Knowledge graph construction method for business process instructed by prompts

Shijia Gu*, Yujia Qi
Kunlun Digital Technology Co., Ltd., CNPC Data Center, Changping, Beijing 102206, China

ABSTRACT

The technology of generative general artificial intelligence not only revolutionizes machine intelligence but also plays a significant role in enterprise digitalization. Given that natural gas sales companies offer conventional human-centered services, cost and efficiency have emerged as a set of irreconcilable contradictions. The automated intelligent customer service model, based on the new generation of artificial intelligence technology, is like an open key to high-quality development. We propose a method to create a knowledge graph of customer service business processes by guiding a multi-modal large language model to generate the knowledge graph with prompt words. Based on the open-source multi-modal large language model, the general ability of the large model was applied to identify, analyze, and extract the business process table in the "Natural Gas Customer Standardized Service Business Process Guidebook" through customized task prompt design. Ultimately, we successfully constructed a business process knowledge graph, confirming the feasibility and effectiveness of this method. This method aims to optimize intelligent customer service by providing high-quality answers. The method enhances automation, intelligence, and standardization of customer service, which improves work efficiency and controls costs simultaneously.

Keywords: Multi-modal large language model, prompt engineering, downstream task, business process, knowledge graph

1. INTRODUCTION

In recent years, artificial intelligence technology has been changing the way humans and machines interact. It has also been instrumental in helping enterprises initiate the trend of digital transformation. As for traditional natural gas sales companies, the manual customer service mode has significant limitations in terms of cost control and processing efficiency, which creates a substantial potential for solutions based on artificial intelligence technology.

The emergence of large language models, such as ChatGPT, provides technical support for the intelligent enhancement of customer service models. Automated intelligent customer service robots can overcome the limitations of manual work in several ways: (1) Automation: robots can automatically respond to customer inquiries using natural language understanding and generation technology. Without manual intervention, work efficiency is increased, and the additional costs caused by staff turnover are eliminated. Additionally, the answers provided are reasonable and accurate, based on standard documents and knowledge databases. The service continuously optimizes through big data for further improvement of service quality. Standardization is achieved by eliminating randomness in customer service reception and addressing illusion problems of large language models. Unified answers can be assisted by backtracking based on knowledge base or knowledge graph[1,2], making the customer service process more standardized and reliable.

However, constructing standard knowledge bases and knowledge graphs is often time-consuming. Customer service rules and procedures are typically presented in the form of text, tables, and pictures, necessitating a significant amount of manual work to comprehend and establish a structure suitable for storage in a database. The multi-modal large language model possesses a general understanding of natural language and images. For more specific tasks like extracting key information, understanding dependencies, and generating knowledge graphs, it requires specific prompt words to guide its actions. It's worth trying to apply a multi-modal large language model to understand the customer service process, assist in intelligent customer service to optimize the question-and-answer effect, and enhance the intelligence and standardization of customer service automation.

*gushijia@cnpc.com.cn; phone 18610932730

International Conference on Optics, Electronics, and Communication Engineering (OECE 2024), edited by Yang Yue, Proc. of SPIE Vol. 13395, 133953X · © 2024 SPIE · 0277-786X · Published under a Creative Commons Attribution CC-BY 3.0 License · doi: 10.1117/12.3049117

This paper aims to study and design the utilization of prompt words to guide multi-modal large language models in identifying and analyzing the procedural form in the standardization process manual of customer service. The ultimate goal is to generate code for the flow chart and knowledge graph of natural gas customer service. The output code can be applied to construct and store a knowledge graph. It can be combined with a large language model to develop an intelligent customer Q&A service and process step retrieval system. This can help reduce the labor costs and work repeatability for gas company employees at the sales end.

This paper consists of six parts. The first part is the introduction, which explains the background of the subject and the research content of this paper. The second part presents related works, introducing domestic and foreign works and achievements related to the subject of this paper. The third part explains the guiding principle of the multi-modal large language model to perform the downstream task step by step. The fourth part introduces the framework for the process of table recognition, content understanding, and generating flowcharts and knowledge graphs. The fifth part involves verification and evaluation, using the "Online Account Opening Process Form" and "Customer Information Collection Process Form" from the "Natural Gas Customer Standardized Service Business Process Guidance Manual" as examples. Based on the overall framework outlined in the implementation plan, we provide a detailed description of prompt setting principles and analyze the results obtained from the multi-modal large language model at each step. We also compared the construction efficiency of the knowledge graph using three methods to determine whether utilize the multi-modal large language model is superior. The sixth section is the conclusion and prospects, which briefly summarizes the feasibility and effectiveness of the research results in this paper and suggests possible directions for future improvements and enhancements in the research.

2. RELATED WORK

2.1 Large language model

Before the era of large models, text content extraction from tables, pictures, and paper books relied on Optical Character Recognition (OCR). The extracted text combined with the original image can be used to train the downstream task of Visual Question Answering (VQA), which enhances picture comprehension. Mishra et al.[3] constructed a dataset of book cover pictures, related question-and-answer pairs, and combined OCR and VQA methods to improve the model's ability to understand pictures and reasoning. However, due to limitations such as incomplete data sets, the accuracy rate of text recognition is relatively low, resulting in a low Q&A precision rate. The model lacks the ability to generate content, and its intelligence level is insufficient to assist humans in completing specific tasks.

With the expansion of datasets and the continuous optimization of large language model architectures, there have been significant breakthroughs in image-based information retrieval and content generation. Large language models are a type of natural language processing technology based on deep learning[4]. They have transformed the traditional human-computer interaction pattern, making it easier for machines to comprehend human language. This shift empowers machines to understand and communicate in natural language, thereby tilting the balance of dialogue towards humans. This advancement enables humans and machines to exchange information using human language, making it more conducive for human-computer dialogues and other interaction scenarios[5]. This is due to the dramatic increase in the scale of the model based on the explosion of computing resources, resulting in the first real emergence of intelligence in the deep neural network model[6]. The emergence of large model intelligence is known as the next generation of general intelligence[7]. A large model, composed of a very large-scale deep neural network, is capable of performing basic tasks such as dialogue, question-answering, extraction, generation, and summarization. It is also due to the versatility and multi-modal capability of the large model that it can synthesize, summarize, extend, and even complete many downstream tasks or sub-tasks[8]. For example, under the guidance of manually set prompt words, downstream tasks such as parts-of-speech tagging in NLP, human voice recognition in speech, and object detection in images can be completed. How to design appropriate prompts to guide large models in performing these tasks and continuously optimizing the results is one of the research hot spots in industry and academia.

GPT-4V[9], Gemini-1.5[10] and other multi-modal large language models have demonstrated exceptional comprehensive capabilities. When a screenshot was fed to GPT-4V, the results showed that GPT-4V could accurately recognize the text content in the image, follow the instructions provided, engage in role-playing, and respond accordingly. Gemini-1.5 can actively analyze whether the information in the pictures is accurate and determine if it can address the question posed. For example, when asked, Gemini-1.5 will respond that it cannot determine if someone is a criminal if the figure lacks relevant information such as clothing items. The multi-modal large language model has demonstrated a strong understanding and generation ability. It can serve as an auxiliary tool to assist in completing intermediate steps of specific tasks. In some cases, the entire process can be accomplished by the large model using prompt words. However, each paper only provides a few experimental examples, and does not fully present any specific downstream tasks accomplished by the multi-modal large language model. The work in this paper fills the vacancy here and achieves the downstream task of automating the generation of a knowledge graph from large models. Compared with traditional methods of constructing knowledge graphs, the efficiency is significantly improved.

2.2 Prompt engineering

Prompt engineering involves designing and improving prompt words in structured text to guide large language models for generating more satisfactory results. The goal is to enhance the efficiency and effectiveness of interactions with large language models[11]. When facing complex tasks, large models often fail to produce optimal results in the initial round of dialogue due to insufficient training data and other factors. Shinn et al.[12] proposed the Reflexion framework, which can significantly enhance the accuracy of various tasks (sequence decision, coding, linguistic reasoning) by utilizing only prompt words without updating the parameters of large models. Kojima et al.[13] tested the model on various datasets and demonstrated that by incorporating the phrase “Let’s think step by step” into the prompt, the model could autonomously generate intermediate reasoning and reach the correct answer.

Only the prompt engineering is not enough in many cases; therefore, it is necessary to introduce Chain-of-Thought. Wei et al.[14] proposed the concept of Chain-of-Thought for the first time. They suggested that introducing a series of intermediate reasoning steps can enhance the performance of large language models, particularly in application scenarios that require logical reasoning, leading to better answers. Zhang et al.[15] proposed a multi-modal thought chain paradigm that surpassed the graphic comprehension of multi-modal large language models, outperforming even the best GPT-3.5 model available at that time.

2.3 Knowledge graph

The concept of Knowledge Graph was first introduced by Google[16] in 2012. A knowledge graph is essentially a semantic network where nodes represent entities or concepts, and edges represent various semantic relationships between entities and concepts. A knowledge graph presents information graphically, clearly illustrating relationships between entities. It can be utilized to develop intelligent question-answering systems. At present, the main methods to build a knowledge graph are artificial construction and machine learning construction. Artificial construction methods require the annotation of entities and relationships, as well as verification by experts to ensure the quality and accuracy of the knowledge graph. This method requires a significant amount of manpower, time, and cost, but it can ensure the accuracy and authority of the knowledge graph. The construction of machine learning method combines unsupervised and supervised learning with a large corpus of data for entity recognition and relationship extraction. The way machine learning constructs knowledge graphs can reduce annotation costs, but training models necessitates a significant amount of text data and computational resources. Li et al.[17] proposed an entity-relation joint extraction model based on the LEBERT model, which can automatically extract relational triplets from multi-modal data. Both of these methods require a significant amount of manpower input. However, by leveraging a multi-modal large language model, we can eliminate the need for manual labor. By guiding the model to comprehend unstructured graph data through prompt engineering and thought chains, we can accurately infer the entities and relationships within the data, ultimately generating a structured knowledge graph.

3. PROMPT WORD DESIGN

3.1 Rules

According to OpenAI’s prompt engineering guidelines and practical experience, the design of prompt words in this paper follows three principles:

(1) Clear: clearly describe the tasks to be performed in the way of commands, and clearly request the expected output;

(2) Decomposition: decompose complex tasks into simple tasks, decompose simple tasks into orderly stages;

(3) Guidance: use process explanations or concrete examples to provide guidance for the large model, and perform tasks according to the ideas.

3.2 Principles

In prompt engineering, conditional probability models describe the inputs and outputs of large models:

Input design:

$$P(x|y) = \prod_{i=1}^{n} P(x_i|y)$$

where, $P(x|y)$ is the probability of the input design, x represents the input, y represents the output, n represents the length of the input, and x_i represents the i-th word of the input.

Output design:

$$P(y|x) = \prod_{i=1}^{m} P(y_i|x)$$

where, $P(y|x)$ represents the probability of the output prediction, y represents the output, x represents the input, m represents the length of the output, y_i represents the i-th word of the output.

It can be inferred from the input-output conditional probability model that the output of the large language model depends on the input. On the contrary, in order to obtain the expected output, it is necessary to design the input prompt words skillfully to guide the inference of the large model.

3.3 Structure and examples

According to the above content, this paper organizes the prompt word structure into four parts: roles and tasks, guidance, input, and output. The purpose of this is to clearly inform the large model of the task to be performed. This enables the large model to execute the task based on the input in a suitable manner to produce the output and assists in directing the large model on how to solve the task. In practice, it can be expanded and customized to application scenarios (Table 1).

Table 1. Prompt word component explanation and sample.

	Prompt	**Explain**	**Example**
1	Roles and tasks	Role and task description	You are a picture classifier, responsible for classifying according to the content in the picture.
2	Guidance	This section describes the steps for executing the task	You need to examine the animal image in the picture, and then determine whether the animal in the picture is a cat or a dog.
3	Input	The scope and format of the data	The pictures that need to be processed are the 10 animal pictures in the folder.
4	Output	In what form is the answer given	The classification result is in the form of a two-dimensional table; the first column is the serial number, and the second column is the classification results of the picture.

4. IMPLEMENTATION SCHEME

In this paper, a comprehensive process automation scheme is designed to achieve the full conversion process of transforming the PDF file of the “Natural Gas Customer Standardized Service Business Process Guidance Manual” into a knowledge graph. It is divided into two stages: document pre-processing and knowledge graph generation, the flowchart of which is shown in Figures 1 and 2.

First of all, each page of the PDF document undergoes layout recognition. For the results of layout recognition, preliminary layout classification is carried out. Other layouts are completely disregarded except for table layouts. Next, we evaluate the comprehensiveness of the table content layout. If it is incomplete, it is combined with the table on the next page to create a comprehensive content flow table.

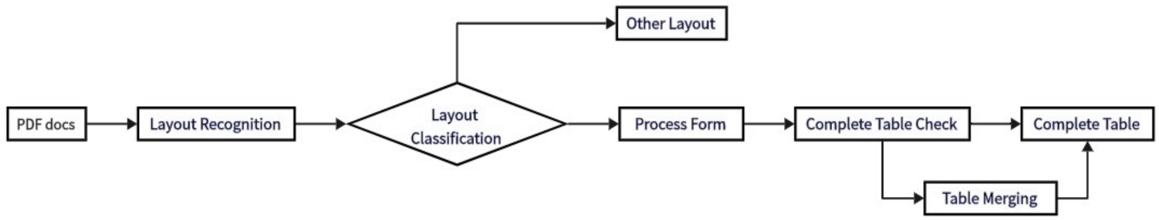

Figure 1. Flow chart of document pre-processing stage.

On the other hand, the content of the PDF document page is identified, all the contents are extracted, the semantic analysis of the table content is carried out in combination with the context provided by the table header, and the complete table layout is integrated to facilitate the understanding of the specific table content in the subsequent step. The most critical aspect of a process is the dependencies between the steps of the process. These two can be obtained through a comprehensive analysis of the process table, and then the descriptive text of the entire customer service process can be generated. Finally, based on the understanding of the process description, the code for the knowledge graph is generated to build the knowledge graph of the customer service process.

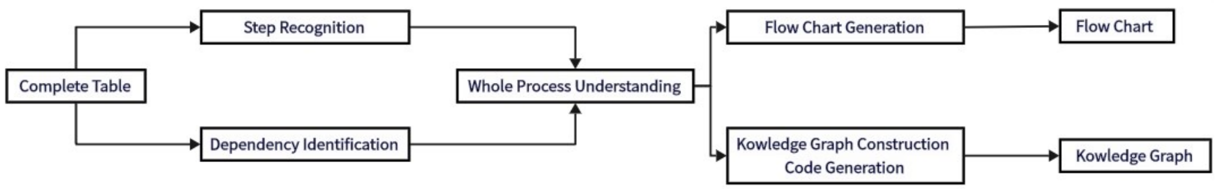

Figure 2. Flow chart of knowledge graph generation stage.

5. EXPERIMENT

5.1 Experimental environment

In this experiment, the overall B/S architecture was used to deploy services, in which the server hardware environment CPU model was Haiguang Hygon C86 7165,16 cores, and the main frequency was 2.0 GHz. Memory 64 GB; The GPU is a single NVDIA Tesla V100S with 32 GB memory. In the software environment, the OS is RedHat Enterprise Linux 7.5×64, CUDA driver version 11.6, and python version 3.10. The client browser is Google Chrome 122.0.6261.95. In particular, in order to prove the effectiveness and applicability of the method proposed in this paper, two widely used multi-modal large language models “Tongyi Qiwen Qwen-VL-Plus” and “IFlyspark Cognition large Model V3.5” were selected as the basic large models in this experiment.

5.2 Pre-processing

According to the document preprocessing process shown in Figure 1, the PDF version of the “Natural Gas Customer Standardized Service Business Process Guidance Manual” comprises a total of 427 pages. First, the document is divided into separate pages, and each page is processed as an image. Then, layout identification is performed for each page picture. We extract layout elements from the page and determine the layout type of the current page. If the document does not include the business process table, it is disregarded. However, if the table is present, it undergoes further evaluation. The integrity of the table content on the current page is verified by cross-referencing it with the layout content of the subsequent page. The content of the two parts is combined into a complete figure of the business process table. For the convenience of explanation, Figure 3 shows the pre-processing results of the “Online Account Opening Process,” a relatively simple online business process selected from the “Business Process Guidance Manual for Standardized Service of Natural Gas Customers.” The entire process is brief, and the connection between the steps is relatively straightforward.

No.	Service category	Activity	Who is operating	Form	Process description
01		Read Account Opening Instructions	Customer		Customers should first read the instructions for opening an account, which mainly includes the conditions and preparation before opening an account.
02		Submit address	Customer		Customers submit the address, and the system checks whether it belongs to the company's business area.
03	Online service	Upload account information	Customer		Resident customers submit copy of the house owner's ID card, real estate proofs; The non-resident customer shall submit the business license, the certificate of the unit, and the photo of the ID card of the legal person.
04		Sign gas supply agreement	Customer		After the data review is completed, the customer will sign the online gas supply agreement.
05		Open an account	Customer		The customer enters the payment number to pay.

Figure 3. Online account opening process table.

Because the integrity of the extracted information in the subsequent steps increases with the sharpness of the PDF file, it is converted to an image with the original resolution, renamed according to the page number, and stored in a unified folder for processing.

5.3 Construction of business process knowledge graph

Next, we analyze and interpret the business process table pictures obtained from the previous process, and generate the business flow chart and knowledge graph.

Step 1: the input prompt words as designed as follows:

You are an extracting tool of the content from the picture	←Role
You need to extract the detailed steps of the business process from the table shown in the picture	←Task
You can analyze and extract the contents of the table columns listed as "business processes"	←Guidance
Finally, all process steps are presented in the form of a two-dimensional table.	←Output

Both large models can return the same result, as shown in Table 2:

Table 2. The step process of online account opening process form.

Procedure	Process step
01	Read Account Opening Instructions
02	Submit address
03	Upload account information
04	Sign gas supply agreement
05	Open an account

Step 2: we combine dependency extraction with process understanding, asking the model to output a detailed description of the business process. The prompt words are:

You need to identify the results based on the contents of this business process table and the steps from the previous step	←Task
The whole business process should be described in details	←Guidance
The description should include all steps thoroughly and the connection between the steps should be explained clearly	←Output

The model returns the result, which accurately describes the whole process with the right logic and order: "This business process mainly involves account opening services, first of all, customers need to read the account opening instructions, understand the relevant conditions and precautions for account opening." Then, the customer submits the address, and the system reviews whether the address is within the company's business scope. After approval, resident customers are required to upload copy of house owner's ID car and real estate proofs, while non-resident customers are required to upload photos of business licenses, unit certificates and copy of legal person's ID card. Next, the customer should sign the gas supply agreement online. Finally, the customer enters the payment number and completes the payment, finishing the account opening process.

Step 3: visualization. The prompt words designed for this step are as follows:

You are a code generation tool	←Role
You need to generate flowchart Mermaid drawing code based on previous business process description	←Task and Output

The output of the large model is as follows:

	Business flow chart: Mermaid drawing code
1	Graph LR
2	A[Read the account opening instructions]-->B[Submit gas address]
3	B-->C[Audit gas address]
4	C-->D[Upload account information]
5	D-->E[Sign gas supply agreement]
6	E-->F[Account opening]

We execute this code in Mermaid to draw the following "online account opening" customer service business process. The dependency and the sequence are show below in Figure 4.

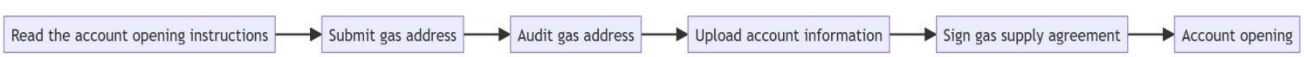

Figure 4. The step process of online account opening process.

Step 4: We generate the building code and create the knowledge graph. The prompt words are designed as follows:

You are a knowledge graph entity and relationship extractor	←Role
You need to build the corresponding knowledge graph according to the description of the Online Account Opening business process	←Task
First you have to extract entities and relationships, and then use the "online account opening service" as a node, all nodes need to start from the root node	←Guidance
Finally, the knowledge graph constructed will be visualized, and Mermaid drawing code will be generated	←Output

The entity list is as follows (Table 3):

Table 3. Entity extracts results of the Online account opening process form.

No.	Entity name	No.	Entity name
1	Online account opening service	9	Real estate proofs
2	Client	10	Non-resident customer
3	Account opening instructions	11	Business license photo
4	Gas address	12	Unit certificate photograph
5	Address auditing system	13	Copy of legal person's ID card
6	Company business scope	14	Gas supply agreement
7	Resident customer	15	Payment number
8	Copy of house owner's ID card	16	Payment link

Mermaid is used to visualize the knowledge graph, and the codes are as follows:

	Knowledge map: Mermaid drawing code
1	graph LR
2	Online Account Opening Service (Online Account Opening Service)-->Customer (Customer)
3	Customer-->Account Opening Instructions (Read Account Opening Instructions)
4	Customer-->|Submit| gas address (gas address)
5	Gas address-->|Audit| Address audit system (address range confirmation)
6	Address audit system-->|approved| Resident customers (resident customers)
7	Address Audit System-->|approved| Non-resident customers (non-resident customers)
8	Resident customers-->|Upload| head of household ID photo (ID photo)
9	Resident customer-->|Upload| real estate certificate photo (real estate certificate photo)
10	Non-resident customer-->|Upload| Business license photo (Business license photo)
11	Non-resident customer-->|Upload| unit proof photo (unit proof photo)
12	Non-resident customer-->|Upload| legal person ID photo (legal person ID photo)
13	Customer-->|Sign| Gas Supply Agreement (Sign online gas Supply agreement)
14	Customer-->|Enter| payment number (enter payment number)
15	Customer-->|Complete| payment process (complete payment)

Finally, we execute the codes to visualize the generated knowledge graph, which is structured in Figure 5 as follows:

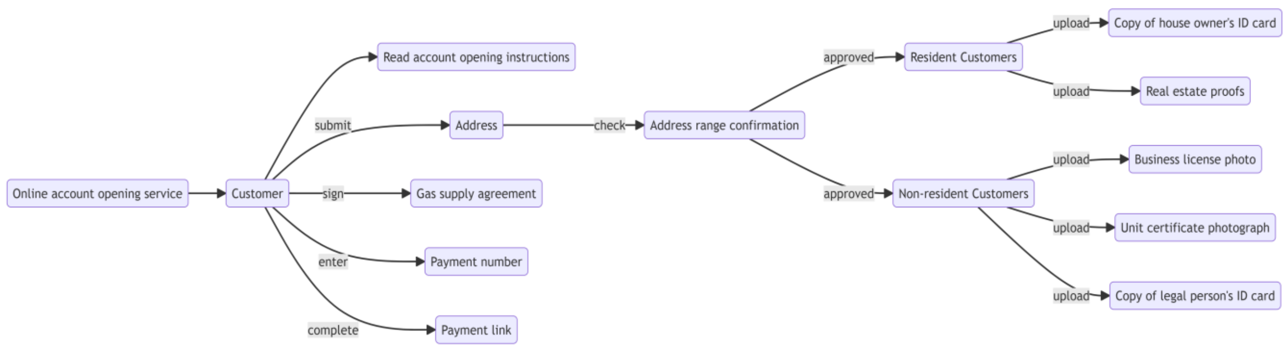

Figure 5. The step knowledge graph of online account opening process.

It is worth noticing that the flow chart splits the business activities in the table. For example, the line number 02 in Figure 2 which the business activity is “submit address” actually contains two sub-steps: “Customer submits gas address” and “system confirms address”, indicating that the flow chart is detailed and perfect, and the large model’s understanding of the table is accurate under the guidance of the prompt words.

In this experiment, it is further verified that the above method can also get better results for slightly complex business processes of the table in Figure 6.

No.	Activity	Who is operating	Form	Process description
01	Collect data	Customer service management department; IT department		Collect customer information from various ways (e.g., customer calls, customer visit and review visit, potential customers, government affairs platform); organize and summarize customer information.
02	Analyse and verify data	Customer service management department; IT department		Analyze customer information, credit status and company value, conduct classification and rating; evaluate the authenticity and accuracy of customer information.
03	Enter customer information database	Customer service management department; IT department	System	Increase or complete customer information database
	Existence	Customer service management department; IT department	System	Existing customers or potential customers not recorded in the system
04	Update data	Customer service management department; IT department	System	Update customer information in time
05	Create new data	Customer service management department; IT department	System	Put customer information into the database
06	Check information database	Customer service management department	System	
	Check result	Customer service management department		If passed, go to step no. 08, otherwise step no. 07
07	Organizational standardization rectification	Customer service management department; IT department		Return to step no. 03
08	Customer information release	Customer service management department		After the release, business related departments and call centers can share customer information.
	End	Customer service management department		

Figure 6. Customer information collection process table.

For the "Customer Information Collection" process in the "Natural Gas Customer Standardized Service Business Process Guidance Manual" shown in Figure 6, we repeat the above steps and use the same prompt words to guide the large model to perform each downstream task. The resulting visual flow chart and knowledge graph are shown in Figure 7 as follows.

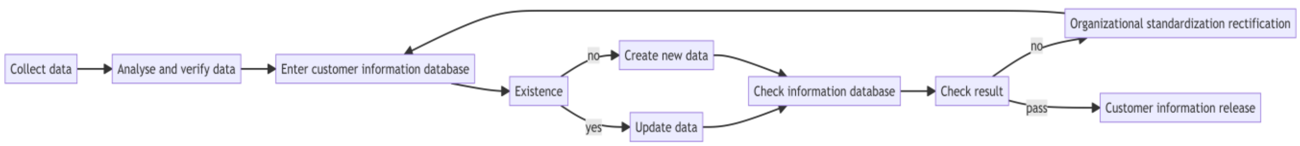

Figure 7. The step knowledge graph of customer information collection process table.

The above flow chart reflects two sets of judgment branches, which prove that under the effective guidance of prompt words, the large language model has the ability to understand the business process and execute the downstream task.

Finally, based on these entities and relationships, the knowledge graph can then be created using specialized knowledge graph building tools such as Neo4j, Stardog, or Apache Jena. In a knowledge graph, each entity is a node, and a relationship is the edge that connects those nodes. For example, the customer node is connected to the gas address node through the "submit" relation, and the original owner information table node through the "fill" relation. This structure enables the knowledge graph to clearly show the logical relationship between the relevant information and steps of the entire transfer process.

5.4 Result analysis

Through experimental verification, the method of building business process knowledge graph guided by prompt words proposed in this paper realizes the generation and construction from the business process description table to the corresponding flow chart and knowledge graph. The design of prompt words is more effective for the execution of downstream tasks, and at the same time, the expected output is obtained, and the goal of each stage is completed.

We randomly extract 10 tables and apply 3 methods to compare: manual construction method, machine learning method and prompt word-guided construction method. The results are shown in Table 4, where we can see that the prompt-based guided process steps and dependency extraction (calling the closed-source large model API) approach performed best in terms of total time. In terms of recognition rate, the positive recognition rate of process nodes reaches 90%, and the recognition rate of process relationships reaches 92.2%. Although it cannot reach 100% accuracy of manual construction, it is the best choice from the perspective of efficiency and recognition accuracy. Other methods differ in computational resource consumption or human involvement complexity, but their overall time and recognition rate are not as efficient and accurate as the "prompt word-guided process step and dependency extraction (calling the closed source large model API)" approach. Therefore, this method is the best choice for application scenarios that require efficient knowledge graph generation and high recognition rate.

Table 4. Comparison between different methods to construct knowledge graph.

Method	Total time cost	Computational resource	Entity recognition accuracy rate	Process relationship recognition accuracy rate	Manual participation complexity
Manual construction	67 min	No	100%	100%	Complex, requiring a large amount of work to annotate entities and relationships
Bert+crf was used for named entity recognition and BiLSTM was used for relationship extraction	30 min	A single NVDIA Tesla V100S with 32 GB memory	75.6%	71.9%	Complex, the model does not accept picture or table input, we need to manually convert the table into descriptions

Method	Total time cost	Computational resource	Entity recognition accuracy rate	Process relationship recognition accuracy rate	Manual participation complexity
Prompt word-guided process steps and dependency extraction (locally deployed Qwen-VL-Chat)	23 min	A single NVDIA Tesla V100S with 32 GB memory	88.3%	86.8%	None
Prompt word-guided process steps and dependency extraction (call Qwen-VL-Plus API)	16 min	no	90%	92.2%	None

In conclusion, the proposed method of constructing knowledge graphs based on multi-modal large language models offers clear advantages: (1) comprehensiveness: multi-modal large language models can typically manage more intricate semantic and inference tasks, enabling them to produce more comprehensive and structured knowledge graphs. (2) Deep understanding: Multi-modal large language models can capture deep semantic information in the text, which helps extract and represent abstract concepts and complex relationships. (3) High efficiency: The multi-modal large language model can directly generate the knowledge graph from the original data, eliminating intermediate steps and data conversion. This significantly reduces the workload while maintaining the flexibility of the knowledge graph. It should be noted that to demonstrate the effectiveness and independence of the method proposed in this paper, two commonly used multi-modal large language models were chosen as the primary models in this experiment. Similar results were achieved by designing structures using the same prompt words.

6. SUMMARY AND FUTURE WORKS

In summary, this paper introduces an innovative approach to constructing a knowledge graph from tables of business processes using prompt word guidance. This approach has been successfully applied to the standardized service business process for natural gas customers. By utilizing the process table data outlined in the Business Process Guidance Manual of Standardized Service for Natural Gas Customers and following specific keywords for design and guidance, a knowledge network that is easily understood by machines is created. This is achieved through key technologies such as step identification, dependency relationship recognition, and process comprehension. This method not only confirms the feasibility and effectiveness of constructing a knowledge graph for industry-specific business processes but also demonstrates its applicability and accuracy. The construction method of the business process knowledge graph essentially involves transforming the complex and changeable business process into a structured and networked knowledge graph representation. This enables the intelligent customer service system to quickly understand customer demands and accurately execute various service operation.

In the natural gas industry, the construction of customer standardized service business process knowledge map not only has strong theoretical significance, but also has great potential in practical application. It can significantly enhance the Q&A accuracy and response speed of intelligent customer service system, making the customer service work achieve a qualitative leap in three dimensions, automation, intelligence and standardization. For example, intelligent customer service robots can use the knowledge graph to accurately understand and quickly answer a series of common questions about registration, fee payment, pipe renovation, complaint, etc., which greatly reduces the work pressure of manual customer service, decreases incautious human error, and also ensures the consistency and efficiency of service quality, thus improve customer satisfaction.

Looking forward to the future, with the breakthrough and application of cutting-edge technologies such as multi-modal large language models and intelligent customer service robots, the entire industry, and even society as a whole are entering a new stage of development. Although the complete autonomous simulation of human cognition is still in progress, overcoming challenges such as emotional perception and deep reasoning is essential. With ongoing research and rapid technological advancements, the future intelligent customer service system is anticipated to achieve a superior level of independent learning and decision-making capabilities. It will also focus on delivering personalized experiences and humanized care. Ultimately reshaping the standards and quality of service in the customer service industry.

REFERENCES

[1] Guu, K., Lee, K., Tung, Z., et al., "Realm: Retrieval augmented language model pre-training," Proceedings of the 37th International Conference on Machine Learning 119, 3929-3938 (2020).

[2] Lewis, P., Perez, E., Piktus, A., et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," Advances in Neural Information Processing Systems 33, 9459-9474 (2020).

[3] Mishra, A., Shekhar, S., Singh, A. K., et al., "OCR-VQA: Visual question answering by reading text in images," 2019 International Conference on Document Analysis and Recognition (ICDAR), 947-952 (2019).

[4] Qiu, X., Sun, T., Xu, Y., et al., "Pre-trained models for natural language processing: A survey," Science China Technological Sciences 63(10), 1872-1897 (2020).

[5] Zhang, X., Yang, X. and Xu, C., "Current state and future development directions of ChatGPT and generative artificial intelligence," National Natural Science Foundation of China 37(5), 743-750 (2023).

[6] Wang, J., Liu, Z., Zhao, L., et al., "Review of large vision models and visual prompt engineering," Meta-Radiology 1, 100047 (2023).

[7] Chen, Y., "Beyond ChatGPT: Opportunities, risks, and challenges from generative AI," Journal of Shandong University (Philosophy and Social Sciences) 3, 127-143 (2023).

[8] Brown, T., Mann, B., Ryder, N., et al., "Language models are few-shot learners," Advances in Neural Information Processing Systems 33, 1877-1901 (2020).

[9] OpenAI. GPT-4V(ision) system card, 25 September 2023, <https://cdn.openai.com/papers/GPTV_System_Card.pdf> (25/4/2024).

[10] Gemini Team, Google, "Gemini 1.5: Unlocking multi-modal understanding across millions of tokens of context," 15 February 2024, <https://storage.googleapis.com/deepmind-media/gemini/gemini_v1_5_report.pdf> (15/2/2024).

[11] Dong, Q., Li, L., Dai, D., et al., "A survey for in-context learning," 31 December 2022, <https://arxiv.org/abs/2301.00234> (25/4/2024).

[12] Shinn, N., Cassano, F., Berman, E., et al., "Reflexion: Language agents with verbal reinforcement learning," Advances in Neural Information Processing Systems 36, 8634-8652 (2023).

[13] Kojima, T., Gu, S. S., Reid, M., et al., "Large language models are zero-shot reasoners," Advances in Neural Information Processing Systems 35, 22199-22213 (2022).

[14] Wei, J., Wang, X., Schuurmans, D., et al., "Chain-of-Thought prompting elicits reasoning in large language models," Advances in Neural Information Processing Systems 35, 24824-24837 (2022).

[15] Zhang, Z., Zhang, A., Li, M., Zhao, H., et al., "Multi-modal Chain-of-Thought reasoning in language models," 2 February 2023, <https://arxiv.org/abs/2302.00923> (25/4/2024).

[16] Ji, S., Pan, S., Cambria, E., et al., "A survey on knowledge graphs: representation, acquisition, and applications," IEEE Transactions on Neural Networks and Learning Systems 33, 494-514 (2021).

[17] Li, H., Fu, Y., Yan, Y., et al., "Construction of multi-modal domain knowledge graph based on LEBERT," Computer Systems and Applications 31(11), 79-90 (2022).