain is particularly interesting due to a domain specific vocabulary and a lack of publicly available domain-specific resources and concept representations.

The FinSim task models the hypernym detection task as a multi class classification problem: Given a concept label (i.e., the hyponym), the correct hypernym is to be found from a set of 10 mutually exclusive classes (i.e., hypernyms). A system participating in this task can return a sorted list of classes. The task is evaluated with two performance metrics: mean rank and accuracy.

The FinMatcher system uses two very broad publicly available knowledge graphs (Wikidata and WebIsALOD) as well as a small linguistic graph resource (WordNet). A knowledge graph contains real world entities from various domains and the relationships that hold between them in a graph format [16]. The system presented in this paper calculates multiple explicit features and uses RDF2vec embeddings obtained from WebIsALOD. The features are concatenated into a feature vector which is presented to a neural classifier which was trained with the provided FinSim training data.

In the following section, related work is introduced. Afterwards, the provided dataset is quickly described. In Section 4, the FinMatcher system is presented. The results of FinSim task are given in Section 5 together with an ablation study. The paper is concluded in Section 6 where future research directions are also presented.

2 RELATED WORK

2.1 Shared Tasks for Hypernym Detection

Hypernym discovery has been addressed before as challenge, for example at SemEval-2018 [1]. Unique to the FinSim task is the focus on the financial services industry. The evaluation campaign premiered in 2020 [4] and has been extended for the 2021 campaign, also referred to as FinSim-2 [11]: Two additional tags have been introduced and the training and evaluation datasets have been extended.

2.2 Knowledge Graphs

FinMatcher uses three external knowledge graphs as background knowledge for the task of hypernym detection.

WordNet [5] is a well known lexical resource. It is a database of English words grouped in sets which represent a particular meaning, called *synsets*; further semantic relations such as hypernymy also exist in the database. The resource is publicly available.[1]

Wikidata is a knowledge graph hosted by the Wikimedia Foundation which is publicly available[2] and maintained by an open community. The graph contains class-like entities, such as "stock market index", and also instance-like entities, such as "MSCI World".

---

[1] see https://wordnet.princeton.edu/download
[2] see https://www.wikidata.org/wiki/Wikidata:Main_Page

An example for a Wikidata statement would be *"MSCI World" instance of "stock market index"*[3]. The graph can be queried using SPARQL[4].

A frequent problem that occurs when working with external background knowledge in the financial services domain is the fact that less common entities – so called *long tail entities* – are not contained within a knowledge base. The *WebIsA* [22] database is an attempt to tackle this problem by providing a dataset which is not based on a single source of knowledge – like *DBpedia* [10] – but instead on the whole Web: The dataset consists of hypernymy relations extracted from the *Common Crawl*[5], a freely downloadable crawl of a significant portion of the Web. For the automated extraction, lexico-syntactic patterns similar to those presented by Hearst [6] were used. Like Wikidata, the graph contains class-like and instance-like concepts. A sample triple from the dataset is *"zerocoupon bond" skos:broader "bond"*[6]. The dataset is also available via a Linked Open Data (LOD) endpoint[7] under the name *WebIsALOD* [7] – hence, it can be queried like Wikidata using SPARQL.

2.3 Knowledge Graph Embeddings

In recent years, latent representations have gained traction not only in natural language processing but also in other data science communities. *RDF2vec* [21] is a knowledge graph embedding approach which allows to obtain a latent representation for the elements of a knowledge graph, i.e. a vector, for each node and each edge in a graph. It applies the *word2vec* [12, 13] model to RDF data: Random walks are performed for each node and are interpreted as sentences. After the walk generation, the sentences are used as input for the word2vec algorithm. As a result, one obtains a vector for each word, i.e., a concept in the RDF graph. Multiple flavors of RDF2vec have been developed in the past such as biased walks [3] or *RDF2vec Light* [20].[8] The calculation of knowledge graph embeddings on large graphs can require a significant amount of resources. Therefore, *KGvec2go*[9] [19] provides pre-trained RDF2vec knowledge graph embeddings through a Web API as well as via download. For the system presented in this paper, a pre-trained embedding of WebIsALOD has been downloaded from KGvec2go.

Both, RDF2vec and WebIsALOD, have been used for integration tasks in the financial services domain before [14, 17].

3 FINSIM DATASET DESCRIPTION

The FinSim dataset consists of 614 hyponym-hypernym pairs. There are 10 class labels (see Figure 1), i.e. hypernyms. The class labels classify concepts not according to their features but instead according to their prototypical kind. The distribution of class labels is not balanced. As shown in Figure 1, the distribution of labels follows a power-law with 286 entries for "equity index" and only 9 entries for "forward". This is a challenging setting for multiple reasons: (i) the training dataset is comparatively small, (ii) the hypernyms are

---
[3]see https://www.wikidata.org/wiki/Q1881843

[4]see https://query.wikidata.org/

[5]see http://commoncrawl.org/

[6]see http://webisa.webdatacommons.org/concept/zero-coupon_bond_

[7]see http://webisa.webdatacommons.org/

[8]For a good overview of the RDF2vec approach and its applications, refer to http://www.rdf2vec.org/
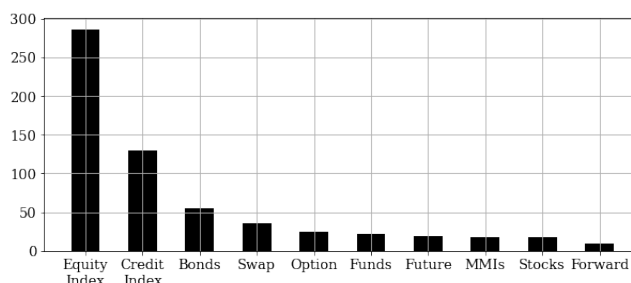
[9]see http://www.kgvec2go.org

Figure 1: Distribution of class labels in the FinSim training dataset

semantically very related, (iii) industry abbreviations are used, and (iv) there are textual overlaps. The FinSim-2 test dataset consists of 212 entries, the distribution of class labels is not known.

Compared to other evaluation campaigns where participants have to submit their implementations, such as the Ontology Alignment Evaluation Initiative (OAEI), participants of the FinSim task run their system on their own premises and submit the predictions made by their system.

4 SYSTEM DESCRIPTION

The FinMatcher system combines explicit and latent features. In total, there are five groups of features which will be presented in the following. The overall architecture is shown in Figure 2.

4.1 Features

Word Overlap. The overlap between hyponym and class label is a strong signal for a match. An example would be "Supranational Bond" which is a "Bond". As such constellations are relatively frequent in the provided dataset, the first feature vector encodes whether the label contains the class label. For this feature minimal text pre-processing is applied including lower-casing and removal of the plural suffix "s". As this step is performed for each class label, a vector of length 10 is obtained. The overlap feature vector is displayed in green in Figure 2.

Wikidata Hypernym Lookup. Wikidata is a large general-purpose knowledge graph which is not tailored to the financial domain. Nonetheless, the data source contains many financial concepts and relations between them. For example, the concept "UCITS" can be linked to "Undertakings for Collective Investment in Transferable Securities" via the *also known as* label; due to the annotated relation *subclass of*, it is easily recognizable that "UCITS" is an "investment fund".[10] This notion is exploited in this set of features: A comprehensive linking mechanism from the MELT framework[11] [8, 9] is used to link classes (the hypernyms) as well as labels (the hyponyms) to Wikidata concepts and then relations $P31$ (instance of) and $P279$ (sublcass of) are followed up to two hops to evaluate whether the class label appears. Distant matches receive a lower signal strength which is calculated through the inverse hop-distance:

---
[10]see https://www.wikidata.org/wiki/Q25323628

[11]The Matching EvaLuation Toolkit (MELT) is a framework for ontology and instance matching (development, evaluation, visualization [18]). However, components can also be exploited for other tasks. For a better overview, see https://github.com/dwslab/melt/

A direct hypernym annotation (as in the UCITS example stated earlier) receives the value $\frac{1}{1} = 1$ whereas a two-hop match would receive a value of $\frac{1}{2} = 0.5$. As this step is performed for each class label, a vector of length 10 is obtained. The Wikidata lookup feature vector is displayed in blue in Figure 2.

WordNet Hypernym Lookup. The same exploitation approach chosen for Wikidata is applied on the WordNet graph: Hypernyms and hyponyms are linked into WordNet and then the inverse hop-distance is used as feature value. This is done for each class label that could be linked. The WordNet lookup feature vector is displayed in yellow in Figure 2.

WebIsALOD Hypernym Lookup. In a similar fashion to the Wiki-data hypernym lookup, class labels as well as hyponym labels are linked to the WebIsALOD graph using a linker from the MELT framework. In this graph, there exists only one significant relation: *skos:broader.* For each hyponym, the broader concepts are obtained and it is checked whether the hypernym appears. Due to a high level of noise, the number of upwards hops is limited to 1. As this step is performed for each class label, a vector of length 10 is obtained. The WebIsALOD lookup feature vector is displayed in purple in Figure 2.

WebIsALOD RDF2vec Similarity. For the embedding feature, each class label as well as each hyponym label is linked again into the WebIsALOD knowledge graph. Each concept in WebIsALOD has an associated embedding vector $v \in I\!\!R^{200}$. For comparisons, the cosine similarity between the hyponym and the class label is calculated.

If the whole concept cannot be linked, multiple sub-concepts are detected and linked. Within this linking process, longer sub-concepts are favored. For example, the string "CDX Emerging Markets" cannot be directly linked – however, the longest substring that can be linked here is "Emerging Markets"; in addition, "CDX" can also be linked. Comparisons in such cases are performed as follows:

$$\frac{\sum_{i=0}^{I} \max_{j=0}^{J}(sim(v_i, v_j))}{|I|} \tag{1}$$

where I represents the set of links of the hyponym, J represents the set of links of the hypernym, v_i and v_j correspond to the vectors of the links and *sim* refers to a similarity function. In this case, the cosine is used as similarity function. As this step is performed for each class label, a vector of length 10 is obtained. The WebIsALOD lookup feature vector is displayed in salmon in Figure 2.

Feature Composition. Each of the features i returns a signal vector $s_i \in I\!\!R^{10}$. All vectors are concatenated to form the final signal vector $S = \|_{i=1}^{5} s_i$, which is used as input for the classifier.

4.2 Classifier

Due to the small total number of training examples, a very simple artificial neural network architecture has been chosen. It is configured with one fully connected layer of size 10 and mean squared error as loss. The network was trained with 100 epochs and a batch size of 25 on a consumer PC. The vector that is to be predicted is of size 10 and represents the one-hot-encoded class label. The neural

Table 1: Ablation Study

Left-out Vector Group	Mean Accuracy	Mean Rank (HITS@10)
Submission	86.69	1.432
No SMOTE	85.55	1.371
Word Overlap	60.51	2.007
Wikidata Hypernyms	85.50	1.490
WordNet Hypernyms	86.66	1.440
ALOD Hypernyms	85.88	1.361
ALOD RDF2vec	85.56	1.481

network classifier performed best among the classifiers evaluated: Naïve Bayes, J48 decision trees, random forests, and a regression.

As the distribution of class labels is skewed (see Figure 1), we applied the synthetic minority oversampling technique (SMOTE) [2] to upsample underrepresented class labels. We experimentally chose 33% of the majority class total as the upsampling barrier; this means that if the majority class in the training split totals to 229 records, upsampling for class labels with less than $\frac{1}{3} * 229 = 76$ records will be performed so that there are 76 records for the underrepresented class label.

5 RESULTS

5.1 Results on the Training Data and Ablation Study

We evaluated our matching system by performing a stratified five-fold cross validation on the training data. We trained each ANN configuration 10 times and report the average results for accuracy and mean rank. We further performed an ablation study by training and evaluating the performance when leaving out each of the five feature groups. The results can be found in Table 1.

It is visible that the most important feature group in terms of accuracy is word overlap. This is not surprising given the high number of labels that contain the hypernym within their name (for example "green bonds" → "bonds") and shows that it is sensible for the task at hand to combine explicit and latent features. The observation that the inclusion of the target label in the term is a significant signal has also been made in the last FinSim campaign [4]. The negligible role of WordNet in terms of accuracy is also comprehensible since this particular external background knowledge dataset contains merely general-purpose class knowledge (such as "call option") but no knowledge about instances (such as "MSCI EMU Index"). For the FinSim dataset, very large knowledge graphs that contain class as well as instance knowledge are more beneficial due to their higher concept coverage. However, the information in the knowledge graphs used also contain some redundancy, as can be observed in Table 1: leaving out a single knowledge graph does not significantly change the results.

To further analyze the contribution of the different signals, we plotted the weights of the input features. As the weight of each input neuron s_i relates to label i, we can directly observe which features the trained model considers relevant to identify which label.

Table 2 shows the summed absolute weight per feature group. This allows to analyze the overall contribution of the individual
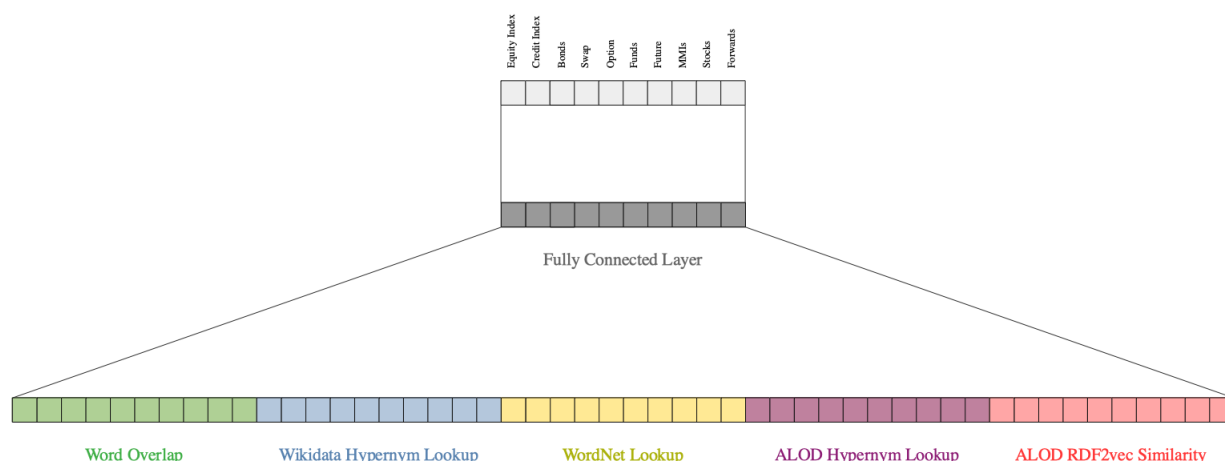
Figure 2: Architecture of FinMatcher

Table 2: Absolute Weights per Feature Group

Vector Group	Weight
Word Overlap	13.64
Wikidata Hypernyms	13.01
WordNet Hypernyms	13.35
ALOD Hypernyms	9.12
ALOD RDF2vec	14.01

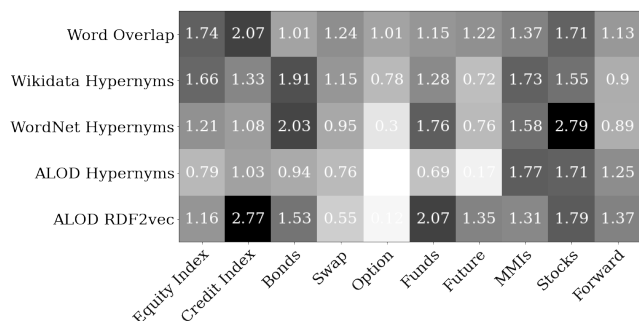|	Equity Index	Credit Index	Bonds	Swap	Option	Funds	Future	MMIs	Stocks	Forward
Word Overlap	1.74	2.07	1.01	1.24	1.01	1.15	1.22	1.37	1.71	1.13
Wikidata Hypernyms	1.66	1.33	1.91	1.15	0.78	1.28	0.72	1.73	1.55	0.9
WordNet Hypernyms	1.21	1.08	2.03	0.95	0.3	1.76	0.76	1.58	2.79	0.89
ALOD Hypernyms	0.79	1.03	0.94	0.76		0.69	0.17	1.77	1.71	1.25
ALOD RDF2vec	1.16	2.77	1.53	0.55	0.12	2.07	1.35	1.31	1.79	1.37

Figure 3: Heatmap of the Absolute Weights per Feature Group and Class Label

feature group. Here, it is visible that the latent RDF2vec feature group receives the highest weight – higher than the word overlap group.

While the word overlap feature is important for the majority labels (equity index, credit index), it is not equally important for all labels and does not have the overall highest weight: Figure 3 shows the summed absolute weight per feature group and class label. The class labels are sorted in descending order by frequency. Here, it is visible that the word overlap has the highest contribution for the equity index as well as a high contribution for the credit index but low weights for the remaining minority classes.

5.2 Results on the Reference Data

FinMatcher participated only with one configuration and achieved an accuracy of 81.1% and a mean rank of 1.415 on the reference data below the expected scores from the training data shown in Table 1.

6 CONCLUSION

In this paper, we presented *FinMatcher*, a hypernym detection system for the financial services domain which exploits multiple knowledge graphs by combining explicit and latent features. We could show that the task can be addressed by including external knowledge in the form of knowledge graphs and that the combination of multiple graphs is overall beneficial.

In the future, we strive to improve the results through the inclusion of more advanced embedding techniques as well as the exploration of additional external datasets.

ACKNOWLEDGMENTS

REFERENCES

[1] José Camacho-Collados, Claudio Delli Bovi, Luis Espinosa Anke, Sergio Oramas, Tommaso Pasini, Enrico Santus, Vered Shwartz, Roberto Navigli, and Horacio Saggion. 2018. SemEval-2018 Task 9: Hypernym Discovery. In *Proceedings of The 12th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2018, New Orleans, Louisiana, USA, June 5-6, 2018*, Marianna Apidianaki, Saif M. Mohammad, Jonathan May, Ekaterina Shutova, Steven Bethard, and Marine Carpuat (Eds.). Association for Computational Linguistics, 712–724. https://doi.org/10.18653/v1/s18-1115

[2] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* 16 (2002), 321–357. https://doi.org/10.1613/jair.953

[3] Michael Cochez, Petar Ristoski, Simone Paolo Ponzetto, and Heiko Paulheim. 2017. Biased graph walks for RDF graph embeddings. In *Proceedings of the 7th International Conference on Web Intelligence, Mining and Semantics, WIMS 2017, Amantea, Italy, June 19-22, 2017*, Rajendra Akerkar, Alfredo Cuzzocrea, Jannong Cao, and Mohand-Said Hacid (Eds.). ACM, 21:1–21:12. https://doi.org/10.1145/3102254.3102279

[4] Ismail El Maarouf, Youness Mansar, Virginie Mouilleron, and Dialekti Valsamou-Stanislawski. 2021. The finsim 2020 shared task: Learning semantic representations for the financial domain. In *Proceedings of the Second Workshop on Financial Technology and Natural Language Processing*. 81–86.

[5] Christiane Fellbaum (Ed.). 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, Massachusetts.

[6] Marti A. Hearst. 1992. Automatic Acquisition of Hyponyms from Large Text Corpora. In *14th International Conference on Computational Linguistics, COLING 1992, Nantes, France, August 23-28, 1992*. 539–545. https://www.aclweb.org/anthology/C92-2082/

[7] Sven Hertling and Heiko Paulheim. 2017. WebIsALOD: Providing Hypernymy Relations Extracted from the Web as Linked Open Data. In *The Semantic Web - ISWC 2017 - 16th International Semantic Web Conference, Vienna, Austria, October 21-25, 2017, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 10588)*, Claudia d'Amato, Miriam Fernández, Valentina A. M. Tamma, Freddy Lécué, Philippe Cudré-Mauroux, Juan F. Sequeda, Christoph Lange, and Jeff Heflin (Eds.). Springer, 111–119. https://doi.org/10.1007/978-3-319-68204-4_11

[8] Sven Hertling, Jan Portisch, and Heiko Paulheim. 2019. MELT - Matching EvaLuation Toolkit. In *Semantic Systems. The Power of AI and Knowledge Graphs - 15th International Conference, SEMANTiCS 2019, Karlsruhe, Germany, September 9-12, 2019, Proceedings (Lecture Notes in Computer Science, Vol. 11702)*, Maribel Acosta, Philippe Cudré-Mauroux, Maria Maleshkova, Tassilo Pellegrini, Harald Sack, and York Sure-Vetter (Eds.). Springer, 231–245. https://doi.org/10.1007/978-3-030-33220-4_17

[9] Sven Hertling, Jan Portisch, and Heiko Paulheim. 2020. Supervised ontology and instance matching with MELT. In *Proceedings of the 15th International Workshop on Ontology Matching co-located with the 19th International Semantic Web Conference (ISWC 2020), Virtual conference (originally planned to be in Athens, Greece), November 2, 2020 (CEUR Workshop Proceedings, Vol. 2788)*, Pavel Shvaiko, Jérôme Euzenat, Ernesto Jiménez-Ruiz, Oktie Hassanzadeh, and Cássia Trojahn (Eds.). CEUR-WS.org, 60–71. http://ceur-ws.org/Vol-2788/om2020_LTpaper6.pdf

[10] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick van Kleef, Sören Auer, and Christian Bizer. 2015. DBpedia - A large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web* 6, 2 (2015), 167–195. https://doi.org/10.3233/SW-140134

[11] Youness Mansar, Juyeon Kang, and Ismaïl El Maarouf. 2021. FinSim-2 2021 Shared Task: Learning Semantic Similarities for the Financial Domain. In *Proceedings of The Web Conference 2021, Virtual Edition*.

[12] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. In *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.). http://arxiv.org/abs/1301.3781

[13] Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. 2013. Distributed Representations of Words and Phrases and their Compositionality. In *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States*, Christopher J. C. Burges, Léon Bottou, Zoubin Ghahramani, and Kilian Q. Weinberger (Eds.). 3111–3119. http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality

[14] Michael Monych, Jan Portisch, Michael Hladik, and Heiko Paulheim. 2020. DESKMatcher. In *Proceedings of the 15th International Workshop on Ontology Matching co-located with the 19th International Semantic Web Conference (ISWC 2020), Virtual conference (originally planned to be in Athens, Greece), November 2, 2020 (CEUR Workshop Proceedings, Vol. 2788)*, Pavel Shvaiko, Jérôme Euzenat, Ernesto Jiménez-Ruiz, Oktie Hassanzadeh, and Cássia Trojahn (Eds.). CEUR-WS.org, 181–186. http://ceur-ws.org/Vol-2788/oaei20_paper7.pdf

[15] Lynne Murphy. 2003. *Semantic relations and the lexicon : antonymy, synonymy, and other paradigms*. Cambridge University Press, Cambridge, UK New York. 2016–236 pages.

[16] Heiko Paulheim. 2017. Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web* 8, 3 (2017), 489–508. https://doi.org/10.3233/SW-160218

[17] Jan Portisch. 2018. Automatic schema matching utilizing hypernymy relations extracted from the web.

[18] Jan Portisch, Sven Hertling, and Heiko Paulheim. 2020. Visual Analysis of Ontology Matching Results with the MELT Dashboard. In *The Semantic Web: ESWC 2020 Satellite Events - ESWC 2020 Satellite Events, Heraklion, Crete, Greece, May 31 - June 4, 2020, Revised Selected Papers (Lecture Notes in Computer Science, Vol. 12124)*, Andreas Harth, Valentina Presutti, Raphaël Troncy, Maribel Acosta, Axel Polleres, Javier D. Fernández, Josiane Xavier Parreira, Olaf Hartig, Katja Hose, and Michael Cochez (Eds.). Springer, 186–190. https://doi.org/10.1007/978-3-030-62327-2_32

[19] Jan Portisch, Michael Hladik, and Heiko Paulheim. 2020. KGvec2go - Knowledge Graph Embeddings as a Service. In *Proceedings of The 12th Language Resources and Evaluation Conference, LREC 2020, Marseille, France, May 11-16, 2020*, Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asunción Moreno, Jan Odijk, and Stelios Piperidis (Eds.). European Language Resources Association, 5641–5647. https://www.aclweb.org/anthology/2020.lrec-1.692/

[20] Jan Portisch, Michael Hladik, and Heiko Paulheim. 2020. RDF2Vec Light - A Lightweight Approachfor Knowledge Graph Embeddings. In *Proceedings of the ISWC 2020 Demos and Industry Tracks: From Novel Ideas to Industrial Practice co-located with 19th International Semantic Web Conference (ISWC 2020), Globally online, November 1-6, 2020 (UTC) (CEUR Workshop Proceedings, Vol. 2721)*, Kerry L. Taylor, Rafael S. Gonçalves, Freddy Lécué, and Jun Yan (Eds.). CEUR-WS.org, 79–84. http://ceur-ws.org/Vol-2721/paper520.pdf

[21] Petar Ristoski, Jessica Rosati, Tommaso Di Noia, Renato De Leone, and Heiko Paulheim. 2019. RDF2Vec: RDF graph embeddings and their applications. *Semantic Web* 10, 4 (2019), 721–752. https://doi.org/10.3233/SW-180317

[22] Julian Seitner, Christian Bizer, Kai Eckert, Stefano Faralli, Robert Meusel, Heiko Paulheim, and Simone Paolo Ponzetto. 2016. A Large DataBase of Hypernymy Relations Extracted from the Web. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC 2016, Portorož, Slovenia, May 23-28, 2016*, Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asunción Moreno, Jan Odijk, and Stelios Piperidis (Eds.). European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2016/summaries/204.html