Semantic Web 10 (2019) 815–822
DOI 10.3233/SW-190364
IOS Press

Special issue on Semantic Deep Learning

Editorial

Dagmar Gromann [a,*], Luis Espinosa Anke [b] and Thierry Declerck [c]
[a] *Centre for Translation Studies, University of Vienna, A-1190 Vienna, Austria*
E-mail: dagmar.gromann@univie.ac.at
[b] *School of Computer Science and Informatics, Cardiff University, Queen's Buildings, Roath, Cardiff, CF24 3AA, United Kingdom*
E-mail: espinosa-ankel@cardiff.ac.uk
[c] *Multilinguality and Language Technology Lab, German Research Centre for Artificial Intelligence (DFKI GmbH), D-66123 Saarbrücken, Germany*
E-mail: declerck@dfki.de

Editors: Pascal Hitzler, Kansas State University, Manhattan, KS, USA; Krzysztof Janowicz, University of California, Santa Barbara, CA, USA

Abstract. Numerous success use cases involving deep learning have recently started to be propagated to the Semantic Web. Approaches range from utilizing structured knowledge in the training process of neural networks to enriching such architectures with ontological reasoning mechanisms. Bridging the neural-symbolic gap by joining deep learning and Semantic Web not only holds the potential of improving performance but also of opening up new avenues of research. This editorial introduces the Semantic Web Journal special issue on Semantic Deep Learning, which brings together Semantic Web and deep learning research. After a general introduction to the topic and a brief overview of recent contributions, we continue to introduce the submissions published in this special issue.

Keywords: Deep learning, Semantic Web, knowledge injection

1. Introduction

Semantic Web technologies and deep learning share the goal of creating intelligent artifacts that emulate human capacities such as reasoning, validating, and predicting. Both fields have been impacting data and knowledge analysis considerably as well as their associated abstract representations. The term *deep learning* is used to refer to deep neural network algorithms that learn data representations by means of transformations with multiple processing layers. Today, such architectures are well studied in the field of Natural Language Processing (NLP), where they have been successfully applied to numerous research challenges. These include low-level tasks, such as part-of-speech [51] and morphological tagging [17], as well as higher-level linguistic problems, such as language modeling [19,43,48,53,61], named entity recognition [38], machine translation [7,37], or direct speech to speech translation [34]. *Semantic Web* technologies and knowledge representation boost the re-use and sharing of knowledge in a structured and machine-readable fashion. Semantic resources such as WikiData [66], Yago [60], BabelNet [45] or DBpedia [5], as well as knowledge base construction and completion methods [9,10], have been successfully applied to improve systems addressing semantically intensive tasks (e.g. Question Answering as in [32]).

There are notable examples showcasing the influence of neural approaches to knowledge acquisition and representation learning on the broad area of Semantic Web technologies. These include, among oth-

*Corresponding author. E-mail: dagmar.gromann@univie.ac.at.

ers, ontology learning [40,49,65], learning structured query languages from natural language [69], ontology alignment [20,28,35,52], ontology annotation [15,58], joined relational and multi-modal knowledge representations [62], and relation prediction [1,59]. Ontologies, on the other hand, have been repeatedly utilized as background knowledge for machine learning tasks. As an example, there is a myriad of hybrid approaches for learning linguistic representations by jointly incorporating corpus-based evidence and semantic resources [13,25,27,33,50]. This interplay between structured knowledge and corpus-based approaches has given way to knowledge graph embeddings, which in turn have proven useful for tasks such as hypernym discovery [21], collocation discovery and classification [22], word sense disambiguation [12,54], joined relational and multi-modal knowledge representations [62] and many others.

In this context, this special issue aims to provide a playground for exploring the interaction between neural NLP and representation learning, on the one hand, and symbolic representation of knowledge and data-driven approaches to pattern recognition, on the other. Specifically, we invited submissions illustrating how Semantic Web resources and technologies benefit from interacting with neural networks. At the same time, we also encouraged submissions showing how knowledge representation would assist in neural NLP tasks, and how knowledge representation systems can build on top of deep learning. The timeliness of this special issue becomes apparent, also, in the potential of symbolic representations of knowledge in the form of ontologies, knowledge graphs, and rules to contribute to the long standing goal of explainable and interpretable Artificial Intelligence [39], for example, for "keep-a-human-in-the-loop" approaches [31] or directly for reasoning about neural network decisions [14].

This special issue builds on and complements a series of workshops dedicated to *Semantic Deep Learning* (SemDeep), co-organized by the editors of this issue. The first workshop, co-chaired by Dagmar Gromann, Thierry Declerck and Georg Heigold, took place as a satellite event to the 14th Extended Semantic Web Conference (ESWC 2017). Based on the success of this first edition of SemDeep, a new edition was submitted to the 12th International Conference on Computational Semantics (IWCS 2017), where SemDeep-2 could reach the computational semantics community. SemDeep-3, now co-chaired by Dagmar Gromann, Thierry Declerck and Luis Espinosa Anke, was co-located with the 27th International Conference on Computational Linguistics (COLING 2018). The fourth edition of SemDeep came back to the Semantic Web community and was co-located with the 17th International Semantic Web Conference (ISWC 2018). SemDeep-5 is a workshop of the 28th International Joint Conference on Artificial Intelligence (IJCAI 2019). For this last workshop, the Organizing Committee was glad to welcome José Camacho Collados and Mohammad Taher Pilehvar, as this new edition of the successful SemDeep workshop series was augmented by a challenge on evaluating contextualized word representations called Word-in-Context (WiC).

2. Recent Semantic Deep Learning approaches

Neural-symbolic approaches (see e.g. [6,8,30] for an overview) represent a relatively young field of research, having only attracted considerable attention within the last few years. The SemDeep series in general, and this special issue in particular, have offered a forum where such methods, from a proof-of-concept stage to a more advanced and robust stage of development, could be presented and discussed.

Specifically, SemDeep has seen contributions on the explicit modeling of lexical and semantic relations stemming from joint neural-symbolic methods [23,44,55]. Additionally, well-defined NLP tasks have also been the focus of several SemDeep papers over the years, covering event detection [11], part-of-speech tagging [67], co-reference resolution [63], sentiment analysis [47], named entity recognition [41] or question answering [32]. Interestingly, another area that has been prominently covered in SemDeep is (formal) knowledge representation, such as the tasks of link prediction in generic knowledge bases as well as domain-specific use cases [3,4,71]. Fewer works focused on more technical aspects of a knowledge-enhanced deep learning pipeline, for example, exploring disjointness in loss functions for classification tasks [57], end-to-end memory networks [36], image-based neural user profiling [68] or Siamese Long Short Term Memory (LSTM) networks [29].

The topic that has attracted most interest in the SemDeep workshop series has been representation learning, and a plethora of submissions were accepted for publication where vector representation of linguistic items, as well as meta-embeddings, were discussed. The concrete topics covered included word and document embeddings [56,70], knowledge graph embeddings [26], joint knowledge graph and text embeddings [16,42], multi-modal approaches [64,68], leveraging external information such as lexical resources [46], embeddings for low resource languages like Igbo [24], and learning structured knowledge [2,18].

3. Overview of this special issue

The paper *Deep learning for noise-tolerant RDFS reasoning* by Bassem Makni and James Hendler presents a noise-tolerant RDFS reasoning approach building on neural machine translation. To this end, they present an embedding approach where RDF graphs are layered and encoded in 3D adjacency matrices where each layer layout forms a graph word. Both input graph and its entailments are represented as a sequence of graph words and RDFS inferences can then be formulated as a machine translation of these graph word sequences. As such the approach seeks to bridge the neural-symbolic graph, adapting the idea of knowledge graph embeddings to RDF graphs, the differences of which are analyzed in detail. Evaluations demonstrate the ability of the approach to learn RDFS rules from a synthetic dataset as well as DBpedia subset and a noise-tolerance not observed in rule-based reasoners.

In *Semantic Referee: A Neural-Symbolic Framework for Enhancing Geospatial Semantic Segmentation* Marjan Alirezaie, Martin Längkvist, Michael Sioutis, and Amy Loutfi propose a neural-symbolic framework in which an ontological reasoner characterizes output errors of a deep learning framework, providing corresponding feedback to improve its performance. In contrast to approaches that seek to integrate neural and symbolic aspects, this approach focuses on the interaction between these two. The reasoner functions as supervisor, a so-called semantic referee, in the training process of a variation of a Convolutional Autoencoder utilized to perform semantic segmentation of satellite images. In order to feed concepts inferred by the reasoner for misclassified regions back to the neural network, such concepts are encoded as image channels and concatenated with the original RGB channels. Additional information from the reasoner in the proposed approach relates to shadow estimation, elevation estimation, and inconsistencies with respect to the ontology constraints. On two real-world datasets and the OntoCity ontology the approach could demonstrate its capacity to reduce classification errors.

The paper *Vecsigrafo: Corpus-based Word-Concept Embeddings – Bridging the Statistic-Symbolic Representational Gap in Natural Language Processing* by José Manuel Gómez-Pérez and Ronald Denaux proposes to jointly learn word and concept embeddings from large semantically annotated corpora. Words are tokenized and disambiguated, where three disambiguation techniques are tested, and associated with concepts from an existing knowledge graph. In an extensive evaluation comparing to word and knowledge graph embeddings on the tasks of word similarity, word prediction, and relation prediction, the merits of the proposed approach are demonstrated. It could show that the joint learning of word and concept embeddings improves the quality over individual word and knowledge graph embeddings and different aspects of such joint vector spaces are discussed in detail in the paper. Furthermore, an extensive ablation study provides interesting insights into variants of Vecsigrafo, such as lemmatization having positive effects on the resulting joint embeddings, filtering improving the coverage of concepts, and disambiguation strategies only marginally impacting lexical embeddings.

In *Studying the Impact of the Full-Network Embedding on Multimodal Pipelines* by Armand Vilalta, Dario Garcia-Gasulla, Ferran Parés, Eduard Ayguadé, Jesus Labarta, E Ulises Moya-Sánchez, and Ulises Cortés, a Full-Network Embedding architecture is evaluated on the task of image annotation and retrieval. This architecture takes an image and its corresponding caption as input and produces a single vector representation that combines the output of a Convolutional Neural Network (CNN) applied to the image with the output of a Gated Recurrent Unit (GRU) applied to the caption. The authors propose, in this empirical study, to determine the fitness of such model as opposed to the one-layer image embeddings typically used in the literature. They report experimental results in three publicly available datasets, namely Flickr8K, Flickr30K and MSCOCO, and discuss different settings, involving hyperparemeter configurations, quality of training data and type of source CNN models.

The paper *Hate Speech Detection: A Solved Problem? The Challenging Case of Long Tail on Twitter* by Ziqi Zhang and Lei Luo first describes a data analysis to quantify and qualify the linguistic characteristics of hate speech in the social media. As a result, the authors show that it is much harder to detect hateful content than non-hate speech in social media, as hateful speech (in Twitter) lacks of unique, discriminative linguistic features. In a second part, the authors propose neural network structures for identifying specific types of hate speech, with a focus on two neural models. One is simulating a skip-gram like feature extraction based on modified a CNN, while the other extracts orderly information between features using a GRU. An evaluation on an English Twitter datasets shows that the described approach can outperform state-of-the-art methods by up to five percentage points.

In *A Convolutional Neural Network-based Model for Knowledge Base Completion and Its Application to Search Personalization* by Dai Quoc Nguyen, Dat Quoc Nguyen, Tu Dinh Nguyen, and Dinh Phung, an embedding model of entities and relationships for knowledge base completion is introduced. This model, named ConvKB, generalizes transitional characteristics in transition-based embedding models. ConvKB has been evaluated on two benchmark datasets, WN18RR and FB15k-237, and better link prediction performance than state-of-the-art embedding models can be reported. ConvKB has been additionally evaluated for triple classification on two benchmark datasets, WN11 and FB13, in order to check if a given triple is valid or not. Here ConvKB also gets better results than state-of-the-art models. Finally, the authors describe the adaptation of ConvKB for search personalization. This application of ConvKB has been positively verified on the query logs of a commercial web search engine.

4. Conclusion and future directions

Contributions to this special issue have focused on utilizing deep learning in connection with reasoning – either making the network itself a reasoner or enabling interaction between deep learning and a reasoner – multi-modal embeddings, feature extraction from natural language, and knowledge base completion. While this enumeration already hints at the large variety of central approaches to *Semantic Deep Learning*, further advances are needed, especially in the area of deep reasoning and inferences.

Combining Semantic Web technologies and deep learning holds the potential to crucially contribute to the recent hype of Explainable Artificial Intelligence (XAI). This might, for instance, take the form of injecting knowledge into training procedures to estimate changes of behaviors depending on utilized knowledge. Another important future direction is further systematic investigations into multi-modal approaches connecting linguistic, visual, and sensory inputs.

Acknowledgements

We would like to thank all the authors of accepted and rejected articles for their efforts and the editors-in-chief of Semantic Web Journal, Pascal Hitzler and Krzysztof Janowicz, for their continued support and help. A special thanks goes to all the people who have made this special issue possible. This is devoted in particular to our reviewers (in alphabetical order): Kemo Adrian, Luu Ahn Tuan, Claudia d'Amato, Miguel Ballesteros, Peter Bloem, Jose Camacho-Collados, Michael Cochez, Stamatia Dasiopoulou, Derek Doran, Cristina España i Bonet, Maarten Grachten, Dario Garcia-Gasulla, Jorge Gracia, Jindrich Helcl, Rezaul Karim, Mayank Kejriwal, Freddy Lecue, Alessandro Lenci, Antonio Lieto, Alessandra Mileo, Sergio Oramas, Petya Osenova, Simone Paolo Ponzetto, Heiko Paulheim, Martin Riedl, Francesco Ronzano, Enrico Santus, Francois Scharffe, Vered Shwartz, Kiril Simov, Michael Spranger, Armand Vilalta, Piek Vossen, and Arkaitz Zubiaga.

The edition of this special issue on Semantic Deep Learning has been supported by the German national project DeepLee, which is partially funded by the German Federal Ministry of Education and Research under the funding code 01IW17001. Responsibility for the content of this workshop is with the editor(s).

References

[1] F. Abedini, M.R. Keyvanpour and M.B. Menhaj, Neural tensor network training using meta-heuristic algorithms for RDF knowledge bases completion, *Applied Artificial Intelligence* **33**(7) (2019), 656–667. doi:10.1080/08839514.2019.1602317.

[2] K. Adrien, From learning through signal processing to argumentation on ontological representations, in: *First Workshop on Semantic Deep Learning (SemDeep-1) at the European Semantic Web Conference*, 2017, http://semdeep.iiia.csic.es/program.html.

[3] A. Agibetov, G. Dorffner and M. Samwald, Using hyperbolic large-margin classifiers for biological link prediction, in: *Proceedings of the 5th Workshop on Semantic Deep Learning (SemDeep-5)*, L. Espinosa-Anke, T. Declerck, D. Gromann, J. Camacho-Collados and M.T. Pilehvar, eds, Association for Computational Linguistics, Macau, China, 2019, pp. 26–30, https://www.aclweb.org/anthology/W19-5805.

[4] A. Agibetov and M. Samwald, Global and local evaluation of link prediction tasks with neural embeddings, in: *Joint Proceedings of ISWC 2018 Workshops SemDeep-4 and NLIWOD-4*, K.-S. Choi, L.E. Anke, T. Declerck, D. Gromann, J.-D. Kim, A.-C. Ngonga-Ngomo, M. Saleem and R. Usbeck, eds, CEUR Workshop Proceedings, Aachen, Germany, 2018, pp. 89–102.

[5] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak and Z. Ives, DBpedia: A nucleus for a web of open data, in: *Proceedings of the 6th International the Semantic Web and 2Nd Asian Conference on Asian Semantic Web Conference*, K. Aberer, P. Cudré-Mauroux, K.-S. Choi, N. Noy, D. Allemang, K.-I. Lee, L. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi and G. Schreiber, eds, ISWC'07/ASWC'07, Springer-Verlag, Busan, Korea, 2007, pp. 722–735, http://dl.acm.org/citation.cfm?id=1785162.1785216. ISBN 3-540-76297-3.

[6] S. Bader and P. Hitzler, Dimensions of neural-symbolic integration – A structured survey, in: *We Will Show Them: Essays in Honour of Dov Gabbay*, B. Artemov, A.S. d'Avila Garcez, L.C. Lamb and J. Woods, eds, Vol. 1, International Federation for Computational Logic, College Publications, London, UK, pp. 167–194.

[7] D. Bahdanau, K. Cho and Y. Bengio, Neural machine translation by jointly learning to align and translate, in: *3rd International Conference on Learning Representations, ICLR 2015, Conference Track Proceedings*, San Diego, CA, USA, May 7–9, 2015, Y. Bengio and Y. LeCun, eds, 2015, http://arxiv.org/abs/1409.0473.

[8] T.R. Besold, A.S. d'Avila Garcez, S. Bader, H. Bowman, P.M. Domingos, P. Hitzler, K. Kühnberger, L.C. Lamb, D. Lowd, P.M.V. Lima, L. de Penning, G. Pinkas, H. Poon and G. Zaverucha, Neural-symbolic learning and reasoning: A survey and interpretation, *CoRR* 2017, http://arxiv.org/abs/1711.03902.

[9] A. Bordes, S. Chopra and J. Weston, Question answering with subgraph embeddings, in: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, A. Moschitti, B. Pang and W. Daelemans, eds, Association for Computational Linguistics, Doha, Qatar, 2014, pp. 615–620, https://www.aclweb.org/anthology/D14-1067. doi:10.3115/v1/D14-1067.

[10] A. Bordes, J. Weston, R. Collobert and Y. Bengio, Learning structured embeddings of knowledge bases, in: *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, D. Fox, W. Burgard and D. Roth, eds, AAAI'11, AAAI Press, Palo Alto, CA, USA, 2011, pp. 301–306, http://dl.acm.org/citation.cfm?id=2900423.2900470.

[11] G. Burel, H. Saif, M. Fernandez and H. Alani, On semantics and deep learning for event detection in crisis situations, in: *First Workshop on Semantic Deep Learning (SemDeep-1) at the European Semantic Web Conference*, 2017, http://semdeep.iiia.csic.es/program.html.

[12] J. Camacho-Collados, C.D. Bovi, A. Raganato and R. Navigli, A large-scale multilingual disambiguation of glosses, in: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk and S. Piperidis, eds, European Language Resources Association (ELRA), Portorož, Slovenia, 2016, pp. 1701–1708, https://www.aclweb.org/anthology/L16-1269.

[13] J. Camacho-Collados, M.T. Pilehvar and R. Navigli, NASARI: A novel approach to a semantically-aware representation of items, in: *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, R. Mihalcea, J. Chai and A. Sarkar, eds, Association for Computational Linguistics, Denver, Colorado, 2015, pp. 567–577, https://www.aclweb.org/anthology/N15-1059. doi:10.3115/v1/N15-1059.

[14] J. Chen, F. Lecue, J.Z. Pan, I. Horrocks and H. Chen, Knowledge-based transfer learning explanation, in: *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference*, M. Thielscher, F. Toni and F. Wolter, eds, AAAI Press, Palo Alto, CA, USA, 2018, pp. 349–358. ISBN 978-1-57735-803-9.

[15] D. Chicco, P. Sadowski and P. Baldi, Deep autoencoder neural networks for gene ontology annotation predictions, in: *Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*, P. Baldi and W. Wang, eds, BCB '14, ACM, New York, NY, USA, 2014, pp. 533–540. ISBN 978-1-4503-2894-4. doi:10.1145/2649387.2649442.

[16] M. Cochez, M. Garofalo, J. Lenßen and M.A. Pellegrino, A first experiment on including text literals in KGloVe, in: *Joint Proceedings of ISWC 2018 Workshops SemDeep-4 and NLIWOD-4*, K.-S. Choi, L.E. Anke, T. Declerck, D. Gromann, J.-D. Kim, A.-C. Ngonga-Ngomo, M. Saleem and R. Usbeck, eds, CEUR Workshop Proceedings, Aachen, Germany, 2018, pp. 103–106.

[17] R. Cotterell and G. Heigold, Cross-lingual character-level neural morphological tagging, in: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, M. Palmer, R. Hwa and S. Riedel, eds, Association for Computational Linguistics, Copenhagen, Denmark, 2017, pp. 748–759, https://www.aclweb.org/anthology/D17-1078. doi:10.18653/v1/D17-1078.

[18] J. Deshmukh, K.M. Annervaz and S. Sengupta, A sequence modeling approach for structured data extraction from unstructured text, in: *Proceedings of the 5th Workshop on Semantic Deep Learning (SemDeep-5)*, L. Espinosa-Anke, T. Declerck, D. Gromann, J. Camacho-Collados and M.T. Pilehvar, eds, Association for Computational Linguistics, Macau, China, 2019, pp. 57–66, https://www.aclweb.org/anthology/W19-5809.

[19] J. Devlin, M.-W. Chang, K. Lee and K. Toutanova, BERT: Pretraining of deep bidirectional transformers for language understanding, in: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, J. Burstein, C. Doran and T. Solorio, eds, Long Papers, Vol. 1, Association for Computational Linguistics, Minneapolis, Minnesota, 2019, pp. 4171–4186, https://www.aclweb.org/anthology/N19-1423. doi:10.18653/v1/N19-1423.

[20] V. Efthymiou, O. Hassanzadeh, M. Rodriguez-Muro and V. Christophides, Matching web tables with knowledge base entities: From entity lookups to entity embeddings, in: *The Semantic Web – ISWC 2017*, C. d'Amato, M. Fernandez, V. Tamma, F. Lecue, P. Cudré-Mauroux, J. Sequeda, C. Lange and J. Heflin, eds, Lecture Notes in Computer Science, Vol. 10587, Springer International Publishing, Vienna, Austria, 2017, pp. 260–277. ISBN 978-3-319-68288-4. doi:10.1007/978-3-319-68288-4_16.

[21] L. Espinosa-Anke, J. Camacho-Collados, C. Delli Bovi and H. Saggion, Supervised distributional hypernym discovery via domain adaptation, in: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, J. Su, K. Duh and X. Carreras, eds, Association for Computational Linguistics, Austin, Texas, 2016,

pp. 424–435, https://www.aclweb.org/anthology/D16-1041. doi:10.18653/v1/D16-1041.

[22] L. Espinosa-Anke, J. Camacho-Collados, S. Rodríguez-Fernández, H. Saggion and L. Wanner, Extending WordNet with fine-grained collocational information via supervised distributional learning, in: *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, Y. Matsumoto and R. Prasad, eds, Association for Computational Linguistics, Osaka, Japan, 2016, pp. 3422–3432, https://www.aclweb.org/anthology/C16-1323.

[23] L. Espinosa-Anke, S. Oramas, H. Saggion and X. Serra, ELMDist: A vector space model with words and MusicBrainz entities, in: *The Semantic Web: ESWC 2017 Satellite Events*, Lecture Notes in Computer Science, Vol. 10577, Springer International Publishing, Cham, Switzerland, 2017, pp. 355–366. ISBN 978-3-319-70407-4. doi:10.1007/978-3-319-70407-4_44.

[24] I. Ezeani, I. Onyenwe and M. Hepple, Transferred embeddings for Igbo similarity, analogy, and diacritic restoration tasks, in: *Proceedings of the Third Workshop on Semantic Deep Learning*, L.E. Anke, D. Gromann and T. Declerck, eds, Association for Computational Linguistics, Santa Fe, New Mexico, 2018, pp. 30–38, https://www.aclweb.org/anthology/W18-4004.

[25] M. Faruqui, J. Dodge, S.K. Jauhar, C. Dyer, E. Hovy and N.A. Smith, Retrofitting word vectors to semantic lexicons, in: *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, R. Mihalcea, J. Chai and A. Sarkar, eds, Association for Computational Linguistics, Denver, Colorado, 2015, pp. 1606–1615, https://www.aclweb.org/anthology/N15-1184. doi:10.3115/v1/N15-1184.

[26] D. Garcia-Gasulla, A. Vilalta, F. Parés, J. Moreno, E. Ayguadé, J. Labarta, U. Cortés and T. Suzumura, Building graph representations of deep vector embeddings, in: *Proceedings of the 2nd Workshop on Semantic Deep Learning (SemDeep-2)*, D. Gromann, T. Declerck and G. Heigl, eds, Association for Computational Linguistics, Montpellier, France, 2017, pp. 9–15, https://www.aclweb.org/anthology/W17-7302.

[27] J. Goikoetxea, E. Agirre and A. Soroa, Single or multiple? Combining word representations independently learned from text and WordNet, in: *Thirtieth AAAI Conference on Artificial Intelligence*, D. Schuurmans and M. Wellman, eds, AAAI Press, Palo Alto, CA, USA, 2016, pp. 2608–2614. ISBN 978-1-57735-760-5.

[28] D. Gromann and T. Declerck, Comparing pretrained multilingual word embeddings on an ontology alignment task, in: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*, N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis and T. Tokunaga, eds, European Language Resources Association (ELRA), Miyazaki, Japan, 2018, pp. 230–236, https://www.aclweb.org/anthology/L18-1034. ISBN 979-10-95546-00-9.

[29] P. Gupta, B. Andrassy and H. Schütze, Replicated Siamese LSTM in ticketing system for similarity learning and retrieval in asymmetric texts, in: *Proceedings of the Third Workshop on Semantic Deep Learning*, L.E. Anke, D. Gromann and T. Declerck, eds, Association for Computational Linguistics, Santa Fe, New Mexico, 2018, pp. 1–11, https://www.aclweb.org/anthology/W18-4001.

[30] P. Hitzler, F. Bianchi, M. Ebrahimi and M.K. Sarker, Neural-symbolic integration and the semantic web, *Semantic Web Journal* (Forthcoming). http://www.semantic-web-journal.net/content/neural-symbolic-integration-and-semantic-web.

[31] A. Holzinger, C. Biemann, C.S. Pattichis and D.B. Kell, What do we need to build explainable AI systems for the medical domain?, *CoRR* (2017), http://arxiv.org/abs/1712.09923.

[32] F. Hommel, P. Cimiano, M. Orlikowski and M. Hartung, Extending neural question answering with linguistic input features, in: *Proceedings of the 5th Workshop on Semantic Deep Learning (SemDeep-5)*, L. Espinosa-Anke, T. Declerck, D. Gromann, J. Camacho-Collados and M.T. Pilehvar, eds, Association for Computational Linguistics, Macau, China, 2019, pp. 31–39, https://www.aclweb.org/anthology/W19-5806.

[33] I. Iacobacci, M.T. Pilehvar and R. Navigli, SensEmbed: Learning sense embeddings for word and relational similarity, in: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, C. Zong and M. Strube, eds, Association for Computational Linguistics, Beijing, China, 2015, pp. 95–105, https://www.aclweb.org/anthology/P15-1010. doi:10.3115/v1/P15-1010.

[34] Y. Jia, R.J. Weiss, F. Biadsy, W. Macherey, M. Johnson, Z. Chen and Y. Wu, Direct speech-to-speech translation with a sequence-to-sequence model, *CoRR* (2019), http://arxiv.org/abs/1904.06037.

[35] E. Jiménez-Ruiz, A. Agibetov, M. Samwald and V. Cross, Breaking-down the ontology alignment task with a lexical index and neural embeddings, *CoRR* (2018), http://arxiv.org/abs/1805.12402.

[36] F.A. Kassar and F. Armetta, Extracting tags from large raw texts using end-to-end memory networks, in: *Proceedings of the 2nd Workshop on Semantic Deep Learning (SemDeep-2)*, D. Gromann, T. Declerck and G. Heigl, eds, Association for Computational Linguistics, Montpellier, France, 2017, pp. 33–40, https://www.aclweb.org/anthology/W17-7305.

[37] Y. Kim, Y. Gao and H. Ney, Effective cross-lingual transfer of neural machine translation models without shared vocabularies, in: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, A. Korhonen, D. Traum and L. Màrquez, eds, Association for Computational Linguistics, Florence, Italy, 2019, pp. 1246–1257, https://www.aclweb.org/anthology/P19-1120.

[38] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami and C. Dyer, Neural architectures for named entity recognition, in: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, K. Knight, A. Nenkova and O. Rambow, eds, Association for Computational Linguistics, San Diego, CA, USA, 2016, pp. 260–270, https://www.aclweb.org/anthology/N16-1030. doi:10.18653/v1/N16-1030.

[39] F. Lecue, On the role of knowledge graphs in explainable AI, *Semantic Web Journal*, http://www.semantic-web-journal.net/content/role-knowledge-graphs-explainable-ai.

[40] A.T. Luu, Y. Tay, S.C. Hui and S.K. Ng, Learning term embeddings for taxonomic relation identification using dynamic weighting neural network, in: *Proceedings of the*

2016 Conference on Empirical Methods in Natural Language Processing, J. Su, K. Duh and X. Carreras, eds, Association for Computational Linguistics, Austin, Texas, 2016, pp. 403–413, https://www.aclweb.org/anthology/D16-1039. doi:10.18653/v1/D16-1039.

[41] S. Magnolini, V. Piccioni, V. Balaraman, M. Guerini and B. Magnini, How to use gazetteers for entity recognition with neural models, in: *Proceedings of the 5th Workshop on Semantic Deep Learning (SemDeep-5)*, L. Espinosa-Anke, T. Declerck, D. Gromann, J. Camacho-Collados and M.T. Pilehvar, eds, Association for Computational Linguistics, Macau, China, 2019, pp. 40–49, https://www.aclweb.org/anthology/W19-5807.

[42] G. Mai, K. Janowicz and B. Yan, Combining text embedding and knowledge graph embedding techniques for academic search engines, in: *Joint Proceedings of ISWC 2018 Workshops SemDeep-4 and NLIWOD-4*, K.-S. Choi, L. Espinosa-Anke, T. Declerck, D. Gromann, J.-D. Kim, A.-C. Ngonga-Ngomo, M. Saleem and R. Usbeck, eds, CEUR Workshop Proceedings, Aachen, Germany, 2018, pp. 77–88.

[43] T. Mikolov, M. Karafiát, L. Burget, J. Cernockỳ and S. Khudanpur, Recurrent neural network based language model, in: *INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association*, T. Kobayashi, K. Hirose and S. Nakamura, eds, Vol. 2, International Speech Communication Association, Makuhari, Japan, 2010, pp. 1045–1048.

[44] T. Murata, Y. Onuki, S. Nukui, S. Inagi, X. Qiu, M. Watanabe and H. Okamoto, Predicting relations between RDF entities by deep neural network, in: *The Semantic Web: ESWC 2017 Satellite Events*, Lecture Notes in Computer Science, Vol. 10577, Springer International Publishing, Cham, Switzerland, 2017, pp. 343–354. ISBN 978-3-319-70407-4. doi:10.1007/978-3-319-70407-4_43.

[45] R. Navigli and S.P. Ponzetto, BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network, *Artificial Intelligence* **193** (2012), 217–250. doi:10.1016/j.artint.2012.07.001.

[46] L. Nieto-Piña and R. Johansson, Automatically linking lexical resources with word sense embedding models, in: *Proceedings of the Third Workshop on Semantic Deep Learning*, L.E. Anke, D. Gromann and T. Declerck, eds, Association for Computational Linguistics, Santa Fe, New Mexico, 2018, pp. 23–29, https://www.aclweb.org/anthology/W18-4003.

[47] S. Parupalli, V.A. Rao and R. Mamidi, Towards enhancing lexical resource and using sense-annotations of OntoSenseNet for sentiment analysis, in: *Proceedings of the Third Workshop on Semantic Deep Learning*, L.E. Anke, D. Gromann and T. Declerck, eds, Association for Computational Linguistics, Santa Fe, New Mexico, 2018, pp. 39–44, https://www.aclweb.org/anthology/W18-4005.

[48] M. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee and L. Zettlemoyer, Deep contextualized word representations, in: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, M. Walker, H. Ji and A. Stent, eds, Long Papers, Vol. 1, Association for Computational Linguistics, New Orleans, Louisiana, 2018, pp. 2227–2237, https://www.aclweb.org/anthology/N18-1202. doi:10.18653/v1/N18-1202.

[49] G. Petrucci, M. Rospocher and C. Ghidini, Expressive ontology learning as neural machine translation, *Journal of Web Semantics* **52** (2018), 66–82. doi:10.1016/j.websem.2018.10.002.

[50] M.T. Pilehvar and N. Collier, De-conflated semantic representations, in: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, J. Su, K. Duh and X. Carreras, eds, Association for Computational Linguistics, Austin, Texas, 2016, pp. 1680–1690, https://www.aclweb.org/anthology/D16-1174. doi:10.18653/v1/D16-1174.

[51] A. Popov, Deep learning architecture for part-of-speech tagging with word and suffix embeddings, in: *Artificial Intelligence: Methodology, Systems, and Applications*, C. Dichev and G. Agre, eds, Lecture Notes in Computer Science, Vol. 9883, Springer International Publishing, Cham, Switzerland, 2016, pp. 68–77. doi:10.1007/978-3-319-44748-3_7.

[52] L. Qiu, J. Yu, Q. Pu and C. Xiang, Knowledge entity learning and representation for ontology matching based on deep neural networks, *Cluster Computing* **20**(2) (2017), 969–977. doi:10.1007/s10586-017-0844-1.

[53] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei and I. Sutskever, Language models are unsupervised multitask learners, *OpenAI Blog* (2019), 1–24, https://d4mucfpksywv.cloudfront.net/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

[54] A. Raganato, C.D. Bovi and R. Navigli, Neural sequence learning models for word sense disambiguation, in: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, M. Palmer, R. Hwa and S. Riedel, eds, Association for Computational Linguistics, Copenhagen, Denmark, 2017, pp. 1156–1167, https://www.aclweb.org/anthology/D17-1120. doi:10.18653/v1/D17-1120.

[55] M.M. Rahman and T. Finin, Understanding and representing the semantics of large structured documents, in: *Joint Proceedings of ISWC 2018 Workshops SemDeep-4 and NLIWOD-4*, K.-S. Choi, L.E. Anke, T. Declerck, D. Gromann, J.-D. Kim, A.-C. Ngonga-Ngomo, M. Saleem and R. Usbeck, eds, Vol. 2241, CEUR Workshop Proceedings, Aachen, Germany, 2018, pp. 65–76.

[56] M. Saeidi, R. Kulkarni, T. Togia and M. Sama, The effect of negative sampling strategy on capturing semantic similarity in document embeddings, in: *Proceedings of the 2nd Workshop on Semantic Deep Learning (SemDeep-2)*, D. Gromann, T. Declerck and G. Heigl, eds, Association for Computational Linguistics, Montpellier, France, 2017, pp. 1–8, https://www.aclweb.org/anthology/W17-7301.

[57] F. Scharffe, Class disjointness constraints as specific objective functions in neural network classifiers, in: *Proceedings of the 2nd Workshop on Semantic Deep Learning (SemDeep-2)*, D. Gromann, T. Declerck and G. Heigl, eds, Association for Computational Linguistics, Montpellier, France, 2017, pp. 16–23, https://www.aclweb.org/anthology/W17-7303.

[58] S. Shekarpour, F. Alshargi, K. Thirunaravan, V.L. Shalin and A. Sheth, CEVO: Comprehensive EVent ontology enhancing cognitive annotation on relations, in: *2019 IEEE 13th International Conference on Semantic Computing (ICSC)*, D. Bulterman, A. Kitazawa, D. Ostrowski, P. Sheu and J. Tsai, eds, IEEE, Newport Beach, CA, USA, 2019, pp. 385–391. doi:10.1109/ICOSC.2019.8665605.

[59] R. Socher, D. Chen, C.D. Manning and A.Y. Ng, Reasoning with neural tensor networks for knowledge base completion, in: *Proceedings of the 26th International Conference on Neural Information Processing Systems – Volume 1*, C.J.C. Burges, L. Bottou, M. Welling, Z. Ghahramani and K.Q. Weinberger, eds, NIPS'13, Curran Associates Inc., USA, 2013, pp. 926–934, http://dl.acm.org/citation.cfm?id=2999611.2999715.

[60] F.M. Suchanek, G. Kasneci and G. Weikum, Yago: A core of semantic knowledge, in: *Proceedings of the 16th International Conference on World Wide Web*, C. Williamson and M.E. Zurko, eds, WWW '07, ACM, New York, NY, USA, 2007, pp. 697–706. ISBN 978-1-59593-654-7. doi:10.1145/1242572.1242667.

[61] I. Sutskever, J. Martens and G. Hinton, Generating text with recurrent neural networks, in: *Proceedings of the 28th International Conference on International Conference on Machine Learning*, L. Getoor and T. Scheffer, eds, ICML'11, Omnipress, Bellevue, Washington, USA, 2011, pp. 1017–1024, http://dl.acm.org/citation.cfm?id=3104482.3104610. ISBN 978-1-4503-0619-5.

[62] S. Thoma, A. Rettinger and F. Both, Towards holistic concept representations: Embedding relational knowledge, visual attributes, and distributional word semantics, in: *The Semantic Web – ISWC 2017*, C. d'Amato, M. Fernandez, V. Tamma, F. Lecue, P. Cudré-Mauroux, J. Sequeda, C. Lange and J. Heflin, eds, Lecture Notes on Computer Science, Vol. 10587, Springer International Publishing, Vienna, Austria, 2017, pp. 694–710. ISBN 978-3-319-68288-4. doi:10.1007/978-3-319-68288-4_41.

[63] R. van Noord and J. Bos, Dealing with co-reference in neural semantic parsing, in: *Proceedings of the 2nd Workshop on Semantic Deep Learning (SemDeep-2)*, D. Gromann, T. Declerck and G. Heigl, eds, Association for Computational Linguistics, Montpellier, France, 2017, pp. 41–49, https://www.aclweb.org/anthology/W17-7306.

[64] A. Vilalta, D. Garcia-Gasulla, F. Parés, J. Moreno, E. Ayguadé, J. Labarta, U. Cortés and T. Suzumura, Full-network embedding in a multimodal embedding pipeline, in: *Proceedings of the 2nd Workshop on Semantic Deep Learning (SemDeep-2)*, D. Gromann, T. Declerck and G. Heigl, eds, Association for Computational Linguistics, Montpellier, France, 2017, pp. 24–32, https://www.aclweb.org/anthology/W17-7304.

[65] J. Völker, D. Fleischhacker and H. Stuckenschmidt, Automatic acquisition of class disjointness, *Journal of Web Semantics* **35**(P2) (2015), 124–139. doi:10.1016/j.websem.2015.07.001.

[66] D. Vrandečić and M. Krötzsch, Wikidata: A free collaborative knowledgebase, *Communications of the ACM* **57**(10) (2014), 78–85. doi:10.1145/2629489.

[67] D. Wang, M. Fang, Y. Song and J. Li, Bridging the gap: Improve part-of-speech tagging for Chinese social media texts with foreign words, in: *Proceedings of the 5th Workshop on Semantic Deep Learning (SemDeep-5)*, L. Espinosa-Anke, T. Declerck, D. Gromann, J. Camacho-Collados and M.T. Pilehvar, eds, Association for Computational Linguistics, Macau, China, 2019, 1–9, https://www.aclweb.org/anthology/W19-5803.

[68] S. Wieczorek, D. Filipiak and A. Filipowska, Semantic image-based profiling of users' interests with neural networks, in: *Emerging Topics in Semantic Technologies – ISWC 2018 Satellite Events [Best Papers from 13 of the Workshops Co-Located with the ISWC 2018 Conference]*, E. Demidova, A. Zaveri and E. Simperl, eds, Studies on the Semantic Web, Vol. 36, IOS Press, 2018, pp. 179–190, ISBN 978-1-61499-893-8. doi:10.3233/978-1-61499-894-5-179.

[69] X. Yin, D. Gromann and S. Rudolph, Neural machine translating from natural language to SPARQL, *CoRR* 2019, http://arxiv.org/abs/1906.09302.

[70] S.-Y. Yoon, A. Loukina, C.M. Lee, M. Mulholland, X. Wang and I. Choi, Word-embedding based content features for automated oral proficiency scoring, in: *Proceedings of the Third Workshop on Semantic Deep Learning*, L.E. Anke, D. Gromann and T. Declerck, eds, Association for Computational Linguistics, Santa Fe, New Mexico, 2018, pp. 12–22, https://www.aclweb.org/anthology/W18-4002.

[71] Y. Zhou, J. Shah and S. Schockaert, Learning household task knowledge from WikiHow descriptions, in: *Proceedings of the 5th Workshop on Semantic Deep Learning (SemDeep-5)*, L. Espinosa-Anke, T. Declerck, D. Gromann, J. Camacho-Collados and M.T. Pilehvar, eds, Association for Computational Linguistics, Macau, China, 2019, pp. 50–56, https://www.aclweb.org/anthology/W19-5808.