IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 June 2024, accepted 10 July 2024, date of publication 19 July 2024, date of current version 2 August 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3430902*

RESEARCH ARTICLE

# Design of Improved Artificial Intelligence Generative Dialogue Algorithm and Dialogue System Model Based on Knowledge Graph

**YIKE GUO**
Chongqing Industry Polytechnic College, Chongqing 401120, China
e-mail: gyk199507@163.com

This work was supported by the Project of Humanities and Social Sciences Research of Chongqing Municipal Education Commission in 2024 - Research on the Deep Integration Mode of Immersive Interactive Technology and Cultural Tourism Industry in the Metaverse - Taking the National 5A-level Scenic Spot Dazu Rock Carvings as an Example under Grant 24SKGH396.

**ABSTRACT** Dialogue systems are an important research direction in artificial intelligence, with broad application prospects and market value. In order to improve system efficiency and user satisfaction, an open domain generative dialogue system integrating knowledge graphs has been developed, which facilitates the utilization of rich background knowledge during dialogue generation, thereby generating more coherent and meaningful dialogue content. At the same time, based on the sequence to sequence model, a bidirectional gated loop unit is introduced to better capture contextual information and improve the model's understanding and generation ability. These results confirmed that the average values of the improved model in the training and validation sets were 98.66% and 87.34%, respectively, with loss values of 0.01 and 0.10. Compared to the baseline model, this improved model improved Hits@1 and Hits@3 by 0.09% and 0.25%, respectively. This improved model had the minimum perplexity of 17.62. The security and diversity of this improved system were 0.80 and 0.82, respectively, taking into account the balance of these two types of performance. Its correlation and fluency were 1.44 and 1.56, respectively. This indicates that this improved model is beneficial for improving the efficiency of generating dialogue and has certain effectiveness, better meeting users' needs and improve user satisfaction. This system can provide users with a better conversation experience and provide technological and innovative features for artificial intelligence dialogue assistants.

**INDEX TERMS** Dialogue system, generative, knowledge graph, Seq2Seq model, bidirectional GRU.

## I. INTRODUCTION

As artificial intelligence (AI) developed, breakthroughs were made in natural language processing, particularly in the Generative Dialogue System (GDS). GDS can simulate human dialogue, providing users with a more natural and boundary interactive experience, which has broad application prospects in multiple fields such as education, healthcare, and technology [1]. Knowledge Graph (KG) combines methods with bibliometric citation analysis, co-occurrence analysis, and other methods in applied mathematics, information science, graphics, and information visualization technology. At the same time, KG visually displays the core structure, development history, cutting-edge fields, and discipline's overall knowledge architecture, achieving the interdisciplinary integration in modern theory [2]. Sequence to Sequence (Seq2Seq) is a model used to learn the mapping relationship between a pair of time series data, that is, input one sequence and output another sequence. The focus of this structure is that the length of the input and output sequences is variable. Seq2Seq is generally used in GDS, but it can only access past conversation context information. Gated Recurrent Unit (GRU) is commonly used in sequence data processing and natural language processing to control the flow of information

The associate editor coordinating the review of this manuscript and approving it for publication was Abdel-Hamid Soliman.

in a sequence through gating mechanisms. GRU has good readability, scalability, and flexibility, and has been widely utilized in computer vision [3]. Bidirectional Recurrent Neural Network (BRNN) can process sequence data, considering both past and future information in a sequence of data. Traditional GDS currently contains limited knowledge content and has certain limitations on the search scope of question answering systems. This leads to the inability to highly meet user needs, a lack of diversity in generated responses, and significant differences from actual responses. Therefore, an open domain generative system is constructed, combined with KG's Improved Generative Dialogue Based on Knowledge Graph (KG-IGD) algorithm, which improves Seq2Seq with bidirectional GRU. The research aims to address the issues of generative dialogue systems in generating diverse responses and improving user satisfaction. If a generative dialogue system can generate diverse responses when interacting with users, it can improve user satisfaction. However, traditional generative models often struggle to flexibly adjust information features, resulting in a lack of diversity and contextual understanding ability in the generated dialogue content. To address this issue, innovative research has been conducted by combining knowledge graphs (KG) with generative models, enabling the models to access and utilize richer knowledge content, thereby enhancing the intelligence level of generative dialogue systems. In addition, a bidirectional GRU is introduced based on the Seq2Seq model, which simultaneously processes the forward and backward information of the input sequence to better understand the contextual information of the dialogue system. Through attention mechanism, the model can flexibly adjust information features during the generation of dialogue, ensuring the diversity of content.

The research mainly includes four parts. Part 1 is a review of research on AI dialogue systems and KG. Part 2 is an introduction to open domain GDS that combines KG-IGD. The first section is about building an open domain GDS that integrates KG. Secondly, an open domain generative dialogue algorithm with bidirectional GRU is designed. Then KG-IGD is built. Part 3 is an analysis of open domain GDS results combined with KG-IGD. Firstly, the generative dialogue KG-IGD's performance is tested. Next is the application analysis of the KG-IGD improvement system. Part 4 summarizes the open domain GDS combined with KG-IGD.

## II. RELATED WORKS

GDS research is supported by fields such as AI, natural language processing, and dialogue systems. Ali et al. built an intelligent medical service dialogue system that combined incremental learning for better medical decision-making. This method offered medical guidance and recommending services, thereby improving the satisfaction of human-computer interaction and the medical knowledge acquisition accuracy. This system had an accuracy rate of 95% [4]. Firdaus et al. developed a conditional variational autoencoder for expressing emotions in dialogue systems by combining multimodal attention. Relevant datasets were established in the experiment to achieve emotional control of dialogue generation tasks, thereby improving user satisfaction and confirming this method's good performance [5]. Zhang et al. developed a variational learning network to predict emotional responses in human-computer conversations. This method learned the potential response distribution of a given session and improved the ability to predict the probability distribution of emotions, confirmingthis network's good performance [6]. Fakhrurroja et al. proposed a dialogue system that combined reinforcement learning and multimodal fusion algorithms to make dialogue systems more contextualized. This method improved the accuracy of system development and ensured a simpler and more natural system, confirming an average accuracy rate of 93.54% [7]. Miandashti et al. collected and detected dialogue scenes through clustering algorithms and function representation to identify conflicts between smart homes and users. This improved user satisfaction and facilitated the design and deployment of smart home systems, confirming this method's reliability [8]. Li et al. focused on reciprocity in human-computer interaction and analyzed the reciprocal cooperation between humans and agents in complex dynamic environments using near-infrared spectroscopy. This promoted a dialogue system for collaborative and collaborative performance and avoided inefficient use of system resources, confirming this technology's practicality [9].

KG, as background knowledge, can achieve semantic search and intelligent recommendation of high-quality learning resources, improving resource acquisition efficiency. Gaur et al. used KG combined with deep learning models to integrate computational knowledge to improve computing power. They effectively analyzed the acquisition of knowledge and avoided excessive reliance on massive data, confirming this method's feasibility [10]. Yuan et al. developed a dialogue generation model based on KG and predicted the knowledge span related to content. Meanwhile, they pre-trained the encoder to filter knowledge selection tasks and merge generated responses, demonstrating this model's good performance [11]. Shen et al. designed a KG group network to understand the impact of environmental exposure and behavioral factors on disease risk. This method visualized the relationship between exposure and outcomes in different publications, promoting knowledge-driven group discovery and dissemination, confirming this method's effectiveness [12]. Fang et al. constructed a KG model that combined reference graph embedding to address the global interdependence between word level information. Character vectors and candidate entities were determined, confirming this method's effectiveness [13]. Shelar et al. extracted information from unstructured text data using KG to improve this model's overall performance. Meanwhile, they imported named entity recognition, which improved spatial accuracy and training accuracy, confirming this method's good performance [14]. Li et al. proposed a KG update specific task for meta learning with few shots. This method adapted to

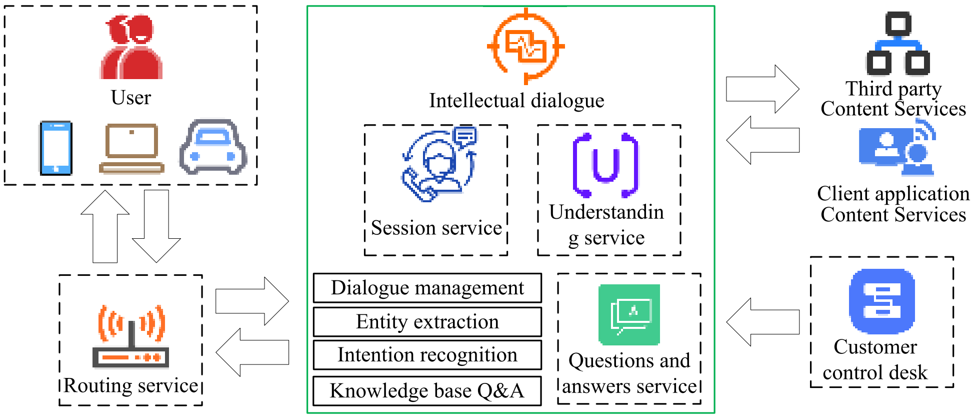


**FIGURE 1. Open domain generative dialogue system.**

new tasks, avoided overfitting, and improved computational efficiency, confirming that this method's advantages were relatively obvious [15].

In summary, significant achievements have been made in GDS and KG. However, in obtaining the current conversation information, GDS cannot guarantee the timeliness and accuracy. Meanwhile, GDS is difficult to determine personalized recommendations for dialogue content in a targeted manner. Therefore, an open domain GDS integrating KG is designed. For the generative algorithm in the system, based on Seq2Seq, the experiment innovatively replaces the original GRU unit with bidirectional GRU to obtain past and future dialogue context information.

## III. AN OPEN DOMAIN GENERATIVE DIALOGUE SYSTEM BASED ON KG-IGD

The performance and advantages and disadvantages of open domain GDS are introduced, and an open domain GDS is built, with the selection of integrating KG to enrich the generated content. Meanwhile, the baseline model Seq2Seq of the encoder-decoder structure is introduced, and a new algorithm KG-IGD is improved by combining bidirectional GRU.

### *A. AN OPEN DOMAIN GENERATIVE DIALOGUE SYSTEM THAT INTEGRATES KNOWLEDGE GRAPHS*

Open domain GDS refers to a dialogue system for non-specific fields or industries, which can engage in free dialogue with users without being limited by knowledge and rules of specific fields or industries. Open domain GDS requires a wider range of language comprehension and generation capabilities to have natural and smooth conversations with users. Open domain GDS needs to handle a wider range of language phenomena and user behavior, as well as more complex natural language understanding and generation tasks [16]. Figure 1 shows the open domain GDS.

In Figure 1, the words spoken by the user are converted into text through automatic speech recognition and enter the dialogue system. After semantic understanding and dialogue decision-making in the dialogue system, the specified content service is called and the text contentsare output. These contents are then converted from text to speech and returned to the user on the client. The dialogue system provides accurate answers directly based on the user's questions, which is a process of information retrieval. A knowledge base needs to be prepared in advance, which can include one or more domains. When a user asks a question, the knowledge base will find semantically matched answers from the knowledge base based on the sentences the user asks. To ensure the diversity of generated content and improve the knowledge utilization of replies, an AIGDS combined with an external knowledge base is built. The decoder can select knowledge that is more suitable for the current conversation content, avoiding noise generated by knowledge unrelated to the current conversation content. To ensure that the dialogue system has a massive external knowledge base, a generative dialogue model and KG fusion are designed. The basic storage form of KG is a triplet of entities, relationships, and entities. The KG construction method involves summarizing and organizing entities, forming underlying concepts, and gradually abstracting them upwards to form upper level concepts [17]. The research needs to construct a dialogue KG for use, as shown in Figure 2.

In Figure 2, conversation data are collected, and the text data are determined. These dataare manually annotated as training data. The remaining data are used as test data. Subsequently, the untrained deep learning extractor completes training under the training model, and triplet data are obtained, thereby constructing a dialogue KG. KG stores a large amount of world knowledge in a structured manner, including entities, concepts, attributes, and their relationships. By accessing KG, this dialogue system can utilize this information to generate richer and more accurate answers, making the dialogue content more profound and extensive. Meanwhile, KG can help dialogue systems better understand user queries and intentions, especially when dealing with complex queries or specific domain problems. This system can utilize the information in KG to gain a deeper semantic understanding of user input, thereby providing more

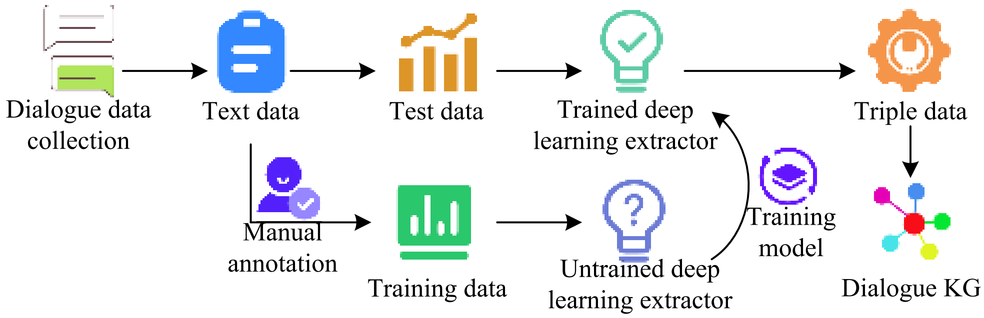


**FIGURE 2.** Process of building dialogue KG.

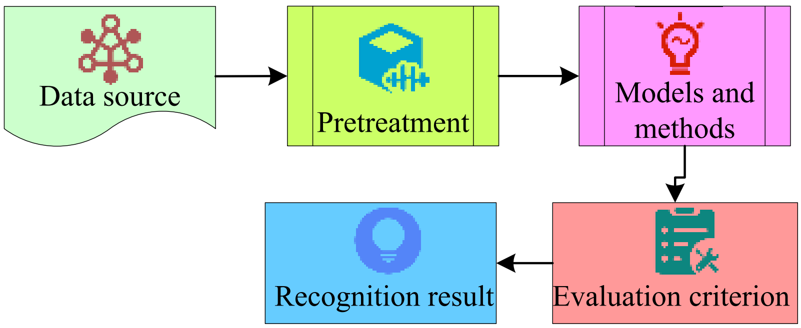


**FIGURE 3.** Named entity recognition process.

accurate responses. Generally, KG can store information about user preferences, historical interactions, etc., enabling dialogue systems to provide a more personalized dialogue experience. By analyzing the interaction between users and the system, this system can learn about their interests and needs, and then use the KG information to provide customized answers. Building a dialogue data KG is constantly updated and iterated, each iteration involving knowledge extracting, fusing, and processing techniques. A key issue in information extraction is how to automatically extract information from heterogeneous data sources to obtain candidate knowledge units. Knowledge extracting is a technology that automatically extracts structured information such as entities, relationships, and entity attributes from structured, semi-structured, and unstructured data [18]. Figure 3 shows the named entity recognition process.

In Figure 3, after preprocessing the data source with word segmentation and part of speech tagging, a model and method are selected, and the recognition results are analyzed based on evaluation criteria. The results of information extraction may own redundant and erroneous information. There is a lack of hierarchy and logic between data, resulting in flattening. Therefore, it should clean up and integrate the results. Knowledge fusing mainly includes entity linking and knowledge merging. After information extracting, knowledge elements such as entities, relationships, and attributes can be obtained from the original corpus. Through knowledge fusing, the ambiguity between entity referents and entity objects can be eliminated, resulting in a series of basic factual expressions. Nevertheless, facts do not equal knowledge. It also needs to go through the knowledge processing to obtain a structured and networked knowledge system [19].

### *B. OPEN DOMAIN GENERATIVE DIALOGUE ALGORITHM BASED ON BIDIRECTIONAL GRU*

After building an open domain generative system, an AI generative dialogue algorithm is designed. Considering that the algorithm needs to reduce the size of the question answering system, it should ensure that this model achieves timeliness and accuracy in obtaining the current dialogue information. Meanwhile, personalized content recommendations need to be provided to users. Seq2Seq is a deep learning architecture primarily designed to handle tasks where both input and output are variable length sequences. For example, it can handle natural languages such as machine translating, text summarizing, speech recognition to text conversion, and chatbot conversations. The core idea of Seq2Seq is to use an encoder decoder structure. Encoder encodes the entire input sequence of dialogue context and other personalized information and outputs a fixed length vector. The decoder decodes the semantic vector output, integrates key information and dialogue context as response content, and outputs the generated response sequence $y$. Its advantage is that there is no need to establish rules and it utilizes self-directed learning to generate new text. However, the response content of this method is uncontrollable. Sometimes grammar errors may occur and there are many universal responses, resulting in

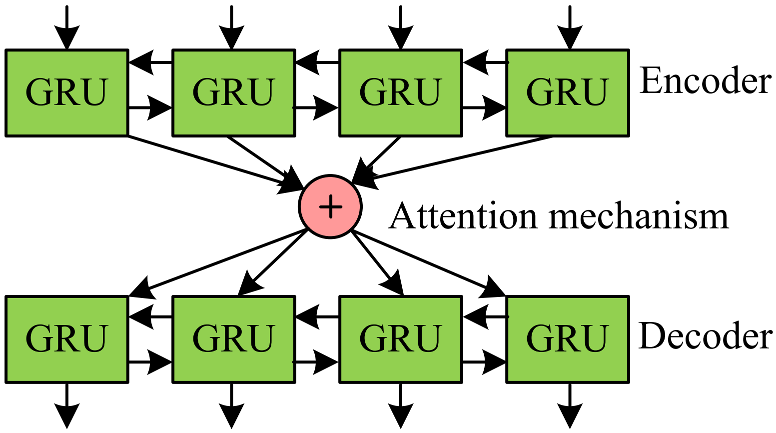


**FIGURE 4. Seq2Seq model combining attention mechanism.**

low diversity and relevance of generated content [20]. The study selects Seq2Seq Model Combining Attention Mechanism (S2SA) that combines attention mechanisms in Figure 4.

In Figure 4, GRU is updated based on historical time state and input to predict continuous values that change over time. Due to the short training time and strong noise resistance of GRU, which can effectively capture semantic associations between long sequences, GRU is chosen as the encoding structure. Attention mechanism can mimic the human visual and cognitive system, allowing Neural Network (NN) to concentrate relevant parts when processing input data. Conmbingthis method, NN can automatically learn and selectively concentrate important information in the input, improving the model's generalization ability. In Seq2Seq, attention mechanisms are typically integrated into the decoder to address the issues encountered by the original Seq2Seq when processing long input sequences. That is, the decoder can only rely on a fixed length context vector output by the encoder when generating each output element. This may not fully capture all the important information of the input sequence. After adding an attention mechanism on the decoder side, the decoder can dynamically focus on different parts of the input sequence when generating each output word. The decoder can calculate a weight distribution at each step of the generation based on the current generated state and previous contextual information. This distribution represents attention to the positions of the input column. The decoder can recombine all hidden states of the encoder in real-time based on these weights, thereby obtaining a more flexible and targeted context vector for the current generation step [21], [22].

The ability of general sequence models to access contextual information only applies to past moments. If this ability is extended to future moments, it will improve model performance. Therefore, this study introduces BRNN, namely, improving the model with bidirectional GRU. BRNN can deal with input data in both two directions and is applied to natural language processing tasks. Moreover, BRNN can acquire contextual dependencies in input data by taking past and future contexts into account [23]. BRNN has two independent RNNs that deal with input data in two directions. These RNNs' outputs are combined together to produce the final output. BRNN adds a hidden layer for reverse information transmission to handle such information more flexibly. Figure 5 shows the architecture of BRNN.

In Figure 5, in general, the forward input passes through the hidden layer and then outputs, followed by downward propagation. In the opposite direction, these two work together to output results. An encoder decoder NN structure composed of BRNN is improved to meet the requirement of different input and output sequence lengths, enabling the conversion from one sequence to another. The forward propagation layer calculates all values from time 1 to time $t$ and obtains each forward output. The forward hidden state is represented by equation (1).

$$h_t = f\left(w_1 x_t + w_2 h_{t-1}\right) \tag{1}$$

The back propagation layer calculates everything in reverse from $t$ to 1. The reverse hidden state is represented by equation (2).

$$h'_t = f\left(w_3 x_t + w_5 h'_{t+1}\right) \tag{2}$$

The output state is represented by equation (3).

$$o_t = g\left(w_4 h_t + w_6 h'_t\right) \tag{3}$$

### C. KG-IGD GENERATIVE DIALOGUE SYSTEM

Based on S2SA, two modules are added, namely dialogue scenes and user information. These methods do not effectively limit the search scope of the question answering system. Therefore, this study introduces dialogue scenario information to narrow down the search scope and determine the dialogue content in a targeted manner. Meanwhile, this model does not fully understand the emotional expression of user discourse, resulting in the generated response not

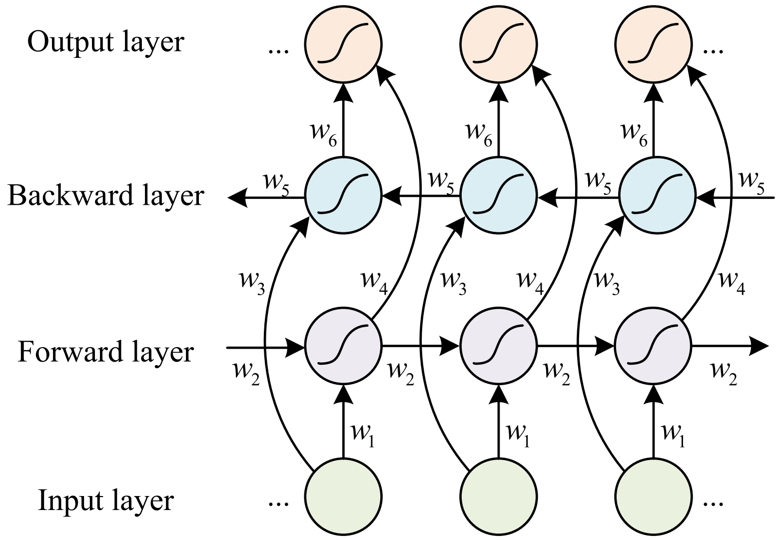


**FIGURE 5.** BRNN architecture.

meeting the true needs of the customer. Therefore, this study proposes a KG-IGD to generate an external knowledge base. During the KG-IGD testing, external knowledge base (K), dialogue scenario (ST), target sequence (G), dialogue context (X), and user information (UP)are input into the Feed forward Neural Network (FNN). During the training phase, these five types of data and corresponding sample data Y are input [24]. Six types of data are input into FNN, and the final output is encoded. The schematic diagram of the KG-IGD algorithm structure is shown in Figure 6.

In Figure 6, KG-IGD selects 5 encoders and completes fusion and normalization processing after encoding. In the encoding of KG-IGD, the update of the hidden layer $h_t$ at each time step is represented by equation (4).

$$h_t = f\left(h_{t-1}, x_t\right) \tag{4}$$

In equation (4), $f$ is a nonlinear activation function. The Softmax function is commonly used in the output layer of multi-classification problems, which can convert the output values into probability distributions and improve the model learning ability and performance. In the encoder, the model encodes a sequence of variable length input from the outside, then outputs a vector of fixed length, and then performs nonlinear transformation on the vector. The hidden state of each hidden layer ultimately becomes a semantic vector $C$, represented by equation (5).

$$C = \varphi\left(h_1, h_2, \cdots, h_t\right) \tag{5}$$

In equation (5), $\varphi$ refers to the operation on the hidden state vector. $t$ is the quantity of input sequences. In the decoding of KG-IGD, regardless of whether the same NN as the encoding stage is used or not, the next output is generated after decoding through $C$ and output sequence, represented by equation (6).

$$h_t = f\left(C, y_{t-1}, h_{t-1}\right) \tag{6}$$

The output sequence is represented by equation (7).

$$P\left(y_t \left| y_1, y_2, \cdots, y_{t-1}, C\right.\right) = g\left(C, y_{t-1}, h_t\right) \tag{7}$$

Under complete training samples, to achieve statistical consistency, KG-IGD maximizes the current probability, represented by equation (8).

$$P\left(y \left| x\right.\right) = \prod\nolimits_{t=1} P\left(y_t \left| y_1, y_2, \cdots, y_{t-1}, x\right.\right) \tag{8}$$

Considering that the result obtained from equation (8) may be smaller than the minimum value that can be represented, the logarithmic likelihood of the function is optimized through maximum likelihood, represented by equation (9).

$$\max \frac{1}{N} \sum\nolimits_{n=1}^{N} \log p\left(y_n \left| x_n\right.\right) \tag{9}$$

When decoding each $y$, if the current decoding sequence length matches the specified decoding, or if there is a specified special ending character in the decoding, the decoding of the current conversation is terminated [25]. By inputting the vector obtained after decoding into the Softmax function, the currently decoded word's true probability distribution can be determined. Compared with all values, the highest probability value is the $y$ obtained through decoding [26]. KG-IGD has 5 encoders, 1 knowledge selector, and 2 decoders. Figure 7 shows the principle of internal feedback.

In Figure 7, the model is equipped with three loss functions: Bag of Words (BOW), Negative Log Likelihood (NLL), and Kullback Leibler Divergence (KLDiv). To ensure the accuracy of sampling knowledge $k_i$, the BOW loss function utilizes the representation of the correlation between the target answer and the mandatory knowledge [27]. It is used to capture the more subtle semantic relationship between the target response and knowledge and to measure the difference between the model's output and the target output. The BOW

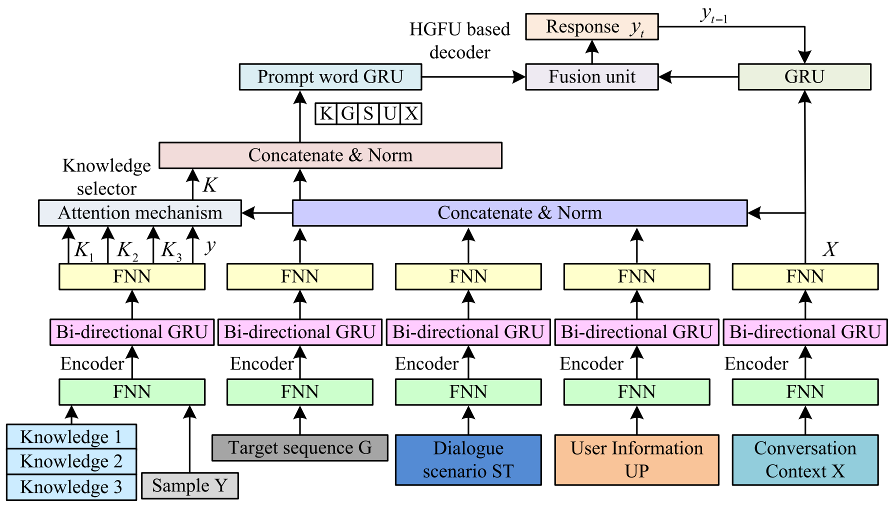


**FIGURE 6.** Schematic diagram of KG-IGD algorithm structure.

loss function is represented by equation (10).

$$L_{BOW}(\theta) = -\frac{1}{m}\sum\nolimits_{t=1}^{m} \log p_\theta (y_t | k_i) \tag{10}$$

The BOW loss function allows the model to focus more on content generation during learning by relaxing the precise matching requirements at the position, thereby improving the diversity and quality of generated text. NLL Loss is used to determine the system generated and actual responses' difference, directly related to the probability distribution of the generated text. NLL Loss is generally used in conjunction with Softmax layers. This combination often appears in text generation classification tasks. During training, the model parameters are updated using the gradient descent algorithm to minimize NLL loss [28]. By minimizing NLL Loss, the generated text sequence is more in line with the distribution of real data, thereby improving the ability of the generative language model to generate reasonable text sequences. The relationship of NLL Loss is represented by equation (11).

$$L_{NLL}(\theta) = -\frac{1}{m}\sum\nolimits_{t=1}^{m} \log p_\theta (y_t | y < t, x, k_i) \tag{11}$$

KLDivLoss can be used to measure the data generated by the generator and the real data's difference. This helps to reduce pattern collapse and improve the richness of generated content. By minimizing KLDivLoss, the generator can generate samples that are closer to real data and better match the generation path of the target distribution, thereby increasing the diversity of answers. The posterior knowledge distribution $p(k_i | x, y, g_c)$ is represented by equation (12).

$$p(k_i | x, y, g_c) = \frac{\exp(k_i \cdot MLP([x; y; g_c]))}{\sum_{j=1}^{N} \exp\left(k_j \cdot MLP([x; y; g_c])\right)} \tag{12}$$

In equation (12), the prior knowledge distribution $p(k_i | x, g_c)$ is represented by equation (13).

$$p(k_i | x, g_c) = \frac{\exp(k_i \cdot MLP([x; g_c]))}{\sum_{j=1}^{N} \exp\left(k_j \cdot MLP([x; g_c])\right)} \tag{13}$$

The relationship of KLDivLoss is represented by equation (14).

$$L_{KL}(\theta) = \frac{1}{N}\sum\nolimits_{i=1}^{N} p(k_i | x, y, g_c) \log \frac{p(k_i | x, y, g_c)}{p(k_i | x, g_c)} \tag{14}$$

The final loss function is represented by equation (15).

$$L(\theta) = \alpha \cdot L_{KL}(\theta) + \alpha \cdot L_{NLL}(\theta) + L_{BOW}(\theta) \tag{15}$$

In equation (15), $\alpha$ is a trainable parameter.

## IV. RESULT ANALYSIS OF AN OPEN DOMAIN GENERATIVE DIALOGUE SYSTEM BASED ON KG-IGD

Performance testing of generative dialogue KG-IGD was conducted, and performance results were compared with different models based on different evaluation indicators. Meanwhile, the KG-IGD improved system's practicality was verified from the actual test results.

### A. PERFORMANCE TESTING OF GENERATIVE DIALOGUE KG-IGD

The study aims to evaluate the performance of an open domain generative dialogue system based on the KG-IGD algorithm and verify its practicality in practical applications. The experimental design mainly includes steps such as dataset preparation, model training, performance testing and comparison, as well as practical application analysis. The selected datasets for the study are DuRecDials [29] and Wizard of Wikipedia [30]. Both are designed for open domain dialogue systems, aimingat enabling dialogue systems to generate rich and informative dialogues using the information contained within them. The dataset was divided into 80% for training, 10% for testing, and 10% for validation. Table 1 presents the experimental hardware environment parameters.

The research set batch size to 1, embedding size to 300, learning rate to 0.0005, objective function to Adam, pretraining iteration to 5, iteration to 100, dropout to 0.3. The

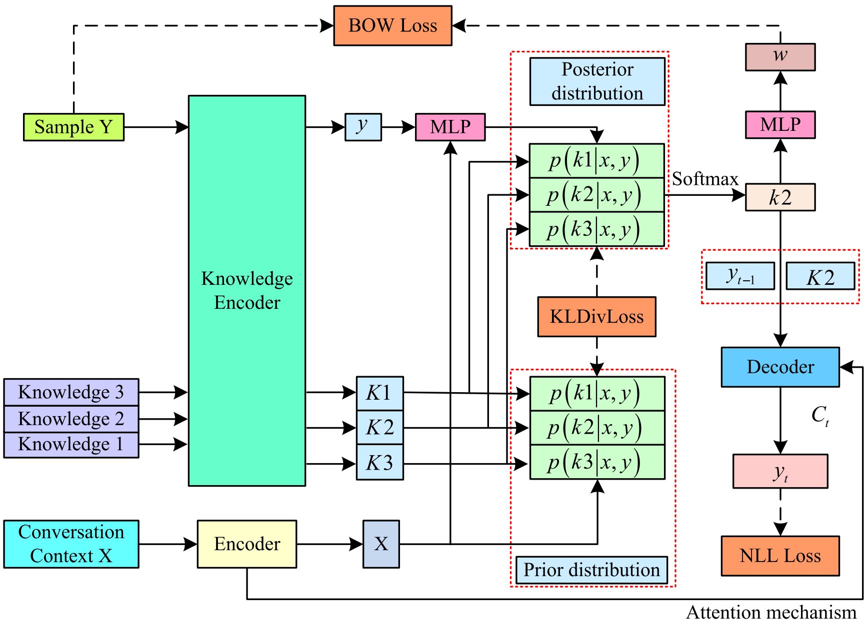


**FIGURE 7.** Internal schematic diagram.

**TABLE 1.** Experimental hardware environment parameters.

| Name | Parameter |
|---|---|
| System | Ubuntu 16.04 |
| Framework | PaddlePaddle |
| Graphics card | NVIDIA RTX Titan |
| Language | Python |
| CPU | Intel i9 7900X |

evaluation indicators are: Hits@1, Hits@3, F1, and Bilingual Evaluation Understudy (BLEU2). The indicator for evaluating language models is Perplexity (PPL), and Knowledge P/R/F1 evaluates the model's ability to select knowledge. Among them, Hits@1 and Hits@3 are commonly used indicators in knowledge graphs, and F1 can clearly reflect the correctness and accuracy of the model. BLEU2 measures the accuracy of word level machine translation and the fluency of sentences. PPL is a commonly used indicator to measure the quality of language models. Knowledge P/R/F1 is the application of Precision, Recall, and F1 measurement methods to evaluate the quality of knowledge in a knowledge graph. DIST-2 is used to measure the diversity of generated text. To verify KG-IGD, research was conducted on training KG-IGD. Figure 8 shows the accuracy and loss curves obtained.

In Figure 8 (a), the average accuracy of KG-IGD on the training set and test set was 96.45% and 83.16%. The fluctuation range of the variation curve was large, considering the presence of noise in the selected dataset. In Figure 8 (b), after the convergence of the loss curves for the training and testing sets, the KG-IGD losses were 0.03 and 0.07, respectively. In the first 20 iterations, the loss curve showed a rapid downward trend and exhibited severe fluctuations. To obtain more accurate results, KG-IGD underwent iterative training with epoch = 300. Figure 9 shows the accuracy, loss values, and PPL in two datasets.

In Figure 9 (a), as the iteration increased, the accuracy gradually increased, and the average values of the training and validation sets were 98.66% and 87.34%, respectively. This indicated that this improved model improved performance and had certain effectiveness. In Figure 9 (b), the loss value gradually decreased with the increasing iteration. The loss values for the training set and validation set were 0.01 and 0.10. This improved model reduced performance loss and improved the efficiency of generating dialogue. In Figure 9 (c), the gradual decrease in PPL indicated this improved model's feasibility, gradually converging at iteration 90. Then the evaluation index results of different models Seq2Seq, S2SA, and KG-S2SA were compared in Figure 10.

In Figure 10 (a), the Hits@1 and Hits@3 of KG-IGD were the highest, with values of 14.46% and 39.97%, respectively, indicating the improved model's better performance. Compared with the Seq2Seq model, the KG-IGD model Hits@1 and Hits@3 The indicators have increased by 7.71% and 15.45% respectively. The F1 of the KG-IGD model is the highest, at 37.20%, which is 13.25% higher than the Seq2Seq model. KG-S2SA is Seq2Seq that integrates KG and contains attention mechanisms. Hits@1 and Hits@3 of KG-IGD increased by 0.09% and 0.25%, respectively. KG-IGD had the minimum PPL of 17.62, indicating that the response content generated by this model had a high probability of matching sentences. The Seq2Seq model has the highest PPL

**TABLE 2.** Comparison of text results generated by different systems.

| | | | |
|---|---|---|---|
| Post | The weather is really nice today. | Wish you a happy holiday! | It seems that our views are consistent. |
| Seq2Seq | Yes. | Thank you. | That's right. |
| S2SA | Yes, the weather is really nice. | Thank you! | Yes, it's consistent. |
| KG-S2SA | Yes, the weather is really nice. | Thank you for your blessings! | Yes, the idea is consistent. |
| KG-IGD | Yes, this weather is suitable for going out for a walk. | Thank you, I'm glad to receive your greetings today! | Well, our ideas happen to coincide. |

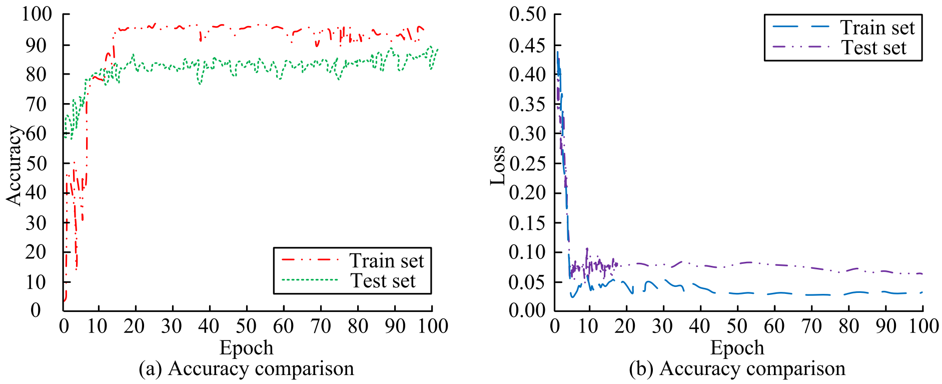


**FIGURE 8.** Changes in accuracy and loss results of KG-IGD.

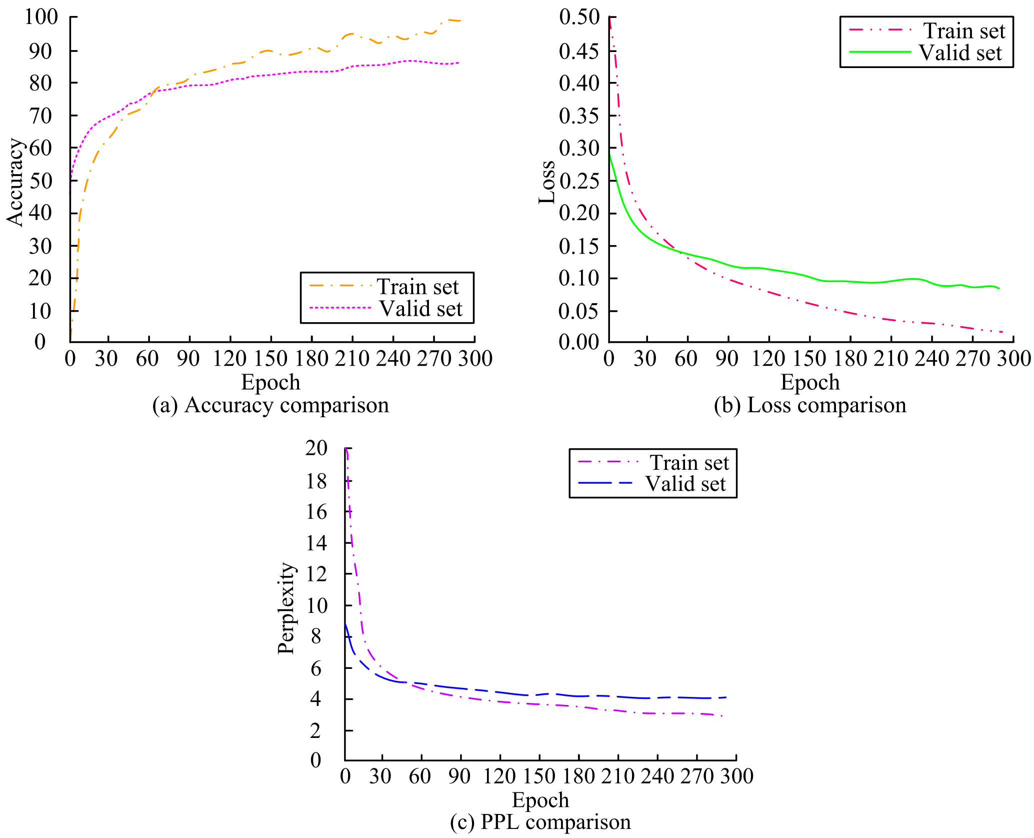


**FIGURE 9.** Accuracy, loss value, and PPL comparison of KG-IGD.

of 27.33, while the KG-IGD model has a 35.53% decrease in PPL in comparison. In Figure 10 (b), the Knowledge P/R/F1 of KG-IGD was the highest, with values of 0.408, 0.387, and 0.397, respectively, indicating that this model had the

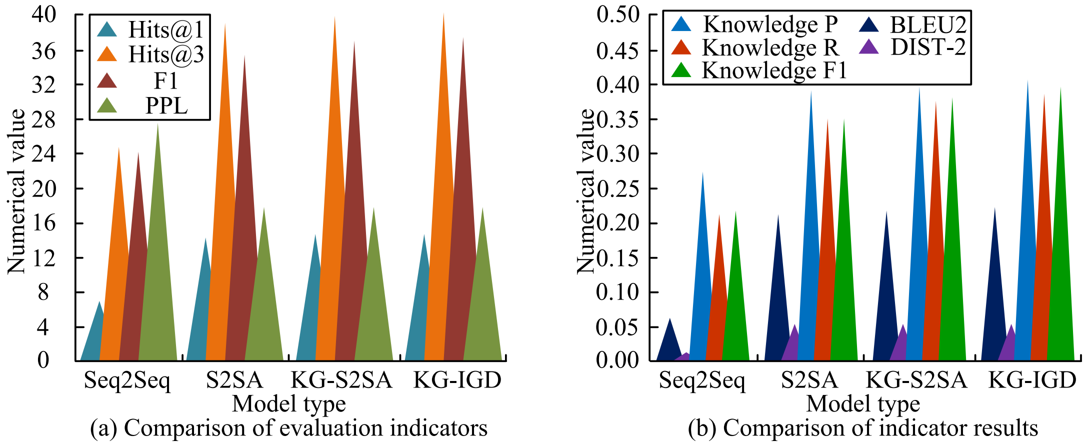


**FIGURE 10. Comparison of indicator results of different models.**

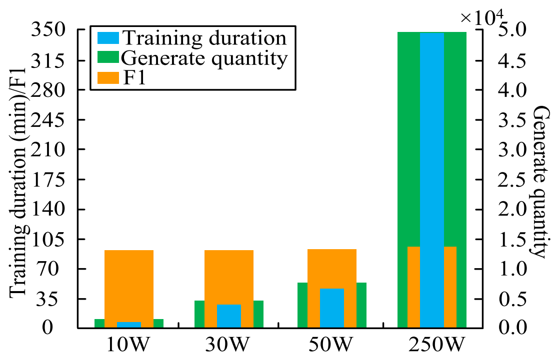


**FIGURE 11. Comparison of experimental results on datasets of different sizes.**

best utilization of knowledge. The DIST-2 of the Seq2Seq model is the smallest, at 0.011.The DIST-2 index of KG-S2SA and KG-IGD was the highest, both of which were 0.052, indicating that the response content generated by the two models was more diverse. The reason is that both the KG-S2SA model and the KG-IGD model integrate knowledge graphs, which can utilize more contextual information and knowledge nodes while generating dialogues. The BLEU2 of KG-IGD was 0.220, indicating that this model had more high-quality responses and better met personalized recommendations.

### B. APPLICATION ANALYSIS OF OPEN DOMAIN GENERATIVE DIALOGUE SYSTEM BASED ON KG-IGD

Under the same experimental environment, KG-IGD was iteratively trained with epoch = 30 on datasets of different sizes of 10W, 30W, 50W, and 250W. The 10W size dataset is called Daily Dialogue, which is a high-quality multi-disciplinary dialogue dataset that includes various topics in daily conversations. The 30W dataset is Persona Chat, which contains high-quality datasets for personalized conversations between characters. The 50W dataset is Twitter Dialogue Corpus, which is dialogue data collected from the Twitter platform and contains a large amount of real social media conversations. The 250W dataset is the Reddit Large Scale Dataset, which is a large-scale conversation data collected on the Reddit platform and includes various topics and discussions. Figure 11 shows the training duration, generated conversations, and F1 value of the test set obtained by the system.

In Figure 11, as the data increased, the F1 value of the test set for KG-IGD gradually increased, resulting in an increase in generated dialogues, while the proportion of training duration growth remained at 1.2 times. On the 250W dataset, the F1 value of the KG-IGD test set was 96.18%, indicating this system's good performance. When training in small datasets, the shorter the training time, the lower the cost of manual annotation. The iterative ability of the model was significantly improved. The study analyzed whether the generated dialogues affected the system's generation performance in Figure 12.

In Figure 12 (a), as the generate quantity increased, the improved system's accuracy did not differ significantly, indicating stability and reliability. When the generate quantity was 0, the improved system improved Hits@1 and Hits@3, with both increased by 7.63% and 9.01%, respectively. In Figure 12 (b), as the generate quantity increased, the BLEU2 and DIST-2 result curves first increased and then decreased, indicating the diversity increase of the system generated content. However, excessive generation may result in irrelevant noise, leading to a decrease in metric results and affecting performance. To analyze the practicality, different algorithms chose a single round dialogue mode and compared the final response text generated in Table 2.

In Table 2, different models had vary ingdialogue responses based on dialogue input. The response of the baseline model Seq2Seq was relatively simple. However, S2SA's response only affirmed the repeatability of the input content, which was relatively standardized. In contrast, KG-IGD's responses were more diverse, and responses to good weather were tailored to user habits, significantly improving

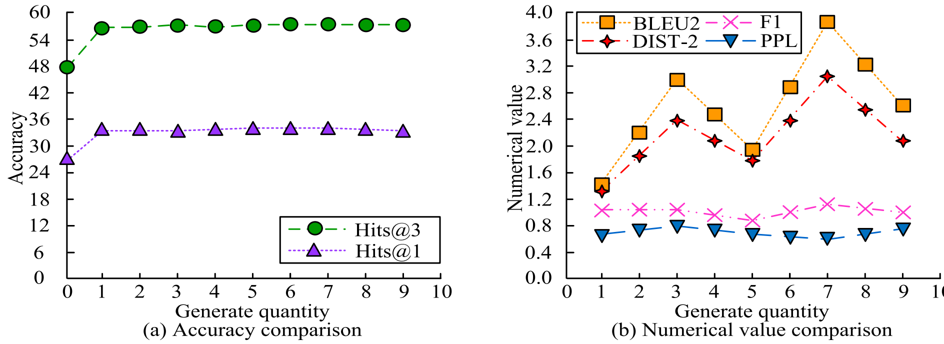


**FIGURE 12. Generate quantity comparison.**

**TABLE 3. Comparison of user experience scores for different models.**

| System type | Seq2Seq | S2SA | KG-S2SA | KG-IGD |
|---|---|---|---|---|
| Safety | 0.20 | 0.32 | 0.46 | 0.80 |
| Correlation | 0.83 | 0.92 | 1.13 | 1.44 |
| Fluency | 0.71 | 0.89 | 1.13 | 1.56 |
| Diversity | 0.20 | 0.29 | 0.72 | 0.82 |

personalized recommendations. This indicated that this improved model incorporating an external knowledge base increased the richness of vocabulary and reference information. The study randomly selected 300 users aged 18-45 to test the design system. Table 3 shows the user experience ratings for generating replies.

In Table 3, the security, diversity, relevance, and fluency of this improved system were all at their highest values, indicating good performance in the test application. The security and diversity of this improved system were 0.80 and 0.82, respectively, indicating that in GDS, users paid more attention to privacy regarding Q&A content. Because the more dialogue content there is, the more personal information data the design user has. The correlation and fluency of this improved system were 1.44 and 1.56, respectively, indicating that this improved algorithm obviously improved performance and users were more satisfied with the increase in system efficiency.

## V. CONCLUSION

As a popular natural language processing task, dialogue systems have broad application prospects in real life. In order to ensure the diversity of responses generated by the dialogue system, research is being conducted on building an open domain generative system, combined with KG's improved generative dialogue algorithm to enhance the richness of dialogue content. Simultaneously selecting bidirectional GRU to improve the Seq2Seq model ensures higher fluency and accuracy. The results indicate that when the generation quantity is 0, the improved system's Hits@1 and Hits@3 Both showed significant improvements, with increases of 7.63% and 9.01% respectively. The F1 of the KG-IGD model is 37.20%, and the PPL is 17.62. Compared with the Seq2Seq model, the PPL of the KG-IGD model has decreased by 35.53%. The Knowledge P/R/F1 index of the KG-IGD model is the highest, with values of 0.408, 0.387, and 0.397, respectively. The DIST-2 and BLEU2 indicators of the KG-IGD model are 0.052 and 0.220, respectively. The response content generated by this improved algorithm was more diverse and met personalized recommendations, making it feasible and effective in GDS. The user's ratings for the security and diversity of the improved system are 0.80 and 0.82, respectively. Users pay more attention to the privacy protection of the system and the diversity of generated content. This indicates its applicability in practical intelligent dialogue scenarios, such as enhancing user interaction experience in smart home appliance control, personal health management, and life service robots. The proposed method has good scalability and stability, which is of great reference value for generative dialogue systems when dealing with large-scale datasets. This method generates rich dialogue content, which to some extent enhances the user interaction experience. However, the study does not provide a detailed analysis of dialogue generation applications in real-life scenarios. Future experiments can delve into data privacy issues related to smart home appliance control or user personal health management.

## REFERENCES

[1] P. P. Groumpos, "A critical historic overview of artificial intelligence: Issues, challenges, opportunities, and threats," *Artif. Intell. Appl.*, vol. 1, no. 4, pp. 197–213, Oct. 2023, doi: 10.47852/bonviewaia3202689.

[2] D Baidoo-Anu and L. O. Ansah, "Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning," *J. AI*, vol. 7, no. 1, pp. 52–62, Dec. 2023, doi: 10.61969/jai.1337500.

[3] F. Cui, Q. Cui, and Y. Song, "A survey on learning-based approaches for modeling and classification of human–machine dialog systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 4, pp. 1418–1432, Apr. 2021, doi: 10.1109/TNNLS.2020.2985588.

[4] T. Ali, J. Hussain, M. B. Amin, M. Hussain, U. Akhtar, W. A. Khan, S. Lee, B. H. Kang, M. Hussain, M. Afzal, H. W. Yu, U. U. Rehman, H.-S. Han, J. Y. Choi, and A. Jamshed, "The intelligent medical platform: A novel dialogue-based platform for health-care services," *Computer*, vol. 53, no. 2, pp. 35–45, Feb. 2020, doi: 10.1109/mc.2019.2924393.

[5] M. Firdaus, H. Chauhan, A. Ekbal, and P. Bhattacharyya, ''EmoSen: Generating sentiment and emotion controlled responses in a multimodal dialogue system,'' *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1555–1566, Jul. 2022, doi: 10.1109/TAFFC.2020.3015491.

[6] R. Zhang, Z. Wang, Z. Huang, L. Li, and M. Zheng, ''Predicting emotion reactions for human–computer conversation: A variational approach,'' *IEEE Trans. Hum.-Mach. Syst.*, vol. 51, no. 4, pp. 279–287, Aug. 2021, doi: 10.1109/THMS.2020.3044975.

[7] H. Fakhrurroja, C. Machbub, A. S. Prihatmanto, and A. Purwarianti, ''Multimodal fusion algorithm and reinforcement learning-based dialog system in human–machine interaction,'' *Int. J. Electr. Eng. Informat.*, vol. 12, no. 4, pp. 1016–1046, Dec. 2020, doi: 10.15676/ijeei.2020.12.4.19.

[8] F. J. Miandashti, M. Izadi, A. A. N. Shirehjini, and S. Shirmohammadi, ''An empirical approach to modeling user-system interaction conflicts in smart homes,'' *IEEE Trans. Hum.-Mach. Syst.*, vol. 50, no. 6, pp. 573–583, Dec. 2020, doi: 10.1109/THMS.2020.3017784.

[9] J. Li, S. Dong, E. K. Chiou, and J. Xu, ''Reciprocity and its neurological correlates in human-agent cooperation,'' *IEEE Trans. Hum.-Mach. Syst.*, vol. 50, no. 5, pp. 384–394, Oct. 2020, doi: 10.1109/THMS.2020.2992224.

[10] M. Gaur, K. Faldu, and A. Sheth, ''Semantics of the black-box: Can knowledge graphs help make deep learning systems more interpretable and explainable?'' *IEEE Internet Comput.*, vol. 25, no. 1, pp. 51–59, Jan. 2021, doi: 10.1109/MIC.2020.3031769.

[11] Y. Yuan and X. Cai, ''A human–machine interaction scheme based on background knowledge in 6G-enabled IoT environment,'' *IEEE Internet Things J.*, vol. 8, no. 20, pp. 15292–15302, Oct. 2021, doi: 10.1109/JIOT.2021.3050880.

[12] Y. Shen, M.-A. Kioumourtzoglou, H. Wu, P. Vokonas, A. Spiro, A. Navas-Acien, A. A. Baccarelli, and F. Gao, ''Cohort network: A knowledge graph toward data dissemination and knowledge-driven discovery for cohort studies,'' *Environ. Sci. Technol.*, vol. 57, no. 22, pp. 8236–8244, May 2023, doi: 10.1021/acs.est.2c08174.

[13] Z. Fang, Q. Zhang, S. Kok, L. Li, A. Wang, and S. Yang, ''Referent graph embedding model for name entity recognition of Chinese car reviews,'' *Knowl.-Based Syst.*, vol. 233, Dec. 2021, Art. no. 107558, doi: 10.1016/j.knosys.2021.107558.

[14] H. Shelar, G. Kaur, N. Heda, and P. Agrawal, ''Named entity recognition approaches and their comparison for custom NER model,'' *Sci. Technol. Libraries*, vol. 39, no. 3, pp. 324–337, May 2020, doi: 10.1080/0194262x.2020.1759479.

[15] J. Li, B. Chiu, S. Feng, and H. Wang, ''Few-shot named entity recognition via meta-learning,'' *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 9, pp. 4245–4256, Sep. 2022, doi: 10.1109/TKDE.2020.3038670.

[16] R. Yan, J. Li, and Z. Yu, ''Deep learning for dialogue systems: Chit-chat and beyond,'' *Found. Trends Inf. Retr.*, vol. 15, no. 5, pp. 417–589, 2022, doi: 10.1561/1500000083.

[17] Z. Zhang, R. Takanobu, Q. Zhu, M. Huang, and X. Zhu, ''Recent advances and challenges in task-oriented dialog systems,'' *Sci. China Technol. Sci.*, vol. 63, no. 10, pp. 2011–2027, Sep. 2020, doi: 10.1007/s11431-020-1692-3.

[18] L. Wu, Y. Chen, K. Shen, X. Guo, H. Gao, S. Li, J. Pei, and B. Long, ''Graph neural networks for natural language processing: A survey,'' *Found. Trends Mach. Learn.*, vol. 16, no. 2, pp. 119–328, 2023, doi: 10.1561/2200000096.

[19] J. Li, A. Sun, J. Han, and C. Li, ''A survey on deep learning for named entity recognition,'' *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 1, pp. 50–70, Jan. 2022, doi: 10.1109/TKDE.2020.2981314.

[20] M. Boussakssou, H. Ezzikouri, and M. Erritali, ''Chatbot in Arabic language using seq to seq model,'' *Multimedia Tools Appl.*, vol. 81, no. 2, pp. 2859–2871, Jan. 2022, doi: 10.1007/s11042-021-11709-y.

[21] L. Weber, M. Sänger, J. Münchmeyer, M. Habibi, U. Leser, and A. Akbik, ''HunFlair: An easy-to-use tool for state-of-the-art biomedical named entity recognition,'' *Bioinformatics*, vol. 37, no. 17, pp. 2792–2794, Sep. 2021, doi: 10.1093/bioinformatics/btab042.

[22] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, ''A survey on knowledge graphs: Representation, acquisition, and applications,'' *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 2, pp. 494–514, Feb. 2022, doi: 10.1109/TNNLS.2021.3070843.

[23] S. Mao and E. Sejdic, ''A review of recurrent neural network-based methods in computational physiology,'' *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6983–7003, Feb. 2022, doi: 10.1109/TNNLS.2022.3145365.

[24] M. Nardo, D. Forino, and T. Murino, ''The evolution of man–machine interaction: The role of human in industry 4.0 paradigm,'' *Prod. Manuf. Res.*, vol. 8, no. 1, pp. 20–34, Jan. 2020, doi: 10.1080/21693277.2020.1737592.

[25] Z. Ma, R. Ni, and K. Yu, ''Recent advances, key techniques and future challenges of knowledge graph,'' *Chin. J. Eng.*, vol. 42, no. 10, pp. 1254–1266, Oct. 2020, doi: 10.13374/j.issn2095-9389.2020.02.28.001.

[26] M. Ren, N. Chen, and H. Qiu, ''Human–machine collaborative decision-making: An evolutionary roadmap based on cognitive intelligence,'' *Int. J. Social Robot.*, vol. 15, no. 7, pp. 1101–1114, Jun. 2023, doi: 10.1007/s12369-023-01020-1.

[27] B. Xue and L. Zou, ''Knowledge graph quality management: A comprehensive survey,'' *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 5, pp. 4969–4988, May 2023, doi: 10.1109/TKDE.2022.3150080.

[28] F.-Y. Wang, J. Guo, G. Bu, and J. J. Zhang, ''Mutually trustworthy human–machine knowledge automation and hybrid augmented intelligence: Mechanisms and applications of cognition, management, and control for complex systems,'' *Frontiers Inform. Technol. Electron. Eng.*, vol. 23, no. 8, pp. 1142–1157, Jun. 2022, doi: 10.1631/FITEE.2100418.

[29] J. Ni, T. Young, V. Pandelea, F. Xue, and E. Cambria, ''Recent advances in deep learning based dialogue systems: A systematic survey,'' *Artif. Intell. Rev.*, vol. 56, no. 4, pp. 3055–3155, Apr. 2023, doi: 10.1007/s10462-022-10248-8.

[30] Y.-J. Tang, Y.-H. Pang, and B. Liu, ''IDP-Seq2Seq: Identification of intrinsically disordered regions based on sequence to sequence learning,'' *Bioinformatics*, vol. 36, no. 21, pp. 5177–5186, Jan. 2021, doi: 10.1093/bioinformatics/btaa667.

**YIKE GUO** was born in Chongqing, Han Nationality, China, in July 1995. She received the B.A. degree from Chongqing Normal University, in 2017, and the M.F.A. degree from Sichuan University of Science and Engineering, in 2021. She is currently pursuing the Ph.D. degree in visual art and design with Suan Sunandha Rajabhat University.

Since 2021, she has been a Lecturer with the Design College, Chongqing Industry Polytechnic College. She has published ten academic articles, two professional textbooks, conducted five research projects, and holds three patents.