# *Construction and Application of a Multi-modal Knowledge Graph Integrated with Large Language Models in the Field of Manufacturing Processes*

Xiaogui Tian
School of Mechanical Engineering
Northwestern Polytechnical University
Xi'an, China
2022100183@mail.nwpu.edu.cn

Jianxin Xu
School of Mechanical Engineering
Northwestern Polytechnical University
Xi'an, China
xujx@nwpu.edu.cn

Shuqin Wang
School of Management
Northwestern Polytechnical University
Xi'an, China
Wangsq2022@mail.nwpu.edu.cn

Yongsheng Zhou
School of Mechanical Engineering
Northwestern Polytechnical University
Xi'an, China
zhouyongsheng@mail.nwpu.edu.cn

***Abstract*—The manufacturing process of a product serves as a bridge connecting product design to production, and it is the cornerstone ensuring product quality in the manufacturing sector. The efficient application of knowledge in the field of manufacturing processes is a crucial means of enhancing manufacturing processes. Nevertheless, currently, knowledge in this field often exists in fragmented and multimodal forms within texts or databases, making it difficult to effectively mine and uniformly represent. The poor interconnectivity among knowledge pieces hinders the provision of high-quality knowledge support for the intelligent transformation of the manufacturing process sector. The advancements in knowledge graphs (KGs) and large language models (LLMs) within vertical domains have introduced novel research avenues for the construction and application of knowledge in the manufacturing process sector, laying a solid foundation for the realization of intelligent manufacturing. Therefore, this study addresses the problems of multimodality, irregularity, difficult representation, and lack of effective application of manufacturing-process domain knowledge. It combines the development of LLMs and their collaborative paradigm with multi-modal knowledge graphs (MMKGs), explores the construction of MMKGs in manufacturing process domains augmented by LLMs to realize the unified expression of process knowledge, and utilizes the excellent processing capability of LLMs to explore the LLM-enhanced multi-modal process knowledge. Additionally, we explore the effective application of multi-modal process knowledge enhanced by the LLM. This will provide inspiration and practical application references for related enterprises to further conduct research on the collaboration between LLM and KG in the manufacturing field.***

***Keywords—Multi-Modal Knowledge Graph, Large Language Model, Knowledge Construction, Knowledge Application***

## I. INTRODUCTION

The quality of manufacturing processes serves as a pivotal factor influencing product performance, and the efficient application of manufacturing-process knowledge is a crucial means of enhancing the quality of product manufacturing process design. Currently, knowledge in the field of manufacturing processes is stored in a fragmented and multimodal format within texts or databases, characterized by inconsistent knowledge representation and poor inter-knowledge connectivity, thereby posing challenges in meeting the demands of the transformation and upgrading of the manufacturing industry [1] Consequently, unifying the representation of knowledge and leveraging its connectivity to achieve efficient application has emerged as one of the core bottlenecks in realizing the goal of intelligent manufacturing.

As a descriptive framework for representing structured KGs are capable of expressing descriptive and instance data in a concise manner. They have been widely utilized in major fields such as electric power, finance, healthcare, and aviation [2]-[5]. With their semantic interconnection, structured, and instantiated features, KGs can deeply mine and effectively represent the knowledge implicit in manufacturing-process data, thereby laying the foundation for efficient knowledge application. Currently, the construction of KGs in the field of manufacturing processes primarily focuses on textual and structured data, neglecting research efforts aimed at constructing KGs from multi-modal data such as three-dimensional models, images, videos, audios, etc. This has led to an overall incomplete construction of knowledge in the manufacturing-process domain. Therefore, further exploration into the construction of knowledge graphs based on multi-modal manufacturing-process data is of significant importance for enhancing the quality of knowledge in this domain.

LLMs represented by BERT and GPT have achieved remarkable successes in various natural language processing tasks [6], laying a solid foundation for the processing of multi-modal data and broadening the application scope of such models in general domains. Given their powerful capabilities in data processing, comprehension, learning, and expression, the introduction of LLMs into the manufacturing-process domain offers the potential for handling and applying multi-modal data.

The current construction of knowledge in the manufacturing -process domain has overlooked the representation of multi-modal data features and contextual information, resulting in incomplete and low quality knowledge bases that hinder designers from effectively applying high - quality knowledge in the process of manufacturing-process design. Consequently, this study, building upon KG construction techniques, integrates the unique characteristics of multi-modal knowledge in the manufacturing-process domain with the robust semantic understanding, learning, and expression capabilities

demonstrated by LLMs in information processing. It explores the process of enhancing MMKG construction through LLMs and investigates efficient avenues for leveraging this enhanced knowledge in the manufacturing-process domain. Ultimately, this research aims to provide relevant enterprises with references and insights into the practical application of LLMs and MMKGs.

In summary, the main contributions of this paper are as follows:

(1) This paper explores the collaborative paradigm of MMKG and LLM, laying the foundation for the enhanced application of LLM in the professional domain.

(2) This paper explores the construction process of MMKG enhanced by LLM, laying the foundation for the automated construction of MMKG.

(3) This paper explores the mutual- enhancement of MMKG and LLM for enhanced application in the field of manufacturing process, laying the foundation for the industrial large-model enhancement application in manufacturing enterprises.

## II. Architecture of the Relationship Between LLMs and MMKGs

### A. LLM Development

LLMs are pre-trained language models with massive parameters constructed using the Transformer architecture. By virtue of their rich knowledge reserves and powerful reasoning capabilities [7], they have demonstrated impressive performance in natural language understanding and complex interactive tasks. Since the concept of OpenAI's language models was proposed, models such as GPT - 3 and GPT - 4 [8], trained on a larger scale of data, have been successively introduced. This has enabled AI technologies to better support various application scenarios and has injected new impetus into the development of AI. The relevant model is shown in Figure 1.

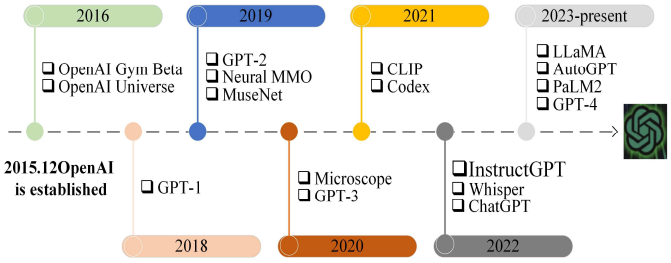


Fig. 1. Model of OpenAI related products

As a knowledge base with large-scale parameters, LLMs are capable of solving most complex problems and providing reliable assistance for the construction and effective application of knowledge in the manufacturing-process field. By utilizing application methods such as contextual learning, chain of thought, and external resource injection, LLMs have made significant progress in information processing and complex problem analysis, demonstrating greater comprehension and flexibility in handling multi-modal data processing.

### B. LLMs and MMKG collaboration paradigm

A MMKG is a graph - structured knowledge repository that encompasses structured attributes numerical features, model - related images, videos, and audio. Its representation mainly falls into two categories: one in which multi-modal information serves as independent entities, and the other in which multi-modal information constitutes attributes of entities. LLMs exhibit a high degree of overlap with KGs in terms of their information processing, semantic understanding, and generation capabilities, demonstrating a symbiotic relationship and the potential for integrated usage. Consequently, exploring the dependency between LLMs and MMKGs plays a pivotal role in facilitating the efficient application of manufacturing process knowledge. Specifically, based on the characteristics of LLMs and MMKGs, their collaborative paradigms can be classified into several categories, as shown in Figure 2.

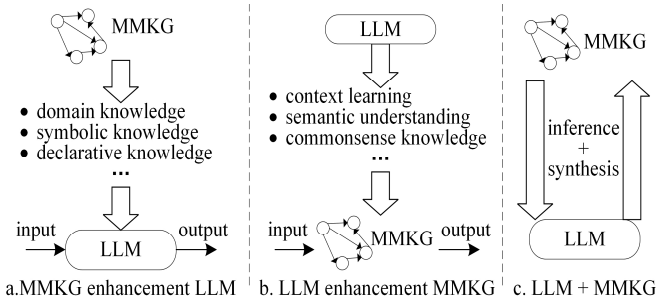


Fig. 2. Collaborative modeling of LLMs and KGs

①MMKGs enhance LLM training. MMKGs can effectively represent fragmented domain knowledge. Injecting domain knowledge into the training process of LLMs can enhance their comprehension of knowledge to a certain extent; ②LLMs enhance the construction of MMKGs. LLMs exhibit strong generalization abilities in natural language processing tasks and pattern recognition. Incorporating LLMs into the KG construction stage can fully leverage their capabilities in small sample learning, information extraction, and reasoning to assist in the creation, fusion, verification, and characterization of graph knowledge; ③Collaborative application of LLMs and MMKGs. Different methods and ideas are employed to transform and integrate the two. This can enhance the overall knowledge's characterization abilities, generalization capabilities, and reasoning powers.

Among them, the MMKG and the LLM fusion architecture is linked by the architecture of LangChain. The application principles of the two are as follows:

Firstly, in the application process, the big model parses the backbone of the user retrieval input questions, transforms the parsing results into corresponding Cypher statements to obtain the key entity information in the multimodal knowledge graph; Secondly, after parsing by LangChain, the retrieval results will be fed back to the LLM to collate the outputs.

## III. A Construction Method for LLMs to Enhance MMKGs

The construction of domain specific KGs primarily focuses on knowledge extraction from multi - source heterogeneous data

sources. The construction process, as shown in Figure 3, mainly involves knowledge identification and extraction from both structured and unstructured data.

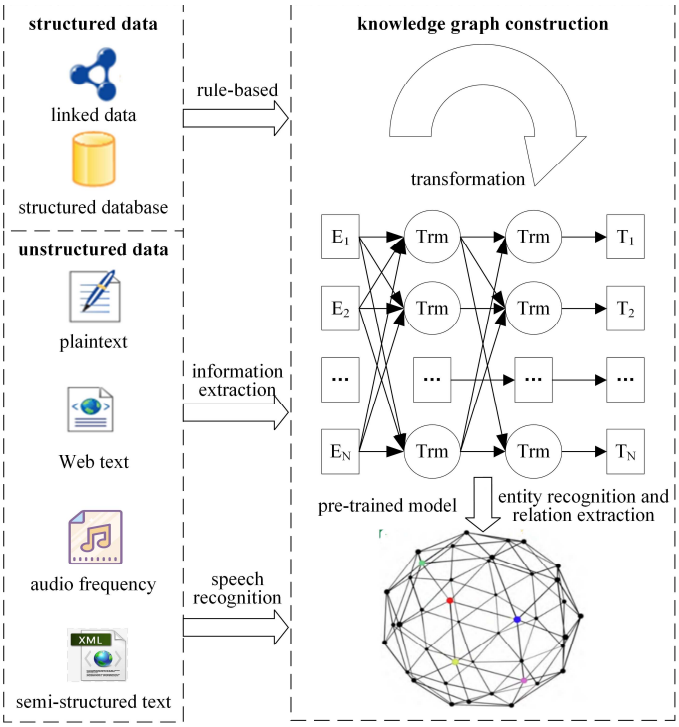


Fig. 3. KG construction process

The construction of MMKG mainly contains two parts: multi-modal entity recognition and multi-modal relationship extraction. In multi-modal entity recognition, YU [9] proposed a multi-modal interaction module and a text entity detection module, and designed a unified multi-modal transformer network to realize entity extraction; Mohamed [10] realized entity recognition in medical documents based on rule-based and hybrid approaches. In multi-modal relationship extraction, HVPNNeT [11] considered visual representations as visual prefixes that can be inserted to guide error-insensitive predictive decision-making for textual representations, achieving more efficient and powerful multi-modal entity and relationship extraction; MoRe [12] introduced text retrieval and image retrieval to multi-modal entity relationship extraction. It achieved the extraction of information with target relevance by introducing external knowledge, which improved the effect of extraction.

In addition, in terms of combining LLMs to construct MMKGs, AGRAWAL [13] used the InstructGPT LLM to learn from zero-sample and few-sample data, and combined its learning ability to realize entity relationship extraction in medical notes; YAN [14] used a pre-trained BERT model to realize the relationship extraction and task reasoning, and to identify the entity-relationships in the graph.

The aforementioned research has achieved corresponding outcomes in the construction of MMKGs, requiring minimal annotation of datasets and demonstrating strong adaptability to diverse domain data. However, due to the lack of domain knowledge guidance in pre-trained models, the direct application of LLMs in the manufacturing process domain for the construction of MMKGs yields suboptimal results. Consequently, based on the previous research, this paper adopts a top-down approach, guided by domain experts, to construct the ontology layer of the MMKG. We design prompt information to constrain the LLM's optimization and processing of domain data. Through knowledge extraction, we instantiate and populate the data layer of the KG, thereby realizing the construction of a LLM-enhanced multi-modal domain specific KG. The construction process is shown in Figure 4.

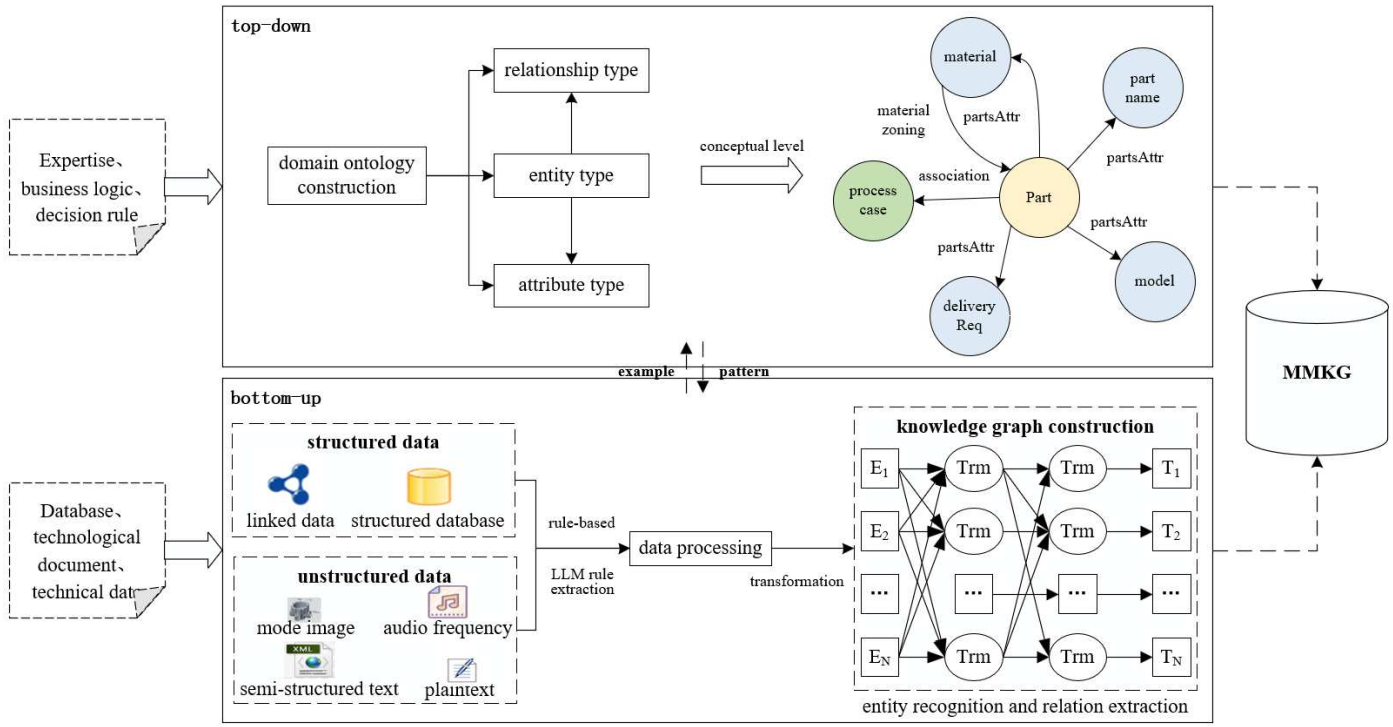

Fig. 4. MMKG construction process with LLM enhancement

Firstly, design the ontology layer for a MMKG. The ontology layer is a crucial element in the construction of a MMKG. Given that the manufacturing process field data is fragmented and multi-modal, and the process has high requirements for knowledge level, the ontology should be constructed according to its role in the manufacturing process, under the guidance of experts. Secondly, collect manufacturing process field data. The data sources include relevant engineering databases in the process domain, process documents, technical information, etc. Then, perform pre - processing work, such as data cleaning, on these data. Moreover, the design of cue information guides the LLM to extract rule information from the manufacturing process data. Finally, augment deep learning algorithms with the rule information extracted by the LLM to identify relevant entities and relationships, thereby realizing the construction of the data layer of the MMKG

## IV. Exploring the Application of Fusing LLMs and MMKGs in Manufacturing Process Fields

Based on the MMKG enhanced by the LLM, this section will further explore the prospect of integrating the LLM and MMKG in the manufacturing process field. This exploration will provide references and inspirations for the practical application of the LLM and MMKG in the intelligent manufacturing field.

### *A. Fusing LLMs with MMKGs for information retrieval*

Information retrieval is a prevalent application form of KGs. In the context of information retrieval based on LLMs and MMKGs, LLMs facilitate the semantic analysis and structural transformation of user queries, while the structured knowledge embedded within MMKGs is leveraged through information retrieval methods tailored for KGs. This approach enables the matching of user needs with appropriate information sets, providing comprehensive and accurate results. The application of information retrieval based on LLMs and MMKGs is shown in Figure 5. Compared to traditional information retrieval methods that directly utilize KGs, this approach better comprehends user demands, yielding more comprehensive and precise knowledge retrieval outcomes.

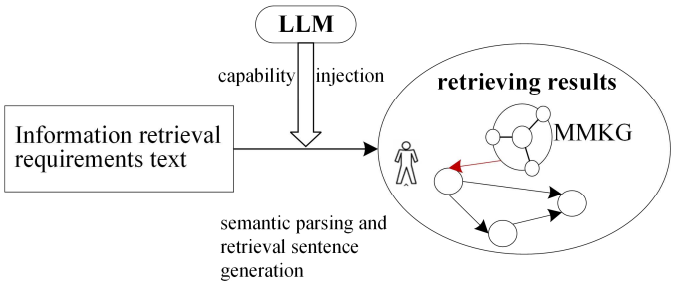


Fig. 5. Information retrieval

Existing LLMs and MMKGs related theories can be applied to knowledge matching and recommendation push technology in the field of intelligent manufacturing. Semantic parsing and understanding are performed for users' questions. By combining with the MMKG, the system can automatically answer the natural language questions asked by users.

### *B. Fusing LLMs with MMKGs for process knowledge question and answer*

Manufacturing processes are mostly based on human participation in designing and solving various complex problems, and it is difficult to effectively utilize fragmented process knowledge due to users' varying design experience and skill levels. Knowledge question and answer (Q&A) is a way to effectively utilize knowledge [15], which helps users to complete various complex tasks by understanding the user's intention and providing Q&A knowledge.

The traditional knowledge Q&A process using keyword matching is no longer applicable to the refined requirements for knowledge in the manufacturing process domain because it cannot accurately recognize user intent. Therefore, the process knowledge Q&A that integrates the LLM and MMKG knowledge base into the Q&A process is the mainstream of current research. Its specialized knowledge Q&A method is shown in Figure 6. The LLM is utilized to accurately identify the user's intention, match the accurate process knowledge from the expertise domain knowledge base, and answer a variety of complex questions in a form that is easily understood by the user.

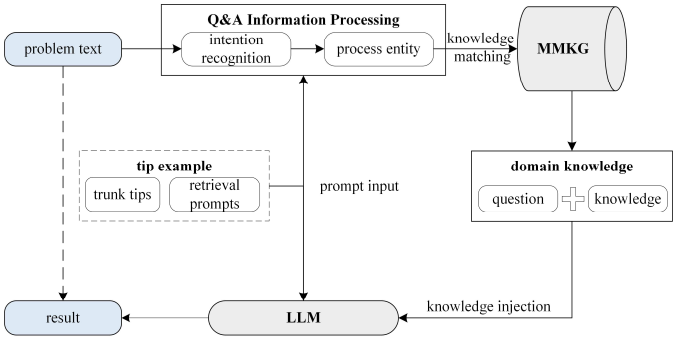


Fig. 6. Manufacturing process knowledge Q&A process

Existing theories related to LLMs and MMKGs can be applied to knowledge matching and knowledge Q&A in the field of intelligent manufacturing. In response to the natural language questions raised by the user, an LLM is used to semantically parse and understand the interrogative sentences and generate MMKG retrieval statements, which are matched with specialized knowledge from the constructed specialized-domain knowledge base and then fed back to the user.

### *C. Process knowledge reasoning by fusing LLMs with MMKGs*

Knowledge reasoning serves as a crucial step in applying design process knowledge within the manufacturing technology domain. Traditionally, people predominantly achieved knowledge reasoning through rule based methods. However, given the vastness and complexity of knowledge in the manufacturing technology domain, current approach of employing rule based methods is no longer suitable for process knowledge reasoning in the context of intelligent manufacturing environments. Therefore, to address this issue, this study adopts a fusion method of LLMs and MMKGs to realize process

knowledge reasoning under the paradigm of intelligent manufacturing. The reasoning process is shown in Figure 7.

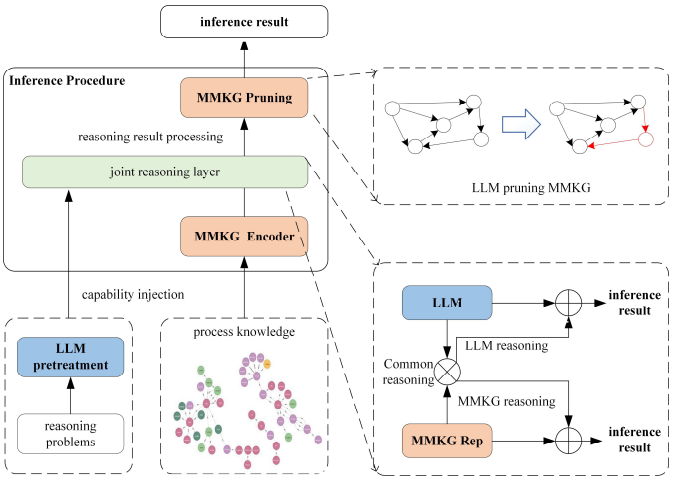


Fig. 7. Manufacturing process knowledge reasoning

First, we parse the reasoning problem through the LLM. Then, we combine the reasoning rules related to the process-domain knowledge in the MMKG. Next, we map the key information in the problem to the entities and relationships in the MMKG and construct the MMKG reasoning path. After that, we utilize the defined knowledge reasoning path, execute pattern matching and path querying in the MMKG. Finally, we prune the reasoning path through the LLM. In this way, we can realize accurate knowledge reasoning to solve problems in manufacturing-process design.

Compared with the traditional knowledge reasoning process, this method is no longer based on individual expert experience for formulating rules. Instead, this method diversifies the forms of knowledge reasoning, providing powerful support for the application of process knowledge.

Existing theories related to LLMs and MMKGs can be applied to knowledge mining, knowledge recommendation, process optimization and decision-making in the field of intelligent manufacturing. First, identify the subject entities in the input problem through the LLM. Then, use the inference-mapping ability of the MMKG to retrieve the relevant paths and get the inference answer. Finally, reason again in combination with the LLM. In this way, we can solve complex problems.

## V. CONCLUSION

Knowledge serves as a pivotal asset that significantly impacts the development of the manufacturing industry. A unified representation of domain knowledge provides crucial impetus for driving the transformation and upgrading of the manufacturing sector. The development of LLMs and KGs has boosted the manufacturing process domain, laying the foundation for realizing the unified expression and effective application of knowledge. This study summarizes the paradigm of collaboration between LLMs and KGs, analyzes the utilization of LLMs to enhance the construction of MMKGs, and collaboratively explores the effective application of LLMs and MMKGs to knowledge in the manufacturing process field. It provides inspiration for further research on LLMs and MMKGs in the field of intelligent manufacturing, thereby offering strong support for the realization of manufacturing modernization and high-quality development.

In the future, LLMs will develop rapidly along the direction of multimodality and cross - domain, and this study needs to further optimize the LLMs iteratively. This includes fine - tuning the LLM using fine - tuning techniques, enhancing the LLM using hints from the chain-of-thinking method, and enhancing the reasoning ability of the LLM through knowledge injection, thus mitigating the model's hallucination problem.

## REFERENCES


[1] Wang Baicun, Zang Jiyuan, Qu Xianming, Dong Jingchen, and Zhou Yanhong, “Research on New-Generation Intelligent Manufacturing based on Human-Cyber-Physical Systems,” Strategic Study of Chinese Academy of Engineering, vol. 20, no. 4, pp. 29-34, 2018.

[2] Na X, et al. “Research on the construction and implementation of power grid fault handling KGs,” Energy Reports, vol. 9, pp. 182, 2023.

[3] AIPIZAR-CHACON I, SOSNOVSKY S. “What's in an index: Extracting domain-specific KGs from text-books,” in Proceedings of the ACM Web Conference, New York, NY, USA, April 2022, pp. 966-976.

[4] AYOUB H, et al. "BERT based clinical knowledge extraction for biomedical KG construction and analysis," Computer Methods and Programs in Biomedicine Update, vol. 1, pp. 100042, 2021.

[5] HANG Peixin, Zhao Xiang, Fang Yang Zhu Huiming, and Xiao Weidong, "End-to-end knowledge triplet extraction combined with adversarial training," Journal of Computer Research and Development, vol. 56, no. 12, pp. 2536-2548, 2019.

[6] Guu K, Lee K, Tung Z, Pasupat P, and Chang M, “Retrieval augmented language model pre-training,” in Proceedings of the 37th International Conference on Machine Learning, Vienna, Jul 12-18, 2020. New York: PMLR, 2020, pp. 3929–3938.

[7] Lu, J. W., et al. "The Chatgpt after: opportunities and challenges of very large scale pre-trained models," Acta Automatica Sinica, vol. 49, no. 4, pp. 705-717, 2023.

[8] Achiam J, Adler S, Agarwal S, et al. "Gpt - 4 technical report", arXiv preprint arXiv:2303.08774, 2023. [Online]. Available: https://arxiv.org/abs/2303.08774

[9] YU J, JIANG J, and YANG L, "Improving multimodal named entity recognition via entity span detection with unified multimodal transformer," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 2020, pp. 3342-3352.

[10] MOHAMED Y L, LOBNA H, and LOTFI B R, “Information extraction from electronic medical documents: state of the art and future research directions,” Knowledge and Information Systems, vol. 65, no. 2, pp. 463–516, 2023.

[11] CHEN X, et al. “Good visual guidance makes a better extractor: hierarchical visual prefix for multi-modal entity and relation extraction,” 2022, arXiv:2205.03521. [Online]. Available: https://arxiv.org/abs/2205.03521

[12] WANG X, et al. “Named entity and relation extraction with multi-modal retrieval,” 2022, arXiv:2212.01612. [Online]. Available: https://arxiv.org/abs/2212.01612

[13] Agrawal M, Hegselmann S, Lang H, Kim Y, and Sontag D, “Large language models are few-shot clinical information extractors,” 2022, arXiv:2205.12689. [Online]. Available: https://arxiv.org/abs/2205.12689

[14] YAN Y, et al. “Large-scale relation learning for question answering over knowledge bases with pre-trained language models,” in Proceedings of the 2021 conference on empirical methods in natural language processing. Association for Computational Linguistics, 2021, pp. 3653–3660.

[15] Zaib Munazza, Zhang Wei Emma, Sheng Quan Z, Mahmood Adnan , and Zhang Yang, “Conversational question answering: a survey,” Knowledge and Information Systems, vol. 64, no. 12, pp. 3151–3195, 2022.