Model	In-KB. Acc.
Yamada et al. (2016) [8]	91.5
Ganea&Hofmann (2017) [18]	92.22±0.14
Yang et al. (2018) [14]	93.0
Le&Titov (2018) [20]	93.07±0.27
DeepType (2018) [22]	94.88
Fang et al. (2019) [15]	94.3
Le& Titov (2019) [20]	89.66±0.16
DCA-SL (2019)[12]	94.64±0.2
Chen et al. (2020) [10]	93.54±0.12
DCA-SL + Triples(ours)	**94.94±0.2**

Table 3: ***Generalizability Study: Comparison of KG Context based model against baselines on the AIDA-CONLL dataset. Best value in bold and previous SOTA value is underline.***

from the extra context. The amount of data fed as the context in our models is minimal (up to 15 1-hop triples). In contrast, the best performing model from work in [3], was fed up to 1500 1+2-hop triples. Our best performance can then be attributed to the quality of textual context learned by the transformers as well as the optimal choice of KG-triples context.

Generalizing KG Context: We induced 1-hop KG context in DCA-SL model [12] for candidate entities. The replacement of the unstructured Wikipedia description with structured KG triple context containing entity aliases, entity types, consolidated entity description, etc. has a positive impact on the performance. Our proposed change (DCA-SL + Triples) outperforms the baselines for Wikipedia entity disambiguation(cf. Table 3). Please note, out of 207,544 total entities of AIDA-CoNLL datasets, 7591 entities have no corresponding Wikidata IDs. Even if we do not feed the KG context for 7591 entities, the performance increases. It validates our third research question (**RQ3**), and we conclude KG triple context can be standardized for the NED task for Wikipedia.

6 CONCLUSION

In this paper, we study three closely related research questions. We demonstrate that pretrained Transformer models, although powerful, are limited to capturing context available purely on the texts concerning the original training corpus. We observe that an extra task-specific KG context improved the performance. However, there is a limit to the number of triples as the context that can improve performance. We note that 2-hop triples resulted in negative or little impact on transformer performance. Our triple context can be generalized (for Wikipedia) and observes a positive effect on the NED model for Wikipedia, leading into a new SOTA for AIDA-CoNLL dataset. For the future work, it would be interesting to understand which triples negatively impact the context and how to select the "optimal choice of KG-triples context," considering we rely on the triple in the same order of the SPARQL endpoint returned results. As a viable next step, we plan to study independent effect of various KG attributes (entity properties such as aliases, descriptions, Instance-of, etc.) on NED models' performance.

7 ACKNOWLEDGMENTS

This work is co-funded by the Federal Ministry of Education and Research's (BMBF) Software Campus initiative under the Answer-KiNG Project.

REFERENCES

[1] Krisztian Balog. 2018. *Entity Linking*. Springer International.
[2] Kurt D. Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. 2008. Freebase: a collaboratively created graph database for structuring human knowledge. In *ACM SIGMOD*.
[3] Alberto Cetoli et al. 2019. A Neural Approach to Entity Linking on Wikidata. In *ECIR*. 78–86.
[4] Shuang Chen, Jinpeng Wang, Feng Jiang, and Chin-Yew Lin. 2020. Improving Entity Linking by Modeling Latent Entity Type Information. *CoRR* (2020).
[5] Antonin Delpeuch. 2019. OpenTapioca: Lightweight Entity Linking for Wikidata. *CoRR* (2019).
[6] Adam Lerer et al. 2019. PyTorch-BigGraph: A Large-scale Graph Embedding System. *CoRR* (2019).
[7] Isaiah Mulang' et al. 2020. encoding knowledge graph entity aliases in attentive neural network for wikidata entity linking. *WISE* (2020).
[8] Ikuya Yamada et al. 2016. Joint Learning of the Embedding of Words and Entities for Named Entity Disambiguation. In *CoNLL*.
[9] Johannes Hoffart et al. 2011. Robust Disambiguation of Named Entities in Text. In *EMNLP*.
[10] Shuang Chen et al. 2020. Improving Entity Linking by Modeling Latent Entity Type Information. In *AAAI 2020*.
[11] Weijie Liu et al. 2019. K-BERT: Enabling Language Representation with Knowledge Graph. *ArXiv* (2019).
[12] Xiyuan Yang et al. 2019. Learning Dynamic Context Augmentation for Global Entity Linking. In *EMNLP*.
[13] Yinhan Liu et al. 2019. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *CoRR* (2019).
[14] Yi Yang et al. 2018. Collective Entity Disambiguation with Structured Gradient Tree Boosting. In *NAACL*.
[15] Zheng Fang et al. 2019. Joint Entity Linking with Deep Reinforcement Learning. In *WWW*.
[16] Zhilin Yang et al. 2019. XLNet: Generalized Autoregressive Pretraining for Language Understanding. *CoRR* (2019).
[17] Paolo Ferragina and Ugo Scaiella. 2010. TAGME: on-the-fly annotation of short text fragments (by Wikipedia entities). *CoRR* (2010).
[18] Octavian-Eugen Ganea and Thomas Hofmann. 2017. Deep Joint Entity Disambiguation with Local Neural Attention. In *EMNLP*.
[19] Nikolaos Kolitsas, Octavian-Eugen Ganea, and Thomas Hofmann. 2018. End-to-End Neural Entity Linking. In *In CoNLL*.
[20] Phong Le and Ivan Titov. 2019. Boosting Entity Linking Performance by Leveraging Unlabeled Documents. In *ACL*.
[21] 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. *Semantic Web Journal* (2015).
[22] Jonathan Raiman and Olivier Raiman. 2018. DeepType: Multilingual Entity Linking by Neural Type System Evolution. In *AAAI 2018*.
[23] Ahmad Sakor, Isaiah Onando Mulang', Kuldeep Singh, Saeedeh Shekarpour, Maria Esther Vidal, Jens Lehmann, and Sören Auer. 2019. Old is gold: linguistic driven approach for entity and relation linking of short text. In *NAACL HLT*. 2336–2346.
[24] Ahmad Sakor, Kuldeep Singh, Anery Patel, and Maria-Esther Vidal. 2019. Falcon 2.0: An Entity and Relation Linking Tool over Wikidata. *arXiv preprint arXiv:1912.11270* (2019).
[25] W. Shen, J. Wang, and J. Han. 2015. Entity Linking with a Knowledge Base: Issues, Techniques, and Solutions. *IEEE Transactions on Knowledge and Data Engineering* 27, 2 (2015), 443–460.
[26] Kuldeep Singh, Arun Sethupat Radhakrishna, Andreas Both, Saeedeh Shekarpour, Ioanna Lytra, Ricardo Usbeck, Akhilesh Vyas, Akmal Khikmatullaev, Dharmen Punjani, Christoph Lange, Maria-Esther Vidal, Jens Lehmann, and Sören Auer. 2018. Why Reinvent the Wheel: Let's Build Question Answering Systems Together. In *The Web Conference*.
[27] Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. 2007. Yago: A Core of Semantic Knowledge. In *16th International Conference on the World Wide Web*. 697–706.
[28] Ricardo Usbeck, Axel-Cyrille Ngonga Ngomo, Michael Röder, Daniel Gerber, Sandro Athaide Coelho, Sören Auer, and Andreas Both. 2014. AGDISTIS - Graph-Based Disambiguation of Named Entities Using Linked Data. In *The Semantic Web - ISWC 2014*. Springer, 457–471.
[29] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention Is All You Need. *CoRR* abs/1706.03762 (2017).
[30] Denny Vrandecic. 2012. Wikidata: a new platform for collaborative data collection. In *WWW (Companion Volume)*.