# LLM+Graph: 2$^{nd}$ International Workshop on Data Management Opportunities in Bringing LLMs with Graph Data


Yixiang Fang
The Chinese University of Hong Kong, Shenzhen, China
fangyixiang@cuhk.edu.cn

Arijit Khan
Aalborg University
Aalborg, Denmark
arijitk@cs.aau.dk

Tianxing Wu
Southeast University
Nanjing, China
tianxingwu@seu.edu.cn

Da Yan
Indiana University Bloomington
Bloomington, USA
yanda@iu.edu



## ABSTRACT

Large Language Models (LLMs) have seen rapid development such as ChatGPT and LLaMA, and have received tremendous attention from both industry and academia. While they have shown remarkable success over text data, recent works have demonstrated their limitations in reasoning with structured data including graphs. Graph-structured data is ubiquitous in the real world ranging from social and biological networks to financial transactions, knowledge bases, and transportation systems – they permeate our daily lives. Therefore, understanding how to utilize graph data optimally with LLMs is a crucial research question. Recently, exploring the synergy between graphs and LLMs are attracting increasing interest in the data management and AI communities. On one hand, LLMs can be enhanced with graph computing techniques to provide answers with more contextualized facts, e.g., graph-based retrieval-augmented generation (graph RAG). On the other hand, downstream tasks, e.g., knowledge graph (KG) construction, graph data management and mining can also benefit by adopting LLMs, such as via LLM-graph DB (graph databases) and LLM-GNN (graph neural network) collaborations. It is, therefore, a timely opportunity to explore effective ways of interactions between LLMs and graph data. Our workshop "**LLM+Graph**" targets data management and data science researchers, aiming to inspect effective algorithms and systems to bringing LLMs, graph data management, and graph ML together in real applications.



**VLDB Workshop Reference Format:**
Yixiang Fang, Arijit Khan, Tianxing Wu, and Da Yan. LLM+Graph: 2$^{nd}$ International Workshop on Data Management Opportunities in Bringing LLMs with Graph Data. VLDB 2025 Workshop: LLM+Graph.


## 1 MOTIVATIONS, TOPICS, AND GOALS

The emergence of LLMs, e.g., ChatGPT, PaLM, and LLaMA, provides promising capabilities in artificial general intelligence (AGI), demonstrating excellent performance in various natural language processing (NLP) tasks in domains including customer support, healthcare, finance, law, education, and myriads of data science applications [2, 9]. General-purpose LLMs are pre-trained on massive text corpora in a self-supervised manner – they circumvent expensive task-specific supervision and the need for labeled data due to simpler prompt-based interactions – making them easy to operate and iterate over LLM pipelines for downstream applications [4]. While LLMs are proficient at learning probabilistic language patterns, they may not explicitly store consistent representations of knowledge, hence can output unreliable and incoherent responses, and often experience hallucinations by generating factually incorrect, inconsistent, or harmful contents.

In the meantime, graph data have been prevalent in various applications such as knowledge graphs, social media, e-commerce platforms, financial transactions, molecular, and biology, and they are often used to model complex relationships among entities. While LLMs' primary focus has been on text sequences, there is a growing interest in enhancing and exploiting the emergent abilities of LLMs to process graph data. Since LLMs offer parametric knowledge, while graphs store explicit and contextual information, they can complement each other in data science applications. Efforts have been made to unify LLMs and graphs by leveraging their respective strengths [1, 12, 18, 20]. For example, KGs assist in the pre-training and inference phases of LLMs, e.g., through retrieval augmented generation (KG-RAG) methods, to provide external knowledge for reducing hallucinations and improving accuracy and interpretability. LLMs, on the other hand, facilitate in various downstream data management and mining tasks over graphs. LLMs as graph models (a.k.a. graph foundation models) is a newly trending direction. Nevertheless, when it comes to graphs, LLMs are yet to demonstrate the same level of success as in NLP and computer vision. The key reason is that LLMs are not able to utilize graph structures effectively in graph reasoning tasks [6]. Hence, our workshop will provide an excellent avenue to investigate this emerging topic.

### 1.1 Graph Data for LLMs

Offering external knowledge through graphs is becoming prevalent for enhancing the accuracy and overall capabilities of LLMs.

• **Graph-enhanced Pre-training:** Adding graph data to training corpus improves pre-training data quality and context, thereby improving LLMs' accuracy. In the phase of pre-training, the KGs and text are aligned (via local subgraph extraction and entity linking) and interacted to jointly train the language models for complex question answering (QA) tasks [23]. To address knowledge forgetting during knowledge integration, InfuserKI [22] introduces the adaptive selection of new knowledge that is integrated with LLMs.

• **Graph-enhanced Fine-tuning:** KGs can fine-tune language models to update and expand their internal knowledge for domain-specific tasks. KG-Adapter [21] improves parameter-efficient fine-tuning of LLMs by introducing a knowledge adaptation layer to LLMs. GAIL [25] fine-tunes LLMs for learning lightweight KGQA models based on retrieved SPARQL-question pairs from KGs.



Proceedings of the VLDB Endowment. ISSN 2150-8097.

• **Graph-enhanced Inference:** The graph-based RAG techniques are effective in improving an LLM's accuracy by providing relevant contexts through relationships. For example, Microsoft has developed GraphRAG [5] which leverages structural graphs to enhance LLM performance by first building KGs from documents and then constructing community-level summaries. KG-RAG4SM [16] proposes external knowledge graph-based retrieval-augmented generation model for data curation tasks such as schema matching.
• **Graph-enhanced Refiners and Validators:** KGs can enhance LLMs in downstream tasks by serving as refiners and validators, providing structured knowledge to verify answers against factual relationships. ACT-Selection [19] filters and re-ranks candidate answers based on their types extracted from Wikidata. KGR [8] extracts and validates factual statements in model outputs, significantly boosting performance on factual QA benchmarks.

## 1.2 LLMs for Graphs

As a powerful tool, LLMs can be used to solve many graph data management and mining problems.
• **LLM-enhanced Graph Querying:** The language understanding capacity of LLMs makes them suitable for processing natural language questions (NLQs) over structured graphs [7, 15, 17]. LLMs assist in translating NLQs into GraphQL or Cypher queries to analyze and query graph data. For example, Neo4j (a top-ranking commercial graph DB) has integrated LLMs to automatically convert natural language questions into Cypher queries [17].
• **LLM-enhanced Graph Mining:** Due to reasoning and code generation capabilities of LLMs, they can be used to extract graph properties, e.g., for graph classification and generate code to solve complex graph mining problems, e.g., the shortest path computation [3, 10, 13, 24]. For instance, GraphWiz [3] demonstrated LLMs' ability to comprehend graph structures and perform graph analysis tasks such as cycle detection and subgraph matching.
• **LLM-enhanced Graph Learning:** Graph intelligence represents a crucial step towards achieving general artificial intelligence. Recently, LLMs have demonstrated remarkable zero-shot capabilities in processing text-attributed graphs. Following this trend, graph foundation models have emerged as a promising direction [13, 24]. Liu et al. [13] proposed a unified model capable of handling various graph ML tasks, such as node classification and link prediction, thereby eliminating the need for task-specific architectures.

## 1.3 Opportunities for Database Researchers

The combination of LLMs and graphs provides various exciting opportunities for data management and data science research.
• **Data and Input Modeling:** The graph structures need to be serialized as part of LLMs' input, either by verbalizing the graph structure in natural languages, or by encoding the sparse structure in dense vector forms. Some interesting open problems include: How to integrate graph structure with other multi-modal data, e.g., text, tables, and images, as input to LLMs? How to extract relevant subgraphs from KGs for specific downstream tasks? How to design and learn prompts with graph data for better generalization?
• **Data Cleaning, Integration, and Augmentation:** Data cleaning (e.g., error detection, repairing) and integration (e.g., entity and relation extraction, entity resolution, linking) are fundamental to data management. The unification of LLMs and domain-specific knowledge graphs provide new opportunities in this domain.
• **Vector Data Management:** Vector DBs support efficient top-$k$ retrievals in RAG for LLMs. With the emergence of graph RAG, there is a need for unifying graph DBs and vector DBs as external memory of LLMs [14]. How to design indexes, search algorithms, and systems for more complex and hybrid vector search, including graph traversal with vector retrieval? How graph DBs can be used as semantic caches of LLMs by indexing previous question-answer pairs into a graph or vector space, enabling semantic matching with new queries instead of more expensive LLM API calls?
• **Accuracy and Consistency:** Enhancing LLMs' accuracy and consistency, reducing hallucinations and the generation of harmful contents, fake news detection, fact checking, etc. with knowledge-grounded techniques are emerging research directions.
• **Explainability and Provenance:** KGs offer explanability to LLMs' responses by probing them and grounding their reasoning with external knowledge. It is critical to develop techniques that can associate LLM-generated contents with its provenance information.
• **Security and Privacy:** LLMs can inadvertently reveal confidential information in its responses, leading to unauthorized data access and security breaches. As the usage of graph data in LLMs expands, so does the concern for privacy and security. Ensuring the confidentiality of sensitive graph information, while still extracting knowledge for LLMs, poses an exciting challenge.
• **Benchmarking and Ground Truth:** In many emerging domains, the integration of LLMs and graphs have demonstrated incredible promises, so it is important to have ground-truth datasets and experimental benchmarks to facilitate future research in these domains.

## 1.4 Goals: Why Important? Why Now?

LLMs have become a powerful tool for interacting with data. LLM adoption is rapidly accelerating in the industry – prominent players include Open AI (ChatGPT), Google (PaLM), Meta (LLaMA), AI21Labs (Jurassic), Cohere, Anthropic (Claude), Microsoft (TuringNLG, Orca), Huawei (Pangu), Naver (HyperCLOVA), Tencent (Hunyuan), Yandex (YaLM), Amazon (Titan, Olympus), Bytedance (Doubao), etc. Most teams using LLMs are investing in prompt engineering, vector databases, and LLMs' monitoring (e.g., Responsible AI). The global LLM market size in terms of revenue is projected to reach 259,886.45 million US dollars by 2029 from 1,302.93 million in 2023, with a compound annual growth rate (CAGR) 141.72% during 2023-2029[1]. Given that the technology is so new, it is an exciting time to work at the forefront of LLM research and development. This technology also created several opportunities for applications in general data management and AI. However, the synergy between LLMs and graph data has received insufficient attention by the DB community. This workshop's objective is to draw attention to this emerging topic, which has the potential to not only deepen LLMs' impact in real-world graph data and applications, but also enhance the performance of LLMs using graph data. Therefore, our workshop is timely and relevant.

The LLM+KG@VLDB 2024 workshop [11] drew over 150 participants—the largest at VLDB 2024. As the second installment

[1] https://www.giiresearch.com/report/qyr1384359-global-large-language-modelllm-market-research.html

in our ongoing LLM+Graph series, this edition expands beyond knowledge-graph management to cover data management and mining across graph computing, and is expected to attract an even larger audience.

## 2 WORKSHOP PROGRAM

This workshop features seven accepted papers including two vision papers, three keynote talks, four industry talks, and and a panel discussion exploring cutting-edge research and emerging directions for unresolved challenges at the intersection of large-language models and graph technologies.

**Session 1**
Opening remark by PC co-chairs.
Keynote 1. *Exploring the Duality Between Large Language Models and Database Systems* – M. Tamer Özsu (University of Waterloo).
Industry Talk 1. *Applications and Challenges of GraphRAG and Graph Foundation Models at ByteDance* – Cheng Chen (ByteDance).
Paper Presentation 1: *LLM-assisted Construction of the United States Legislative Graph* - Francesco Cambria and Andrea Colombo.
Paper Presentation 2: *Scalable Graph-based Retrieval-Augmented Generation via Locality-Sensitive Hashing* - Fangyuan Zhang, Zhengjun Huang, Yingli Zhou, Qingtian Guo, Wensheng Luo, and Xiaofang Zhou.

**Session 2**
Keynote 2. *Towards Graph Foundation Models with Riemannian Geometry* – Philip S. Yu (University of Illinons, Chicago).
Industry Talk 2. *Retrieval and Reasoning with LLMs on Neo4j: Progress and Challenges* – Brian Shi (Neo4J).
Paper Presentation 3: *LLM-Hype: A Targeted Evaluation Framework for Hypernym-Hyponym Identification in Large Language Models* - Qiu Ji, Pengfei Zhu, Haolei Zhu, Yang Sheng, Guilin Qi, Lianlong Wu, Kang Xu, and Yuan Meng.
Paper Presentation 4: *Graph-Enhanced Large Language Models for Spatial Search [Vision]* - Nicole Schneider, Kent O'Sullivan, and Hanan Samet.

**Session 3**
Keynote 3. *Reasoning over Property Graphs: Leveraging Large Language Models for Automated Data Consistency* – Angela Bonifati (Lyon 1 University).
Industry Talk 3. *Chat2Graph: A Graph Native Agentic System* – Heng Lin (AntGroup).
Paper Presentation 5: *xpSHACL: Explainable SHACL Validation using Retrieval-Augmented Generation and Large Language Models* - Gustavo Publio and Jose Emilio Labra Gayo.
Paper Presentation 6: *Automatic Prompt Optimization for Knowledge Graph Construction: Insights from an Empirical Study* - Nandana Mihindukulasooriya, Niharika DSouza, Faisal Chowdhury, and Horst Samulowitz.
Paper Presentation 7: *Towards the Next Generation of Agent Systems: From RAG to Agentic AI [Vision]* - Yingli Zhou and Shu Wang.

**Session 4**
Panel: *LLMs vs. Graphs: Supercharge or Supersede Graph Data Management, Mining, and Learning?* - Panelists: M. Tamer Özsu (University of Waterloo), Philip S. Yu (University of Illinons, Chicago), Angela Bonifati (Lyon 1 University), Chen Cheng (ByteDance), Heng Lin (AntGroup).

## 3 PROGRAM COMMITTEE

Cuneyt Akcora (University of Central Florida)
Naheed Anjum Arafat (Howard University)
Ariful Azad (Texas A&M University)
Sheng Bi (Southeast University)
Lei Cao (University of Arizona)
Joydeep Chandra (Indian Institute of Technology Patna)
Lijun Chang (The University of Sydney)
Yubo Chen (Institute of Automation, Chinese Academy of Sciences)
Rainer Gemulla (Universität Mannheim)
Guimu Guo (Rowan University)
Xin Huang (Hong Kong Baptist University)
Jiahui Jin (Southeast University)
Longbin Lai (Alibaba Corporation)
Wolfgang Lehner (TU Dresden)
Cheng Long (Nanyang Technological University)
Chenhao Ma (The Chinese University of Hong Kong, Shenzhen)
Chuangtao Ma (Aalborg University)
Yudong Niu (Singapore Management University)
Peng Peng (Hunan University)
Lu Qin (University of Technology Sydney)
A. Erdem Sariyuce (University at Buffalo)
Juan Sequeda (data.world)
Jieming Shi (The Hong Kong Polytechnic University)
Yihan Sun (University of California, Riverside)
Kai Wang (Shanghai Jiao Tong University)
Meng Wang (Tongji University)
Dong Wen (University of New South Wales)
Yinglong Xia (Meta)
Shiyu Yang (Guangzhou University)
Long Yuan (Wuhan University of Technology)
Yuanyuan Zeng (Chinese University of Hong Kong, Shenzhen)
Yanfeng Zhang (Northeastern University)
Fan Zhang (Guangzhou University)
Yingli Zhou (The Chinese University of Hong Kong, Shenzhen)
Rong Zhu (Alibaba Group)

## 4 WORKSHOP CO-CHAIRS

**Yixiang Fang** is an Associate Professor at the School of Data Science in the Chinese University of Hong Kong, Shenzhen, China. His research interests focus on data management, data mining, and artificial intelligence over big graph data. He has published extensively in the areas of database and data mining, including One of the Best Papers in SIGMOD 2020, and most of them are published in top-tier conferences and journals. He was awarded the 2021 ACM SIGMOD Research Highlight Award. He is an editorial board member of the journal of Information & Processing Management (IPM). He has also served as program committee members for several top conferences (e.g., PVLDB, ICDE, and KDD) and invited reviewers for top journals (e.g., TKDE and VLDBJ) in the areas of database and data mining. He is a member of ACM, IEEE, and CCF. More information at https://fangyixiang.github.io/.

**Arijit Khan** is an IEEE senior member, an ACM distinguished speaker, and an associate professor in the Department of Computer Science, Aalborg University, Denmark. Arijit is the recipient of the

IBM Ph.D. Fellowship (2012-13), a VLDB Distinguished Reviewer award (2022), and a SIGMOD Distinguished PC award (2024). He published over 100 papers in premier data management and mining venues, e.g., SIGMOD, VLDB, KDD, ICLR, TKDE, ICDE, WWW, SDM, EDBT, CIKM, WSDM and TKDD. Arijit co-presented tutorials at VLDB, ICDE, CIKM, and DSAA; and served/ is serving as an associate editor of TKDE and TKDD. Arijit served as the co-chair of LLM+KG workshop (co-located w/ VLDB 2024), LLM+Vector Data workshop (co-located w/ ICDE 2025), KG+Responsible AI workshop (co-located w/ ESWC 2025, CIKM 2024), and Big-O(Q) workshop (co-located w/ VLDB 2015), and wrote a book on uncertain graphs in the Morgan & Claypool's Synthesis Lectures on Data Management. More information at https://homes.cs.aau.dk/~Arijit/index.html.

**Tianxing Wu** is an Associate Professor working at the School of Computer Science and Engineering of Southeast University, China. He is one of the main contributors to build Chinese large-scale encyclopedic knowledge graph: Zhishi.me and schema knowledge graph: Linked Open Schema. He has achieved SIGIR 2024 Best Short Paper Nominee, CCKS 2022 Best Paper Award, WISA 2024 Best Student Paper Award. His research interests include knowledge graph, knowledge representation and reasoning, and data mining. He has published over 60 papers in top-tier venues, e.g., SIGMOD, ICDE, SIGIR, ACL, AAAI, IJCAI, TKDE, and TKDD. He is an editorial board member of International Journal on Semantic Web and Information Systems, and Data Intelligence. He has served as the (senior) program committee member of AAAI, IJCAI, ACL, KDD, The WebConf, ICML, and ICLR. More information at https://tianxing-wu.github.io/.

**Da Yan** is an Associate Professor in the Department of Computer Science of the Luddy School of Informatics, Computing, and Engineering (SICE) at Indiana University Bloomington. He is a DOE Early Career Research Program (ECRP) awardee in 2023, the sole winner of the Hong Kong 2015 Young Scientist Award in Physical/Mathematical Science, and senior members of ACM and IEEE. His research interests include parallel and distributed systems for big data analytics, data mining, and machine learning. He frequently publishes in top databases and AI conferences and journals, and he serves extensively in the major top databases and AI conferences and journals as reviewers. Dr. Yan leads the BIOKDD workshop with SIGKDD (now in its 23$^{rd}$ year), and co-organized a Dagstuhl seminar, and a few top conferences. He also served as guest editors of journals such as IEEE/ACM TCBB, BMC Bioinformatics, and IEEE CG&A. More information at https://homes.luddy.indiana.edu/yanda/home.html.

## REFERENCES


[1] Garima Agrawal, Tharindu Kumarage, Zeyad Alghamdi, and Huan Liu. 2024. Can knowledge graphs reduce hallucinations in LLMs?: A survey. In *North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers) (NAACL-HLT)*. 3947–3960.

[2] Sihem Amer-Yahia, Angela Bonifati, Lei Chen, Guoliang Li, Kyuseok Shim, Jianliang Xu, and Xiaochun Yang. 2023. From large language models to databases and back: A discussion on research and education. *SIGMOD Rec.* 52, 3 (2023), 49–56.

[3] Nuo Chen, Yuhan Li, Jianheng Tang, and Jia Li. 2024. Graphwiz: An instruction-following language model for graph computational problems. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*. 353–364.

[4] Xi Chen, Wei Hu, Arijit Khan, Shreya Shankar, Haofen Wang, Jianguo Wang, and Tianxing Wu. 2024. Large language models, knowledge graphs, and vector databases: Synergy and opportunities for data management. ACM SIGMOD Blog. https://wp.sigmod.org/?p=3813

[5] Darren Edge, Ha Trinh, Newman Cheng, Joshua Bradley, Alex Chao, Apurva Mody, Steven Truitt, and Jonathan Larson. 2024. From local to global: A graph RAG approach to query-focused summarization. *arXiv:2404.16130* (2024).

[6] Bahare Fatemi, Jonathan Halcrow, and Bryan Perozzi. 2024. Talk like a graph: Encoding graphs for large language models. In *International Conference on Learning Representations (ICLR)*.

[7] Balaji Ganesan, Sambit Ghosh, Nitin Gupta, Manish Kesarwani, Sameep Mehta, and Renuka Sindhgatta. 2024. LLM-powered GraphQL generator for data retrieval. In *International Joint Conference on Artificial Intelligence (IJCAI)*. 8657–8660.

[8] Xinyan Guan, Yanjiang Liu, Hongyu Lin, Yaojie Lu, Ben He, Xianpei Han, and Le Sun. 2024. Mitigating large language model hallucinations via autonomous knowledge graph-based retrofitting. In *AAAI Conference on Artificial Intelligence*. 18126–18134.

[9] Alon Y. Halevy, Yejin Choi, Avrilia Floratou, Michael J. Franklin, Natasha F. Noy, and Haixun Wang. 2023. Will LLMs reshape, supercharge, or kill data science? *Proc. VLDB Endow.* 16, 12 (2023), 4114–4115.

[10] Xuanwen Huang, Kaiqiao Han, Yang Yang, Dezheng Bao, Quanjin Tao, Ziwei Chai, and Qi Zhu. 2024. Can GNN be good adapter for LLMs?. In *ACM Web Conference (WWW)*.

[11] Arijit Khan, Tianxing Wu, and Xi Chen. 2024. LLM+KG: Data management opportunities in unifying large language models + knowledge graphs. In *Proceedings of Workshops at the International Conference on Very Large Data Bases (VLDB)*.

[12] Yuhan Li, Zhixun Li, Peisong Wang, Jia Li, Xiangguo Sun, Hong Cheng, and Jeffrey Xu Yu. 2024. A survey of graph meets large language model: Progress and future directions. In *International Joint Conference on Artificial Intelligence (IJCAI)*. 8123–8131.

[13] Hao Liu, Jiarui Feng, Lecheng Kong, Ningyue Liang, Dacheng Tao, Yixin Chen, and Muhan Zhang. 2024. One for all: Towards training one graph model for all classification tasks. In *International Conference on Learning Representations (ICLR)*.

[14] Shige Liu, Zhifang Zeng, Li Chen, Adil Ainihaer, Arun Ramasami, Songting Chen, Yu Xu, Mingxi Wu, and Jianguo Wang. 2025. TigerVector: Supporting Vector Search in Graph Databases for Advanced RAGs. In *Companion of the International Conference on Management of Data (SIGMOD)*. 553–565.

[15] Yang Liu, Xin Wang, Jiake Ge, Hui Wang, Dawei Xu, and Yongzhe Jia. 2024. Text to Graph Query Using Filter Condition Attributes. In *Proceedings of Workshops at the International Conference on Very Large Data Bases (VLDB)*.

[16] Chuangtao Ma, Sriom Chakrabarti, Arijit Khan, and Bálint Molnár. 2025. Knowledge graph-based retrieval-augmented generation for schema matching. *arXiv:2501.08686* (2025).

[17] neo4j. 2024. neo4j. https://neo4j.com/labs/neodash/2.4/user-guide/extensions/natural-language-queries/.

[18] Jeff Z. Pan, Simon Razniewski, Jan-Christoph Kalo, Sneha Singhania, Jiaoyan Chen, Stefan Dietze, Hajira Jabeen, Janna Omeliyanenko, Wen Zhang, Matteo Lissandrini, Russa Biswas, Gerard de Melo, Angela Bonifati, Edlira Vakaj, Mauro Dragoni, and Damien Graux. 2023. Large language models and knowledge graphs: Opportunities and challenges. *Transactions on Graph Data and Knowledge (TGDK)* 1, 1 (2023), 2:1–2:38.

[19] Mikhail Salnikov, Maria Lysyuk, Pavel Braslavski, Anton Razzhigaev, Valentin Malykh, and Alexander Panchenko. 2023. Answer candidate type selection: Text-to-text language model for closed book question answering meets knowledge graphs. In *Conference on Natural Language Processing (KONVENS)*. 155–164.

[20] Wenbo Shang and Xin Huang. 2024. A survey of large language models on generative graph analytics: Query, learning, and applications. *arXiv:2404.14809* (2024).

[21] Shiyu Tian, Yangyang Luo, Tianze Xu, Caixia Yuan, Huixing Jiang, Chen Wei, and Xiaojie Wang. 2024. KG-Adapter: Enabling knowledge graph integration in large language models through parameter-efficient fine-tuning. In *Findings of the Association for Computational Linguistics (ACL)*. 3813–3828.

[22] Fali Wang, Runxue Bao, Suhang Wang, Wenchao Yu, Yanchi Liu, Wei Cheng, and Haifeng Chen. 2024. InfuserKI: Enhancing large language models with knowledge graphs via infuser-guided knowledge integration. In *Proceedings of Workshops at the International Conference on Very Large Data Bases (VLDB)*.

[23] Michihiro Yasunaga, Antoine Bosselut, Hongyu Ren, Xikun Zhang, Christopher D. Manning, Percy Liang, and Jure Leskovec. 2022. Deep Bidirectional Language-Knowledge Graph Pretraining. In *Advances in Neural Information Processing Systems (NeurIPS)*.

[24] Mengmei Zhang, Mingwei Sun, Peng Wang, Shen Fan, Yanhu Mo, Xiaoxiao Xu, Hong Liu, Cheng Yang, and Chuan Shi. 2024. GraphTranslator: Aligning graph model to large language model for open-ended tasks. In *ACM Web Conference (WWW)*.

[25] Zhiqiang Zhang, Liqiang Wen, and Wen Zhao. 2024. A GAIL fine-tuned LLM enhanced framework for low-resource knowledge graph question answering. In *ACM International Conference on Information and Knowledge Management (CIKM)*. 3300–3309.