Educational Material to Knowledge Graph Conversion: A Methodology to Enhance Digital Education

Miquel Canal-Esteve

University of Alicante, Carretera de San Vicente del Raspeig, s/n, San Vicente del Raspeig, Alicante, España

Abstract

This paper aims to present a line of research focused on the automatization of structuring digital educational content as knowledge graphs (KGs) to enhance natural language processing tasks. Unlike traditional repositories like Moodle, KGs offer a more flexible representation of relationships between concepts, facilitating intuitive navigation and discovery of connections. By integrating effectively with Large Language Models, KGs can improve personalized explanations, answers, and recommendations. This research will explore and develop technologies for creating and editing educational data (both text and multimedia) and technologies that enable students and teachers to utilize this structured knowledge effectively.

Keywords

Educational Material to Knowledge Graph Conversion, Large Language Models, Automated Knowledge Graph Generation, Intelligent Educational Technologies, Personalized Learning

1. Research Justification

Knowledge graphs (KGs) structure complex information into nodes and relationships, allowing an intuitive and manipulable representation of knowledge. This structure facilitates the integration of information from diverse sources, improves the ability to perform precise semantic searches, and enhances the inference of new knowledge from existing data [1, 2]. Given these capabilities, KGs have shown significant potential across various domains, including education [3].

In the educational environment, KGs can transform how educational information is organized and accessed. They integrate data from multiple sources, such as textbooks, research articles, and online resources, to link key concepts, theories, and relevant authors [4]. For example, In molecular biology, a KG can illustrate the connections between "DNA", "transcription" and "protein synthesis" with references to videos, book chapters, and other resources.

Integration with Large Language Models (LLMs) can enhance this approach, enabling detailed explanations and accurate answers [2]. This approach facilitates the search for specific information for students and educators and helps identify hidden relationships between different topics, promoting deeper, interdisciplinary learning [5].

Although many KGs have been proposed in the literature, due to their complexity, they are

Doctoral Symposium on Natural Language Processing, 26 September 2024, Valladolid, Spain.
mikel.canal@ua.es (M. Canal-Esteve)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

often limited to small environments [6]. The construction of KGs has traditionally required laborious data extraction and linking processes based on natural language processing (NLP) and data mining techniques [2]. However, in recent years, LLMs have revolutionized the field of NLP, demonstrating a remarkable ability to understand and generate natural language and programming. The potential of LLMs for automatic KG generation is an emerging area of research [7, 8].

This research address the problem of converting educational materials into KGs for improved content structuring, navigation, and personalization through LLMs.

2. Background and Related Work

2.1. Knowledge Graphs in the Educational Environment

2.1.1. Representation and Efficient Access to Knowledge

According to Dang et al. [4], efficient access to knowledge is crucial in KGs applied in education. These graphs organize large amounts of information, facilitating understanding and retrieval. Abu-Salih and Alotaibi [5] note that KGs enhance semantic searchability, allowing quick access to specific information.

Abu-Salih and Alotaibi [5] also state that KGs are transforming education by enabling personalized learning and improved curriculum planning. However, challenges include lack of standardized formats, limited interoperability, incomplete data, and scalability. Future research should address these limitations and explore advanced language models and multidomain KGs.

2.1.2. Enhancement of Learning and Discovery of Connections

According to Ain et al. [3], KGs enable dynamic representation of concepts, helping students understand connections between topics, improving retention and contextualized learning.

KGs also enhance educational systems' ability to provide personalized recommendations. Chicaiza and Valdiviezo-Diaz [9] show that mapping relationships between concepts and resources optimizes learning by aligning with students' progress and interests, revealing new connections in a non-linear learning environment.

Stancin et al. [10] highlight the role of ontologies in structuring knowledge and managing curricula. Combining various methodologies, researchers have increasingly used ontologies in education, showing their importance and potential.

2.1.3. Personalization and Integration with LLMs

Research by Li et al. [11] shows KGs improve content organization and personalization in online learning platforms, offering recommendations based on learner progress and interests.

KGs are also crucial in intelligent tutoring systems. Li and Wang [12] state that KGs enable virtual tutors to provide tailored explanations.

Khoiruddin et al. [13] reviews the development of e-learning ontologies, emphasizing methodologies like NeON and METHONTOLOGY and metrics like Relationship Richness to assess quality. Proper application of these methods can enhance e-learning systems.

Chen et al. [14] describe KnowEDu, a system that constructs KGs using pedagogical and assessment data via NLP algorithms, providing a foundation for implementing educational KGs. This method is relevant to Section 4.

2.2. Text-to-Knowledge graph conversion models

The first step to convert educational material to KG is to convert text to KG, often using a LLM [15]. Many integrations exist between LLMs and KGs, but these cover only part of the text-to-knowledge graph process, as seen in the review [7]. Below is an analysis of models that perform the complete task of moving from text to KG.

Common features and differences are noted in these models. They are evaluated in Zero-Shot, One-Shot, and Few-Shot scenarios, measuring datasets' accuracy and semantic relatedness. Differences lie in the base LLMs, fine-tuning techniques, and specific architectures used. Results show improvements in some configurations, but there is still room to optimize the accuracy and efficiency of KG generation.

For instance, in the study by Giglou et al. [16] several models are evaluated on the text to OWL conversion task in Zero-Shot, including BERT-Large [17], PubMedBERT [18], BART-Large [19], Flan-T5-Large [20], Flan-T5-XL [20], BLOOM-1b7 [21], BLOOM-3b [21], GPT-3 [22], GPT-3.5 [23], LLaMA [24] and GPT-4 [25]. These models were tested on the term typing task using different datasets: WordNet [26], GeoNames [27], NCI [28], SNOMEDCT_US [29] and MEDCIN [30]. The best results were 91.7 for WordNet [26], but significantly lower for the other datasets, with scores of 43.3, 16.1, 37.7 and 29.8, respectively, evidencing considerable room for improvement in the models' ability for this task. They were also evaluated in the entity classification task with the GeoNames [27], UMLS [31], and schema.org datasets, showing scores of 67.8, 78.1 and 74.4, again suggesting considerable room for improvement. Finally, in the relationship recognition task with the UMLS [31] dataset, a result of 49.5 was obtained, reflecting once again the need for improvement.

Moreover, the same article presents two tuned models: Flan-T5-Large [20] and Flan-T5-XL [20], which show remarkable improvements in several datasets of the evaluated tasks. For example, for the datasets of the first task, the results were improved to 32.8, 43.4 and 51.8. The results improved to 79.3 and 91.7 in the entity classification task, and in the relationship recognition task, 53.1 was achieved.

Similarly, in the study by Mihindukulasooriya et al. [32] Vicuna-13B [33] and Alpaca-LoRA-13B [34, 35] are evaluated in Zero-Shot on the Fact Extraction task using the F1 metric for different subsets of the Wikidata-TekGen [36] and DBpedia-WebNLG [37] datasets. The best result for the Wikidata dataset [36] is 0.38 for Vicuna [33] and 0.28 for Alpaca [34, 35] and for the DBpedia dataset [37] it is 0.3 for Vicuna [33] and 0.25 for Alpaca [34, 35]. As in the previous case, it is evident that there is much room for improvement.

Furthermore, in the study by Zhu et al. [2], a comprehensive evaluation of Extended Language Models (LLMs) such as GPT-4 [25] and ChatGPT[23] in KG construction and reasoning tasks is performed by experiments on eight datasets and four representative tasks: entity and relationship extraction, event extraction, link prediction, and question and answer. The results show that, although GPT-4 achieves an F1 score of 31.03 in relation extraction on DuIE2.0 [11] on zero-shot and 41.91 on one-shot, as well as an F1 score of 34.2 on MAVEN [38] for event extraction on zero-shot, and a hits@1 of 32.0 on FB15K-237 [39] for link prediction on zero-shot, these results are improbable.

The paper by Melnyk et al. [8] presents an innovative approach for generating KGs from text in multiple stages. This approach is divided into two main phases: first, the generation of nodes using the pre-trained language model T5-large [20] and then the construction of edges using the information from the generated nodes. This method seeks to overcome the limitations of traditional graph linearization approaches by breaking the process into manageable and separately optimizable steps. The model was evaluated on three datasets: WebNLG 2020 [40], TEKGEN [41] and New York Times [42], obtaining F1 scores of 0.722, 0.707 and 0.918 respectively, demonstrating its effectiveness. However, it highlights the need for further improvement, especially in edge generation, to optimize the system's performance in various applications.

Finally, in the study by Ain et al. [3], embeddings-based methods, such as SIFRank [43] and SIFRankplus, which is an extension made by the authors, enhanced with SqueezeBERT [44], achieved an F1-score of 40.38% in keyphrase extraction. In concept weighting, the SBERT-based [45] strategy achieved an accuracy of 13.9% and an F1-score of 20.6% for the top ten ranked concepts, superior results to the benchmark models with which they were purchased. Despite these advances, the results highlight the need to improve the accuracy and performance of the techniques to ensure the effective construction of KGs.

3. Hypothesis and Objectives

The main hypothesis of the research is that it is feasible to research and develop technologies that convert teaching materials into KGs and integrate these with large-scale language models. This integration aims to enhance various education-related natural language processing tasks.

The main objective of the research is to design and implement these technologies, focusing on the automatic transformation of teaching materials into KGs and their integration with language models. This will address tasks in education-related natural language processing. To achieve this objective, the following specific goals are proposed:

- To study the state of the art to identify the most relevant existing solution alternatives in the domain and the main evaluation resources.
- To investigate and develop technologies that allow the creation of advanced tools based on language models, designed to language models, designed to convert texts from multiple disciplines into KGs. This approach will have particular applications in the educational environment, facilitating efficient capture and organization of knowledge.
- Research and develop technologies that allow the integration of large-scale language models with previously developed tools to enrich and expand KGs, in addition to and expand KGs, as well as generate personalized text and answers based on the information contained in the graphs.

4. Methodology

This section presents an innovative methodology for automatically using an LLM to generate KGs from educational materials. Existing models like BERT-Large, GPT-4, Vicuna-13B, PubMedBERT, BART-Large, Flan-T5, BLOOM, GPT-3, GPT-3.5, LLaMA, and Alpaca-LoRA-13B have shown progress in converting text to KGs but still have significant limitations, as seen in the previous section. For example, in term typing tasks, scores were 43.3 for GeoNames, 16.1 for NCI, 37.7 for SNOMEDCT_US, and 29.8 for MEDCIN, compared to 91.7 for WordNet. In entity classification, the highest scores were 78.1 for UMLS and 74.4 for schema.org. Fact extraction tasks showed Vicuna-13B scoring 0.38 and Alpaca-LoRA-13B scoring 0.28 on Wikidata-TekGen. These results highlight the need for new strategies to improve model performance in text-to-KG conversion in general and particularly in education.

To address these limitations, it proposes a methodology based on creating an expert model in natural language and KGs, refined to convert learning materials to KGs following a structured learning object for a guided teaching experience with multimedia content. This includes two phases: continual pre-training with a large dataset of KGs in OWL, RDF, and similar formats, and specific fine-tuning with didactic materials. In pre-training, a varied dataset of KGs from various disciplines trains the model using masking and self-supervised learning, enhancing its understanding of semantic relationships and hierarchical structures in KGs, improving its ability to generate coherent and accurate graphs.

Continual pre-training allows the model to become more expert in the domain in which it is pre-trained [46]; in this case, it is believed that it would involve improved semantic understanding, training on structured data, flexibility and generalization, reduction of biases, and leveraging of existing resources.

In the fine-tuning phase, diverse educational materials will be gathered, and their corresponding KGs will be created manually or semi-automatically. This process will require defining a KG scheme or reusing one already described in the literature that fits the proposed use case. For this phase, the schemes, and methodologies described in the studies [47] and [14].

Although KGs are not used in [47], it becomes clear that a small amount of domain-specific data, such as slides and lecture transcripts, can be extremely valuable for building knowledge-based and generative educational chatbots. Slides are enriched with semantic annotations, identifying entities such as definitions, quotes, and examples. This enables knowledge-based to provide accurate and relevant responses by mining directly from this structured data.

Chen et al. [14] describes a system developed to build educational KGs using pedagogical and learning assessment data automatically. The methods used in this study for extracting instructional concepts and identifying meaningful educational relationships will provide a solid foundation for the proposed KG scheme. Integrating these methodologies is expected to improve the system's effectiveness in automatically generating KGs from educational materials.

5. Research Issues to Discuss

In the first phase of the research we are currently in, continual pretraining will be performed on LLaMA3-8b. A dataset with public KGs ranging from 10 to 50GB is being prepared. This dataset is being characterized based on the themes of the KG and other semantic KG such as the number of classes, the depth of the KG, the density of relationships, etc., as well as linguistic metrics like the number of tokens.

Once continual pretraining is completed, the model's ability to complete OWL code and perform other NLP tasks, such as those mentioned in the Background and Related Work, will be evaluated to ensure it has not forgotten natural language. Subsequently, a second phase of fine-tuning will be conducted for specific semantic tasks such as link prediction, entity recognition, and KG completion.

After the model has been trained and evaluated on these tasks, it will be instructed to perform the task of converting educational material into a KG. This will require defining a reference KG and manually (or semi-automatically) populating it with several examples so that the model can learn to perform this task during the instruction phase.

Key issues to discuss in this phase include:

1. **Dataset Preparation**: Ensuring the dataset is diverse and representative of various domains to avoid bias and enhance the model's generalization capabilities.
2. **Evaluation Metrics**: Deciding on appropriate metrics for evaluating the model's performance in OWL code completion and NLP tasks, ensuring comprehensive assessment.
3. **Knowledge Graph Definition and Population**: Developing a robust and flexible reference KG and strategies for its manual or semi-automatic population.
4. **Instruction Phase Design**: Designing an effective instruction phase to train the model on converting educational materials to KGs, including selecting examples and defining evaluation criteria.

These discussions will guide the research process, ensuring methodological rigor and the development of an effective system for converting educational materials into KGs integrated with large-scale language models.

References

[1] M. Kejriwal, Knowledge graphs: A practical review of the research landscape, Information 13 (2022) 161. URL: https://www.mdpi.com/2078-2489/13/4/161. doi:10.3390/info13040161.

[2] Y. Zhu, X. Wang, J. Chen, S. Qiao, Y. Ou, Y. Yao, S. Deng, H. Chen, N. Zhang, LLMs for Knowledge Graph Construction and Reasoning: Recent Capabilities and Future Opportunities, arXiv e-prints (2023) arXiv:2305.13168. doi:10.48550/arXiv.2305.13168. arXiv:2305.13168.

[3] Q. U. Ain, M. A. Chatti, K. G. C. Bakar, S. Joarder, R. Alatrash, Automatic construction of educational knowledge graphs: A word embedding-based approach, Information 14 (2023) 526. URL: https://www.mdpi.com/2078-2489/14/10/526. doi:10.3390/info14100526.

[4] F.-R. Dang, J.-T. Tang, K.-Y. Pang, T. Wang, S.-S. Li, X. Li, Constructing an educational knowledge graph with concepts linked to wikipedia, Journal of Computer Science and Technology 36 (2021) 1200–1211. doi:https://doi.org/10.1007/s11390-020-0328-2.

[5] B. Abu-Salih, S. Alotaibi, A systematic literature review of knowledge graph construction and application in education, Heliyon 10 (2024) e25383. doi:10.1016/j.heliyon.2024.e25383.

[6] X. Yuan, J. Chen, Y. Wang, A. Chen, Y. Huang, W. Zhao, S. Yu, Semantic-enhanced knowledge graph completion, Mathematics 12 (2024). URL: https://www.mdpi.com/2227-7390/12/3/450. doi:10.3390/math12030450.

[7] S. Pan, L. Luo, Y. Wang, C. Chen, J. Wang, X. Wu, Unifying Large Language Models and Knowledge Graphs: A Roadmap, arXiv e-prints (2023) arXiv:2306.08302. doi:10.48550/arXiv.2306.08302. arXiv:2306.08302.
[8] I. Melnyk, P. Dognin, P. Das, Knowledge Graph Generation From Text, arXiv e-prints (2022) arXiv:2211.10511. doi:10.48550/arXiv.2211.10511. arXiv:2211.10511.
[9] J. Chicaiza, P. Valdiviezo-Diaz, A comprehensive survey of knowledge graph-based recommender systems: Technologies, development, and contributions, Information 12 (2021) 232. URL: https://www.mdpi.com/2078-2489/12/6/232. doi:10.3390/info12060232.
[10] K. Stancin, P. Poscic, D. Jaksic, Ontologies in education – state of the art, Education and Information Technologies 25 (2020) 5301–5320. doi:10.1007/s10639-020-10226-z.
[11] S. Li, J. Tang, M. Kan, D. Zhao, S. Li, H. Zan, Duie: A large-scale chinese dataset for information extraction, Natural Language Processing and Chinese Computing 11839 (2019). doi:10.1007/978-3-030-32236-6_72.
[12] L. Li, Z. Wang, Knowledge Graph Enhanced Intelligent Tutoring System Based on Exercise Representativeness and Informativeness, arXiv e-prints (2023) arXiv:2307.15076. doi:10.48550/arXiv.2307.15076. arXiv:2307.15076.
[13] M. Khoiruddin, S. Kusumawardani, I. Hidayah, S. Fauziati, A review of ontology development in the e-learning domain: Methods, roles, evaluation, 2023 International Conference on Computer, Control, Informatics and its Applications (IC3INA) (2023). doi:10.1109/IC3INA60834.2023.10285789.
[14] P. Chen, Y. Lu, V. W. Zheng, X. Chen, B. Yang, Knowedu: A system to construct knowledge graph for education, IEEE Access 6 (2018) 31553–31563. doi:10.1109/ACCESS.2018.2839607.
[15] M. Trajanoska, R. Stojanov, D. Trajanov, Enhancing knowledge graph construction using large language models, ArXiv abs/2305.04676 (2023). URL: https://api.semanticscholar.org/CorpusID:258557103.
[16] H. Giglou, J. D'Souza, S. Auer, LLMs4OL: Large Language Models for Ontology Learning, arXiv e-prints (2023) arXiv:2307.16648. doi:10.48550/arXiv.2307.16648. arXiv:2307.16648.
[17] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding (2019) 4171–4186. URL: https://aclanthology.org/N19-1423. doi:10.18653/v1/N19-1423.
[18] Y. Gu, R. Tinn, H. Cheng, M. Lucas, N. Usuyama, X. Liu, T. Naumann, J. Gao, H. Poon, Domain-specific language model pretraining for biomedical natural language processing, ACM Trans. Comput. Healthcare 3 (2021). URL: https://doi.org/10.1145/3458754. doi:10.1145/3458754.
[19] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, L. Zettlemoyer, BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension (2020) 7871–7880. URL: https://aclanthology.org/2020.acl-main.703. doi:10.18653/v1/2020.acl-main.703.
[20] H. W. Chung, L. Hou, S. Longpre, B. Zoph, Y. Tay, W. Fedus, Y. Li, X. Wang, M. Dehghani, S. Brahma, A. Webson, S. S. Gu, Z. Dai, M. Suzgun, X. Chen, A. Chowdhery, A. Castro-Ros, M. Pellat, K. Robinson, D. Valter, S. Narang, G. Mishra, A. Yu, V. Zhao, Y. Huang, A. Dai, H. Yu, S. Petrov, E. H. Chi, J. Dean, J. Devlin, A. Roberts, D. Zhou, Q. V. Le, J. Wei,

Scaling Instruction-Finetuned Language Models, arXiv e-prints (2022) arXiv:2210.11416. doi:10.48550/arXiv.2210.11416. arXiv:2210.11416.
[21] B. Workshop, T. Le Scao, A. Fan, C. Akiki, E. Pavlick, S. Ilić, D. Hesslow, R. Castagné, A. Sasha Luccioni, F. Yvon, M. Gallé, J. Tow, A. M. Rush, S. Biderman, A. Webson, P. Sasanka Ammanamanchi, T. Wang, B. Sagot, N. Muennighoff, A. Villanova del Moral, et al., BLOOM: A 176B-Parameter Open-Access Multilingual Language Model, arXiv e-prints (2022) arXiv:2211.05100. doi:10.48550/arXiv.2211.05100. arXiv:2211.05100.
[22] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language Models are Few-Shot Learners, arXiv e-prints (2020) arXiv:2005.14165. doi:10.48550/arXiv.2005.14165. arXiv:2005.14165.
[23] OpenAI, ChatGPT: Language Model, 2023. URL: https://www.openai.com/chatgpt, accessed: 2024-05-21.
[24] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, A. Rodriguez, A. Joulin, E. Grave, G. Lample, LLaMA: Open and Efficient Foundation Language Models, arXiv e-prints (2023) arXiv:2302.13971. doi:10.48550/arXiv.2302.13971. arXiv:2302.13971.
[25] OpenAI, J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. Leoni Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, R. Avila, I. Babuschkin, S. Balaji, V. Balcom, P. Baltescu, H. Bao, M. Bavarian, J. Belgum, I. Bello, Berdine, et al., GPT-4 Technical Report, arXiv e-prints (2023) arXiv:2303.08774. doi:10.48550/arXiv.2303.08774. arXiv:2303.08774.
[26] G. A. Miller, Wordnet: a lexical database for english, Commun. ACM 38 (1995) 39–41. URL: https://doi.org/10.1145/219717.219748. doi:10.1145/219717.219748.
[27] T. Rebele, F. Suchanek, J. Hoffart, J. Biega, E. Kuzey, G. Weikum, Yago: A multilingual knowledge base from wikipedia, wordnet, and geonames, International Semantic Web Conference (2016) 177–185. doi:10.1007/978-3-319-46547-0_19.
[28] National Cancer Institute, National Institutes of Health, NCI Thesaurus, 2022. URL: http://ncit.nci.nih.gov, accessed: 2024-05-21.
[29] SNOMED International, US Edition of SNOMED CT, 2023. URL: https://www.nlm.nih.gov/healthit/snomedct/us_edition.html, accessed: 2024-05-21.
[30] Medicomp Systems, MEDCIN, 2023. URL: https://medicomp.com, accessed: 2024-05-21.
[31] O. Bodenreider, The unified medical language system (umls): integrating biomedical terminology, Nucleic acids research 32 (2004) D267–D270. doi:https://doi.org/10.1093/nar/gkh061.
[32] N. Mihindukulasooriya, S. Tiwari, C. F. Enguix, K. Lata, Text2KGBench: A Benchmark for Ontology-Driven Knowledge Graph Generation from Text, arXiv e-prints (2023) arXiv:2308.02357. doi:10.48550/arXiv.2308.02357. arXiv:2308.02357.
[33] Chiang, Wei-Lin and Li, Zhuohan and Lin, Zi and Sheng, Ying and Wu, Zhanghao and Zhang, Hao and Zheng, Lianmin and Zhuang, Siyuan and Zhuang, Yonghao and Gonzalez, Joseph E. and Stoica, Ion and Xing, Eric P., Vicuna: An open-source chatbot impressing GPT-4 with 90% ChatGPT quality, 2023. URL: https://vicuna.lmsys.org, accessed: 2024-05-21.

[34] R. Taori, I. Gulrajani, T. Zhang, Y. Dubois, X. Li, C. Guestrin, P. Liang, T. B. Hashimoto, Stanford Alpaca: An instruction-following LLaMA model, 2023. URL: https://github.com/tatsu-lab/stanford_alpaca, accessed: 2023-05-21.

[35] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen, LoRA: Low-rank adaptation of large language models, 2022. URL: https://openreview.net/forum?id=nZeVKeeFYf9, accessed: 2023-05-21.

[36] D. Vrandečić, M. Krötzsch, Wikidata: a free collaborative knowledgebase, Commun. ACM 57 (2014) 78–85. URL: https://doi.org/10.1145/2629489. doi:10.1145/2629489.

[37] C. Gardent, A. Shimorina, S. Narayan, L. Perez-Beltrachini, Creating training corpora for NLG micro-planners (2017) 179–188. URL: https://aclanthology.org/P17-1017. doi:10.18653/v1/P17-1017.

[38] X. Wang, Z. Wang, X. Han, W. Jiang, R. Han, Z. Liu, J. Li, P. Li, Y. Lin, J. Zhou, MAVEN: A Massive General Domain Event Detection Dataset (2020) 1652–1671. URL: https://aclanthology.org/2020.emnlp-main.129. doi:10.18653/v1/2020.emnlp-main.129.

[39] K. Toutanova, D. Chen, P. Pantel, H. Poon, P. Choudhury, M. Gamon, Representing text for joint embedding of text and knowledge bases (2015) 1499–1509. URL: https://aclanthology.org/D15-1174. doi:10.18653/v1/D15-1174.

[40] T. Castro Ferreira, C. Gardent, N. Ilinykh, C. van der Lee, S. Mille, D. Moussallem, A. Shimorina, The 2020 bilingual, bi-directional WebNLG+ shared task: Overview and evaluation results (WebNLG+ 2020) (2020) 55–76. URL: https://aclanthology.org/2020.webnlg-1.7.

[41] O. Agarwal, H. Ge, S. Shakeri, R. Al-Rfou, Knowledge graph based synthetic corpus generation for knowledge-enhanced language model pre-training (2021) 3554–3565. URL: https://aclanthology.org/2021.naacl-main.278. doi:10.18653/v1/2021.naacl-main.278.

[42] S. Riedel, L. Yao, A. McCallum, Modeling relations and their mentions without labeled text, Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2010. Lecture Notes in Computer Science 6323 (2010). doi:10.1007/978-3-642-15939-8_10.

[43] Y. Sun, H. Qiu, Y. Zheng, Z. Wang, C. Zhang, Sifrank: A new baseline for unsupervised keyphrase extraction based on pre-trained language model, IEEE Access 8 (2020) 10896–10906. doi:10.1109/ACCESS.2020.2965087.

[44] F. Iandola, A. Shaw, R. Krishna, K. Keutzer, SqueezeBERT: What can computer vision teach NLP about efficient neural networks? (2020) 124–135. URL: https://aclanthology.org/2020.sustainlp-1.17. doi:10.18653/v1/2020.sustainlp-1.17.

[45] N. Reimers, I. Gurevych, Sentence-BERT: Sentence embeddings using Siamese BERT-networks (2019) 3982–3992. URL: https://aclanthology.org/D19-1410. doi:10.18653/v1/D19-1410.

[46] T. Wu, L. Luo, Y.-F. Li, S. Pan, T.-T. Vu, G. Haffari, Continual Learning for Large Language Models: A Survey, arXiv e-prints (2024) arXiv:2402.01364. doi:10.48550/arXiv.2402.01364. arXiv:2402.01364.

[47] M. Wölfel, M. B. Shirzad, A. Reich, K. Anderer, Knowledge-based and generative-ai-driven pedagogical conversational agents: A comparative study of grice's cooperative principles and trust, Big Data and Cognitive Computing 8 (2024). URL: https://www.mdpi.com/2504-2289/8/1/2. doi:10.3390/bdcc8010002.