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

Miquel Canal-Esteve and Yoan Gutiérrez
Research Group of Language Processing and Information System
University of Alicante, Spain

Abstract

This article argues that digital educational content should be structured as knowledge graphs (KGs). Unlike traditional repositories such as Moodle, a KG offers a more flexible representation of the relationships between concepts, facilitating intuitive navigation and discovery of connections. In addition, it integrates effectively with Large Language Models, enhancing personalized explanations, answers, and recommendations. This article studies different proposals based on semantics and knowledge modelling to determine the most appropriate ways to strengthen intelligent educational technologies.

1 Introduction

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 (Kejriwal, 2022; Zhu et al., 2023). Given these capabilities, KGs have shown significant potential across various domains, including education (Ain et al., 2023).

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 (Dang et al., 2021). In addition, integration with Large Language Models (LLMs) can enhance this approach, enabling detailed explanations and accurate answers (Zhu et al., 2023). This approach facilitates the search for specific information for students and educators and helps identify hidden relationships between different topics, promoting deeper, interdisciplinary learning (Abu-Salih and Alotaibi, 2024).

Although many KGs have been proposed in the literature, due to their complexity, they are often limited to small environments (Yuan et al., 2024). The construction of KGs has traditionally required laborious data extraction and linking processes based on natural language processing (NLP) and data mining techniques (Zhu et al., 2023). 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 (Pan et al., 2023; Melnyk et al., 2022).

To address the problem of converting educational materials into KGs for improved content structuring, navigation, and personalization with large language models, this paper explores several key areas:

- Identifying the advantages of using KGs in the educational environment.
- Highlighting the most relevant KGs in education and their significant contributions.
- Examining the latest models based on LLMs that facilitate the conversion from text to KG.
- Proposing an innovative approach to enhance the educational material to KG task.

2 Advantages of using knowledge graphs in the educational environment

2.1 Representation and efficient access to knowledge

As indicated in Dang et al. (2021), representation and efficient access to knowledge is fundamental in KGs applied in education. These graphs allow large amounts of information to be organized and visualized in a structured manner, facilitating understanding and retrieval of relevant data. Abu-Salih

Proceedings of the 1st Workshop on Knowledge Graphs and Large Language Models (KaLLM 2024), pages 85–91
August 15, 2024

and Alotaibi (2024) note that KGs significantly improve semantic searchability, allowing students and educators to access the specific information they need quickly.

2.2 Enhancement of learning and discovery of connections

According to Ain et al. (2023), KGs facilitate a more flexible and dynamic representation of concepts and their interrelationships, allowing students to explore and better understand how different topics are connected. This approach improves information retention and fosters deeper and more contextualized learning.

Furthermore, KGs can significantly improve the ability of educational systems to provide personalized and relevant recommendations. Chicaiza and Valdiviezo-Diaz (2021) demonstrate that systems can suggest materials integrated into the student's learning process by mapping the relationships between concepts and educational resources. This optimizes the learning process by aligning with each student's progress and specific interests and facilitates discovering new connections and areas of interest that might not be evident in a more traditional, linear learning environment.

2.3 Personalization and integration with LLMs

Research by Li et al. (2019) analyses the use of KGs in online learning platforms. The authors find that these graphs improve the organization of educational content and facilitate learning personalization. Educational systems using KGs can provide content recommendations based on each learner's progress and interests.

In addition, KGs can play a crucial role in creating intelligent tutoring systems. According to Li and Wang (2023), these graphs enable virtual tutors to provide more detailed explanations tailored to the individual needs of learners.

3 Review of knowledge graphs in education

This section discusses three recent studies that review using KGs and ontologies in education. Each study addresses different aspects and applications of these technologies, assessing their impact and challenges. The conclusions of each of these studies are then presented, providing a comprehensive view of the current and future state of KGs in education. Additionally, we add the article (Chen et al., 2018) that proposes a methodology to build KGs in the educational environment. The proposed scheme will be relevant to the proposed method in Section 5.

Abu-Salih and Alotaibi (2024) conclude that KGs are transforming education by providing personalized learning experiences and enriched data for curriculum planning. However, they face challenges such as a lack of standardized formats, limited interoperability, incomplete data, and scalability issues. Future research is suggested to address these limitations and explore integrating advanced language models and creating multidomain KGs.

Stancin et al. (2020) highlights the crucial role of ontologies in educational systems, facilitating structured knowledge representation and curriculum management. Although there is no single methodology for their construction, researchers combine several methodologies. Recent literature review shows an increase in the use of ontologies in education, highlighting their importance and future potential.

Khoiruddin et al. (2023) reviews the development of ontologies in e-learning, highlighting methodologies such as NeON and METHONTOLOGY, and the roles of domain experts, developers, and end users. It uses metrics such as Relationship Richness to assess the quality of ontologies. He concludes that a proper understanding and application of these methods and metrics can improve the efficiency and effectiveness of e-learning systems.

Finally, Chen et al. (2018) describes a system called KnowEDu developed to automatically construct KGs in education using pedagogical and learning assessment data. KnowEdu uses NLP algorithms to extract meaningful instructional concepts and educational relationships from heterogeneous data. The methods and results of this study provide a solid foundation for the practical implementation of educational KGs. However, this methodology does not allow for an automatic transition from text to KG.

4 Text-to-Knowledge graph conversion models

Many integrations exist between LLMs and KGss, but these only cover one of the text-to-knowledge graph process's tasks, as seen in the review (Pan et al., 2023). An analysis of models that perform the complete task of moving from text to KG is shown below. Several common features and differences are observed in these models. The models are commonly evaluated in Zero-Shot, One-Shot, and Few-Shot scenarios, measuring various datasets' accuracy and semantic relatedness capability. The differences lie in the base LLMs chosen, the fine-tuning techniques applied, and the specific architectures used. The results show that, although there are improvements in certain configurations, there is still ample room to optimize the accuracy and efficiency of KG generation.

For instance, in the study by Giglou et al. (2023) several models are evaluated on the text to OWL conversion task in Zero-Shot, including BERT-Large (Devlin et al., 2019), PubMedBERT (Gu et al., 2021), BART-Large (Lewis et al., 2020), Flan-T5-Large (Chung et al., 2022), Flan-T5-XL (Chung et al., 2022), BLOOM-1b7 (Workshop et al., 2022), BLOOM-3b (Workshop et al., 2022), GPT-3 (Brown et al., 2020), GPT-3.5 (OpenAI, 2023), LLaMA (Touvron et al., 2023) and GPT-4 (OpenAI et al., 2023). These models were tested on the term typing task using different datasets: WordNet (Miller, 1995), GeoNames (Rebele et al., 2016), NCI (National Cancer Institute, National Institutes of Health, 2022), SNOMEDCT_US (SNOMED International, 2023) and MEDCIN (Medicomp Systems, 2023). The best results were 91.7 for WordNet (Miller, 1995), 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 (Rebele et al., 2016), UMLS (Bodenreider, 2004), 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 (Bodenreider, 2004) 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 (Chung et al., 2022) and Flan-T5-XL (Chung et al., 2022), 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. (2023) Vicuna-13B (Chiang et al., 2023) and Alpaca-LoRA-13B (Taori et al., 2023; Hu et al., 2022) are evaluated in Zero-Shot on the Fact Extraction task using the F1 metric for different subsets of the Wikidata-TekGen (Vrandečić and Krötzsch, 2014) and DBpedia-WebNLG (Gardent et al., 2017) datasets. The best result for the Wikidata dataset (Vrandečić and Krötzsch, 2014) is 0.38 for Vicuna (Chiang et al., 2023) and 0.28 for Alpaca (Taori et al., 2023; Hu et al., 2022) and for the DBpedia dataset (Gardent et al., 2017) it is 0.3 for Vicuna (Chiang et al., 2023) and 0.25 for Alpaca (Taori et al., 2023; Hu et al., 2022). As in the previous case, it is evident that there is much room for improvement.

Furthermore, in the study by Zhu et al. (2023), a comprehensive evaluation of Extended Language Models (LLMs) such as GPT-4 (OpenAI et al., 2023) and ChatGPT(OpenAI, 2023) 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 (Li et al., 2019) on zero-shot and 41.91 on one-shot, as well as an F1 score of 34.2 on MAVEN (Wang et al., 2020) for event extraction on zero-shot, and a hits@1 of 32.0 on FB15K-237 (Toutanova et al., 2015) for link prediction on zero-shot, these results are improbable.

The paper by Melnyk et al. (2022) 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 (Chung et al., 2022) 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 (Castro Ferreira et al., 2020), TEKGEN (Agarwal et al., 2021) and New York Times (Riedel et al., 2010), 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. (2023), embeddings-based methods, such as SIFRank (Sun et al., 2020) and SIFRankplus, which is an extension made by the authors, enhanced with SqueezeBERT (Iandola et al., 2020), achieved an F1-score of 40.38% in keyphrase extraction. In concept weighting, the SBERT-based (Reimers and Gurevych, 2019) 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.

5 Proposed 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-knowledge graph conversion in general and particularly in education.

To address these limitations, we propose a methodology that involves creating an expert model in natural language and KG language. This model is subsequently refined to convert learning materials into KGs, following a learning object structure that offers a guided and comprehensive teaching experience with multimedia educational content. The methodology comprises two phases: continuous pre-training using a large dataset of KGs and specific fine-tuning with didactic materials.

During pre-training, a diverse dataset of KGs from sources like Wikidata (Vrandečić and Krötzsch, 2014), DBpedia (Lehmann et al., 2015), and YAGO (Rebele et al., 2016) will be used to train the model with masking and self-supervised learning. This will enhance the model's understanding of semantic relationships and hierarchical structures, improving its ability to generate coherent and accurate graphs.

Continual pre-training allows the model to become more expert in its domain, enhancing semantic understanding, training on structured data, flexibility, generalization, bias reduction, and leveraging existing resources (Wu et al., 2024).

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 necessitate defining a KG schema or leveraging an existing one from the literature that aligns with the proposed use case. Specifically, the of the IEEE Computer Society (2020) provides a comprehensive schema and vocabulary for metadata that could be particularly useful. Alongside this standard, methodologies and schemes described in the studies by (Wölfel et al., 2024) and (Chen et al., 2018) will also be considered.

Although KGs are not used in Wölfel et al. (2024), 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. (2018) 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.

6 Conclusion

In conclusion, this article argues that structuring digital educational content as KGs rather than traditional repositories provides significant advantages. KGs offer a flexible, navigable representation of concept relationships, enhancing learning personalization and integration with LLMs. A methodology to automatically generate KGs from educational texts is proposed, promising to transform access to and organization of educational information for more profound, personalized learning.

References

Bilal Abu-Salih and Salihah Alotaibi. 2024. A systematic literature review of knowledge graph construction and application in education. *Heliyon*, 10(3):e25383.

Oshin Agarwal, Heming Ge, Siamak Shakeri, and Rami Al-Rfou. 2021. Knowledge graph based synthetic corpus generation for knowledge-enhanced language model pre-training. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3554–3565, Online. Association for Computational Linguistics.

Qurat Ul Ain, Mohamed Amine Chatti, Komlan Gluck Charles Bakar, Shoeb Joarder, and Rawaa Alatrash. 2023. Automatic construction of educational knowledge graphs: A word embedding-based approach. *Information*, 14(10):526.

Olivier Bodenreider. 2004. The unified medical language system (umls): integrating biomedical terminology. *Nucleic acids research*, 32(Database issue), D267–D270.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. *arXiv e-prints*, arXiv:2005.14165.

Thiago Castro Ferreira, Claire Gardent, Nikolai Ilinykh, Chris van der Lee, Simon Mille, Diego Moussallem, and Anastasia Shimorina. 2020. The 2020 bilingual, bi-directional WebNLG+ shared task: Overview and evaluation results (WebNLG+ 2020). In *Proceedings of the 3rd International Workshop on Natural Language Generation from the Semantic Web (WebNLG+)*, pages 55–76, Dublin, Ireland (Virtual). Association for Computational Linguistics.

Penghe Chen, Yu Lu, Vincent W. Zheng, Xiyang Chen, and Boda Yang. 2018. Knowedu: A system to construct knowledge graph for education. *IEEE Access*, 6:31553–31563.

Wei-Lin Chiang, Zhuohan Li, Zi Lin, Ying Sheng, Zhanghao Wu, Hao Zhang, Lianmin Zheng, Siyuan Zhuang, Yonghao Zhuang, Joseph E. Gonzalez, Ion Stoica, and Eric P. Xing. 2023.

Janneth Chicaiza and Priscila Valdiviezo-Diaz. 2021. A comprehensive survey of knowledge graph-based recommender systems: Technologies, development, and contributions. *Information*, 12(6):232.

Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellat, Kevin Robinson, Dasha Valter, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. 2022. Scaling Instruction-Finetuned Language Models. *arXiv e-prints*, arXiv:2210.11416.

Fu-Rong Dang, Jin-Tao Tang, Kun-Yuan Pang, Ting Wang, Sha-Sha Li, and Xiao Li. 2021. Constructing an educational knowledge graph with concepts linked to wikipedia. *Journal of Computer Science and Technology*, 36:1200–1211.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training 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, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Claire Gardent, Anastasia Shimorina, Shashi Narayan, and Laura Perez-Beltrachini. 2017. Creating training corpora for NLG micro-planners. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 179–188, Vancouver, Canada. Association for Computational Linguistics.

Hamed Giglou, Jennifer D'Souza, and Sören Auer. 2023. LLMs4OL: Large Language Models for Ontology Learning. *arXiv e-prints*, arXiv:2307.16648.

Yu Gu, Robert Tinn, Hao Cheng, Michael Lucas, Naoto Usuyama, Xiaodong Liu, Tristan Naumann, Jianfeng Gao, and Hoifung Poon. 2021. Domain-specific language model pretraining for biomedical natural language processing. *ACM Trans. Comput. Healthcare*, 3(1).

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. *LoRA: Low-rank adaptation of large language models*. Accessed: 2023-05-21.

Forrest Iandola, Albert Shaw, Ravi Krishna, and Kurt Keutzer. 2020. SqueezeBERT: What can computer vision teach NLP about efficient neural networks? In *Proceedings of SustaiNLP: Workshop on Simple and Efficient Natural Language Processing*, pages 124–135, Online. Association for Computational Linguistics.

Mayank Kejriwal. 2022. Knowledge graphs: A practical review of the research landscape. *Information*, 13(4):161.

Muhammad Khoiruddin, Sri Kusumawardani, Indriana Hidayah, and Silmi Fauziati. 2023. A review of ontology development in the e-learning domain: Methods, roles, evaluation. *2023 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*.

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):167–195.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.

Linqing Li and Zhifeng Wang. 2023. Knowledge Graph Enhanced Intelligent Tutoring System Based on Exercise Representativeness and Informativeness. *arXiv e-prints*, arXiv:2307.15076.

S. Li, J. Tang, M.Y. Kan, D. Zhao, S. Li, and H. Zan. 2019. Duie: A large-scale chinese dataset for information extraction. *Natural Language Processing and Chinese Computing*, 11839.

Medicomp Systems. 2023. MEDCIN. Accessed: 2024-05-21.

Igor Melnyk, Pierre Dognin, and Payel Das. 2022. Knowledge Graph Generation From Text. *arXiv e-prints*, arXiv:2211.10511.

Nandana Mihindukulasooriya, Sanju Tiwari, Carlos F. Enguix, and Kusum Lata. 2023. Text2KGBench: A Benchmark for Ontology-Driven Knowledge Graph Generation from Text. *arXiv e-prints*, arXiv:2308.02357.

George A. Miller. 1995. Wordnet: a lexical database for english. *Commun. ACM*, 38(11):39–41.

National Cancer Institute, National Institutes of Health. 2022. NCI Thesaurus. Accessed: 2024-05-21.

Learning Technology Standards Committee of the IEEE Computer Society. 2020. Ieee standard for learning object metadata. *IEEE Std 1484.12.1-2020*, pages 1–50.

OpenAI. 2023. ChatGPT: Language Model. Accessed: 2024-05-21.

OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, Irwan Bello, Berdine, and et al. 2023. GPT-4 Technical Report. *arXiv e-prints*, arXiv:2303.08774.

Shirui Pan, Linhao Luo, Yufei Wang, Chen Chen, Jiapu Wang, and Xindong Wu. 2023. Unifying Large Language Models and Knowledge Graphs: A Roadmap. *arXiv e-prints*, arXiv:2306.08302.

Thomas Rebele, Fabian Suchanek, Johannes Hoffart, Joanna Biega, Erdal Kuzey, and Gerhard Weikum. 2016. Yago: A multilingual knowledge base from wikipedia, wordnet, and geonames. *International Semantic Web Conference*, pages 177–185.

Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3982–3992, Hong Kong, China. Association for Computational Linguistics.

S. Riedel, L. Yao, and A. McCallum. 2010. Modeling relations and their mentions without labeled text. *Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2010. Lecture Notes in Computer Science*, 6323.

SNOMED International. 2023. US Edition of SNOMED CT. Accessed: 2024-05-21.

K. Stancin, P. Poscic, and D. Jaksic. 2020. Ontologies in education – state of the art. *Education and Information Technologies*, 25:5301–5320.

Yi Sun, Hangping Qiu, Yu Zheng, Zhongwei Wang, and Chaoran Zhang. 2020. Sifrank: A new baseline for unsupervised keyphrase extraction based on pre-trained language model. *IEEE Access*, 8:10896–10906.

Rohan Taori, Ishaan Gulrajani, Tianyi Zhang, Yann Dubois, Xuechen Li, Carlos Guestrin, Percy Liang, and Tatsunori B. Hashimoto. 2023. Stanford Alpaca: An instruction-following LLaMA model. Accessed: 2023-05-21.

Kristina Toutanova, Danqi Chen, Patrick Pantel, Hoifung Poon, Pallavi Choudhury, and Michael Gamon. 2015. Representing text for joint embedding of text and knowledge bases. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1499–1509, Lisbon, Portugal. Association for Computational Linguistics.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. 2023. LLaMA: Open and Efficient Foundation Language Models. *arXiv e-prints*, arXiv:2302.13971.

Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Commun. ACM*, 57(10):78–85.

Xiaozhi Wang, Ziqi Wang, Xu Han, Wangyi Jiang, Rong Han, Zhiyuan Liu, Juanzi Li, Peng Li, Yankai Lin, and Jie Zhou. 2020. MAVEN: A Massive General Domain Event Detection Dataset. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1652–1671, Online. Association for Computational Linguistics.

BigScience Workshop, Teven Le Scao, Angela Fan, Christopher Akiki, Ellie Pavlick, Suzana Ilić, Daniel Hesslow, Roman Castagné, Alexandra Sasha Luccioni, François Yvon, Matthias Gallé, Jonathan Tow, Alexander M. Rush, Stella Biderman, Albert Webson, Pawan Sasanka Ammanamanchi, Thomas Wang, Benoît Sagot, Niklas Muennighoff, Albert Villanova del Moral, and et al. 2022. BLOOM: A 176B-Parameter Open-Access Multilingual Language Model. *arXiv e-prints*, arXiv:2211.05100.

Tongtong Wu, Linhao Luo, Yuan-Fang Li, Shirui Pan, Thuy-Trang Vu, and Gholamreza Haffari. 2024. Continual Learning for Large Language Models: A Survey. *arXiv e-prints*, arXiv:2402.01364.

Matthias Wölfel, Mehrnoush Barani Shirzad, Andreas Reich, and Katharina Anderer. 2024. 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(1).

Xu Yuan, Jiaxi Chen, Yingbo Wang, Anni Chen, Yiou Huang, Wenhong Zhao, and Shuo Yu. 2024. Semantic-enhanced knowledge graph completion. *Mathematics*, 12(3).

Yuqi Zhu, Xiaohan Wang, Jing Chen, Shuofei Qiao, Yixin Ou, Yunzhi Yao, Shumin Deng, Huajun Chen, and Ningyu Zhang. 2023. LLMs for Knowledge Graph Construction and Reasoning: Recent Capabilities and Future Opportunities. *arXiv e-prints*, arXiv:2305.13168.