Knowledge-Based Dialogue System for the Ageing Support on Daily Activities

Julio Vizcarra(✉) and Kristiina Jokinen

National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan
{julio.vizcarra,kristiina.jokinen}@aist.go.jp

Abstract. With the increasing digitalization of society, we need to use a wide range of digitalized services in our daily activities such as searching for events in a calendar, checking the weather forecast, receiving guidance for completing certain tasks or recommendations for certain topics. Assistance for digital services is often needed, and particularly in the ageing stages, support for these tasks from a coach can become valuable. We introduce our work on a dialogue system that is part of a digital coach providing interactive support for elder adults in their daily activities. The work centers on using knowledge graphs to improve coaching interventions and is part of a larger project that focuses on supporting elder people and their healthy active living. Knowledge graphs are models of the domain content, defined by the domain experts, and they are used in the dialogue system to understand the content of the user utterances and to generate appropriate system responses. The dialogue coach can thus personalize conversations with the elder users and provide empathic and informative responses.

Keywords: Dialogue system · Knowledge graph · Machine learning · Natural language processing

1 Introduction

In daily life, with the increasing digitalization of society, we need to use a wide range of digitalized services such as searching for events in a calendar, checking the weather forecast, receiving guidance for completing certain tasks or recommendations for certain topics. Assistance for digital services is often needed, and particularly in the ageing stages, support for these tasks from a coach can become valuable. A digital coach can provide the needed help by enabling natural interaction and thus support elder adults in their daily activities. By analysing the conversations, it is also possible to learn more about the user concerns and needs, and offer more personalized support.

However, current conversational AI systems often ignore the semantic knowledge underlying the utterance content, or they use domain knowledge hardcoded in the system's conversational flow. Consequently, porting of the system to new task-specific domains or customizing dialogue interactions to different users can be difficult. If the expert knowledge can be modelled for the processing of the conversations, it is possible

Q. Gao and J. Zhou (Eds.): HCII 2022, LNCS 13331, pp. 122–133, 2022.
https://doi.org/10.1007/978-3-031-05654-3_8

to improve the system's capability to understand what the user needs, which can result in better assistance, recommendations, and flow in the conversations in general.

Knowledge Graph (KG) [1] has become an effective mechanism of representation that provides explainability, explicit definition of entities and relations, semantic interpretation, and it can contain fine-grained descriptions of the domain. Big advances in AI, NLP, and database technologies have enabled development of knowledge graphs and databases on multiple domains [2]. We have especially explored knowledge graphs as labeled property graphs (LPG) in the Neo4j graph database system [3].

In this paper, we propose a knowledge-based approach that tackles the acquisition of expert knowledge and its integration in KG, to be used in dialogue processing. The work centers the work on a semantic approach and exploitation of large knowledge bases, and their integration in conversational AI aiming to provide coaching for the users in daily living tasks. In the implementation, we can personalize the conversation considering the domain knowledge located in knowledge graphs related to the user profiles as well as the content in the sentences. We use the Rasa Open-Source Conversational AI framework [4] for dialogue modelling and present a dialogue system that supports elderly people in several tasks related to their daily living.

The novelty of the presented work is the conceptual-semantic handling of user utterances and conversations using knowledge graphs. An advantage of our approach is that the design of dialogue systems allows different domains to be included as knowledge graphs, which also adds on the flexibility of the system over hardcoded dialogue strategies by updates of the knowledge base.

The paper is structured as follows. We first give a short overview of the background and related research in Sect. 2. We then briefly describe the system and use cases in Sect. 3, while Sect. 4 focuses on knowledge graph and dialogue modelling, and the integration of the two in the dialogue system. Section 5 provides examples of the system function and finally Sect. 6 provides conclusions and future work.

2 Related Work

Many significant improvements have been achieved recently in dialogue systems using knowledge graphs. The inclusion of prior knowledge in the conversation has led to a better user response and interaction. The most promising use of KG has been in language generation. For instance, Yi-Lin Tuan et al. [5] implemented dynamic knowledge graphs in knowledge-grounded dialogue generation. The goal was to facilitate neural conversation models to learn zero-shot adaptation to updated, unseen knowledge graphs, by exploiting user data related to dialogues, speakers, and scenes. For this task, they presented a TV series conversation corpus (DyKgChat) with facts of the fictitious life of characters, and corresponding knowledge graphs including explicit information such as the relations *friend-of*, *enemy-of*, and *residence-of* as well as the linked entities.

Similarly, Houyu Zhang et al. [6] presented a conversation generation model ConceptFlow (Conversation generation with Concept Flow), that leverages commonsense knowledge graphs to explicitly model conversation flows. The model could generate semantically appropriate and informative responses. while using graph attention mechanism with few parameters. The work experimented with Reddit conversations.

In spoken dialogue understanding, Yi Ma et al. [7] proposed an inference knowledge graph that used semantic knowledge graphs and remapping using Markov Random Fields to create user goal tracking models that could form part of a spoken dialogue system. In their experiments, the authors demonstrated that their model could return more relevant entities to the user than the database lookup baseline. In dialogue generation, Sixing Wu et al. [8] proposed a method that uses a knowledge graph to alleviate the issue of generating boring responses. The commonsense knowledge-aware dialogue generation model (ConKADI) focuses on the knowledge facts that are highly relevant to the context. The main contribution was the use of knowledge graphs in learning models to add context and generate dialogues.

Tackling conversational context, Jaehun Jung et al. [9] proposed a dialogue-conditioned path traversal model called AttnIO that uses knowledge graphs based on attention flows. The authors claimed that they implemented a full use of structural information in the KG. Their method is capable of exploring a broad range of multi-hop knowledge paths, and also flexibly adjusting the varying range of plausible nodes and edges to attend depending on the dialogue context. Moreover, the system can intuitively model knowledge exploration depending on the dialogue characteristics.

In the area of human-robot interaction, Wilcock and Jokinen [10] describe an integration of Neo4j KGs with the Rasa dialogue framework and demonstrate it on the Furhat robot [11]. The generic framework aims to support a variety of social robots to provide high-quality information to users by accessing semantically rich knowledge about multiple different domains.

3 System Overview and Use Cases

We focus on using the knowledge graph to enable a coaching dialogue system that could provide useful information and empathic responses to the user. The research is conducted in the context of the project e-VITA (https://www.e-vita.coach/) a large collaboration project between EU and Japan, which aims to develop a virtual coaching system to support older adult's wellbeing in a smart living environment, promote their active healthy living, and provide personalized recommendations and trustworthy interaction (see background and more details in [12]). Some inspiration is found in Diogo Martinho et al. [13] who proposed a conceptual approach for improving the well-being of elderly people which used social networks to analyze user preferences, detected affective states, and evaluated cognitive capabilities. Similarly, Paulo Menezes et al. [14] proposed a multi-agent system designed to promote physical activity in elderly users.

The general overview of the system is depicted in Fig. 1. In the diagram, our contribution is marked in green dashes. The red dash box represents the Rasa [4] conversational AI framework. The process starts when the user shares a sentence and the module of natural language understanding (NLU) tries to interpret the user intention and extract relevant entities. The dialogue management module (DM) interacts with the action server to interpret the user input with respect to domain knowledge, provided by the domain experts and stored in the knowledge graphs. This process is referred to as "semantic processing" in Fig. 1, and it is carried out by computing queries on a knowledge graph. The result of the KG query is used as the basis of the system response, and the DM module enables interaction by generating a corresponding natural language sentence and sending it to the user.

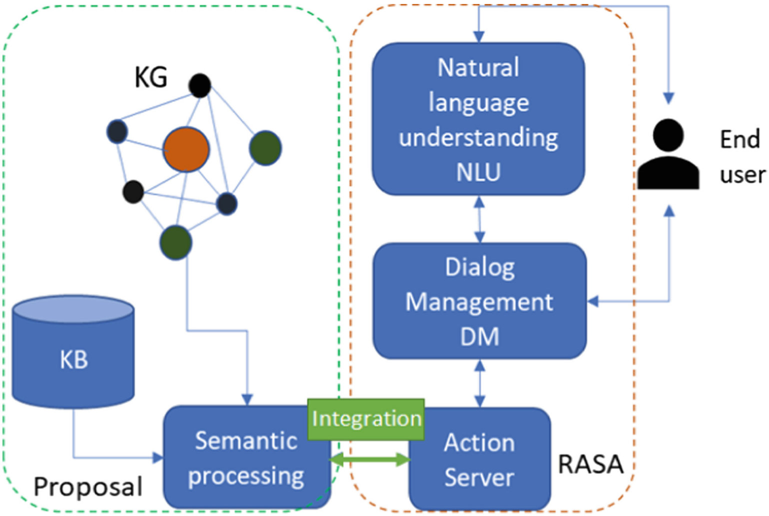

Fig. 1. Overview of the system.

3.1 Examples of Use Cases and Dialogue Design

The full e-VITA coach is envisaged to provide support in a wide variety of use cases ranging from exercise recommendations to daily reminders, providing information on news and weather reports, and emotional support in stressful situations. We have chosen two cases to exemplify the system functionality: the user experiences certain emotion and the user asks for information about exercises. The dialogues were designed in collaboration with the domain experts.

In the first case, the system is expected to take the user's affective state into account when providing its responses. For instance, if the user expresses sadness, the coach aims to provide empathic and consoling companionship. The action is triggered if the user explicitly informs about their emotional state, or if the system infers such a state from the user's voice. In the current version of the system, the emotion recognition is not fully integrated, so the example focuses on conversations where the users explicit express certain sentiment in their verbal utterances. Recognizing the user's intent to convey their mood, the system then responds by providing an appropriate feedback or suggestion that accords with the user's feelings. In the example case, the user's sad mood causes the system to accompany its message by a song that is assumed to improve the mood (Sakura-Sakura is a popular song in Japan), while in a happy case, the system joins the user's experience by providing an encouraging message and a happy song.

Table 1 shows an example dialogue when the user expresses sadness. The coach asks the reason and provides a song that may help. Finally, the coach confirms if the song improved the user's mood, and two paths, depending on the user response, are shown.

In the second case, the system enables factual information providing dialogues, and provides the user with relevant information. For instance, if the system recognizes the user's intention to find out more about exercises and physical training, the system will tell about the exercises and training programs that are stored in its knowledge graphs.

Since the knowledge graphs also allow reasoning about the types and properties of stored knowledge, the system can provide recommendations according to user specifications and available information. If the intent of an elderly user is recognized to solicit a recommendation of physical exercises as a part of their daily routine, the system provide examples of the available training possibilities, and the user can ask for further details or filter suitable particular exercises by their difficulty level.

Table 2 shows the system recommending exercises for the user. The left column shows how the dialogue system lists some training options stored in the knowledge graph, while the right column shows how the user can specify the exercises according to their difficulty.

Table 1. Sample dialogue for empathic responses.

Actor	Message	
User	I feel sad	
Coach	are you OK?. what is the reason you are sad ?	
User	I feel lonely	
Coach	You look sad, would you like me to put on your favorite music?	
User	ok	
Coach	I'll go to play a song that can make you to feel better. Please let me know when to stop the music. [plays the song Sakura Sakura]	
User	please stop	
Coach	Do you feel a bit better now?	
User	Yes	No
Coach	I wish you keep feeling alright ... Is there anything else I can help you with?	I am sorry, I hope I could help you. Is there anything else I can help you with?

4 Modeling

4.1 Knowledge Graph Models

Figure 2 presents a knowledge graph for the actions that the system can perform when it recognizes that the user experiences a certain emotion. The emotion is linked to a particular countermeasure, which can have a property describing recommendation as text or URL. Similarly, Fig. 3 presents an excerpt of a knowledge graph that represents

Table 2. Dialogues for exercises recommendations.

	Story listing all exercises	Story filtering exercises by difficulty
Actor	Message	Message
User	can you recommend me an exercise	can you recommend me an easy exercise
Coach	I know three exercises. According to my knowledge, you can practice Walking, Dancing, or Yoga.	Sure, an exercise with easy difficulty is: Walking

exercises, their level of difficulty, recommended poses, and other indicators that a user can perform. The knowledge graphs are modelled as Neo4j type labeled property graphs [3] and they can be queried using the Cypher query language [19].

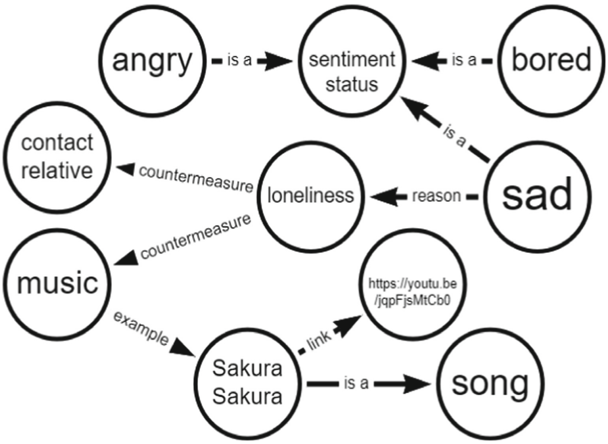

Fig. 2. Excerpt of a LPG representing emotional states and possible countermeasures.

4.2 Dialogue Model

The Rasa-based dialogue modelling [4] requires that example utterances representing the user intents are designed, together with the suitable system responses. Also sample dialogue flows, called stories can be defined. These are used in the training of Rasa's dialogue policy to decide on the suitable dialogue action after the recognition of the user intent.

Table 3 shows some user sentences corresponding to the user intent related to asking for exercises. The examples sentences are also labeled with respect to the important entities that describe the content of the intents. For instance, the word "easy" is labeled as a type of entity "difficulty". Labels are used for Name Entity Extraction (NER).

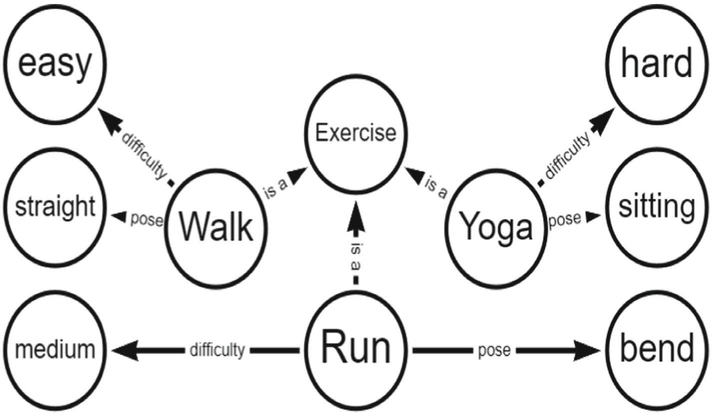

Fig. 3. Excerpt of a LPG representing physical exercises and their properties.

Table 3. Intent classification and entities.

intent: exercise_knowledge_base
-what exercise can you recommend with [low](intensity) intensity? -can you name some exercise please? -can you tell me a [easy](difficulty) exercise?

Table 4. Story for listing exercises.

story: Exercise KG list exercise	story: Exercise KG using difficulty
steps: - intent: exercise_knowledge_base - action: action_exercise_KB	steps: - intent: exercise_knowledge_base entities: - difficulty: hard - slot_was_set: - difficulty: hard - action: action_exercise_KB_difficulty

The story that the dialogue can follow deals with the 'happy path' case where the system successfully lists all the exercises (Table 4 left column) or the case when the user wants exercises filtered by a level of difficulty (Table 4 right column). The story is defined as a dialogue flow of user intents and system actions, and it also specifies the entities required to complete the path. The system actions are the so called knowledgebaseActions, which retrieve data from the KG referred to by its name.

4.3 Knowledge Graph Integration in the Dialogue System

The overall view of the dialogue system modules is presented in Fig. 4. The modules are based on the Rasa architecture and modified according to the project needs. The user interacts with the coach through the coaching device (in our case Nao robot [21]). The user's utterances will be analyzed by the speech recognizer (not shown in Fig. 4), and the text transcription will be the input to Natural Language Understanding (NLU) pipeline which consists of a series of modules that are needed for the preprocessing, entity extraction, and intent classification, and transforms the user input into a vector format. The user input then goes through the "knowledge processing" where the content is analysed with respect to the system's domain knowledge, and semantic inferences are performed. The important part concerns querying the knowledgebase, which in our case uses Cypher queries based on the extracted entities in the graph database. The query results are inserted in the suitable system response form, which is determined by the dialogue model using the learnt state tracking policy. The text sentence thus augmented is sent further to the text-to-speech synthesizer by the dialogue manager, to be ultimately uttered by the coaching agent to the user.

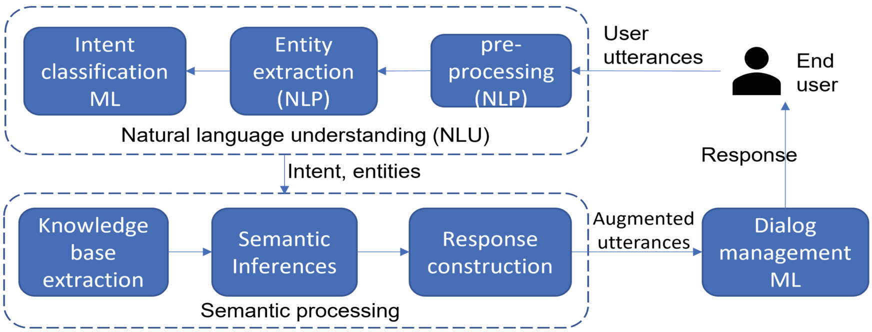

Fig. 4. Content-based processing of user utterances.

We use the Rasa Open-Source Conversational AI [4] to enable the above dialogue processing. We customised Rasa's NLU pipeline by the spaCy [15] tokenizer, entity extractor (NER), and featurizer, together with other featurizers available in Rasa. To extract the intent and relevant entities from the user message, Rasa's Dual Intent and Entity Transformer (DIET) [16] is used. Table 5 shows the NLU pipeline.

To determine the system response, the dialogue manager predicts the best next action using a dialogue policy. We used the Transformer Embedding Dialogue (TED) policy implemented in Rasa [17] (Table 6). The max history was set to 3. We experimented with different context lengths for the policy performance, and the length 3 produced the best performance for our example conversations.

Table 5. The NLU Pipeline used in the study.

pipeline:
- name: SpacyNLP model: "en_core_web_md" - name: SpacyTokenizer - name: "SpacyEntityExtractor" - name: SpacyFeaturizer - name: RegexEntityExtractor - name: LexicalSyntacticFeaturizer - name: CountVectorsFeaturizer - name: DIETClassifier

Table 6. Rasa's dialogue policy for best next action in the study.

policies:
- name: TEDPolicy max_history: 3

The domain knowledge is stored in a Neo4j graph database [18] and retrieved using Cypher querying language [19]. As a simple example, Table 7 shows a Cypher query that retrieves all hard exercises which are of type guideline. The MATCH command searches for nodes which are Exercises such that they have a link to a node Difficulty with the title 'hard', and moreover, have type 'guide-line'.

Table 7. Example Cypher query.

Query that retrieves hard exercises
MATCH (e:Exercise) WHERE exists ((e)-[*..3]->(:Difficulty{title: 'hard'})) and e.type = 'guide_line' RETURN e.title

5 Rasa X Examples

Some examples of the interaction between a user and the e-VITA coach are shown in the screenshots below. The dialogues are conducted through text on a web-based extension Rasa X [20].

Figure 5 depicts a dialogue where the system gives recommendations to the user according to a level of difficulty. Figure 6 presents the case when the user experiences sadness, and receives consolation with a song.

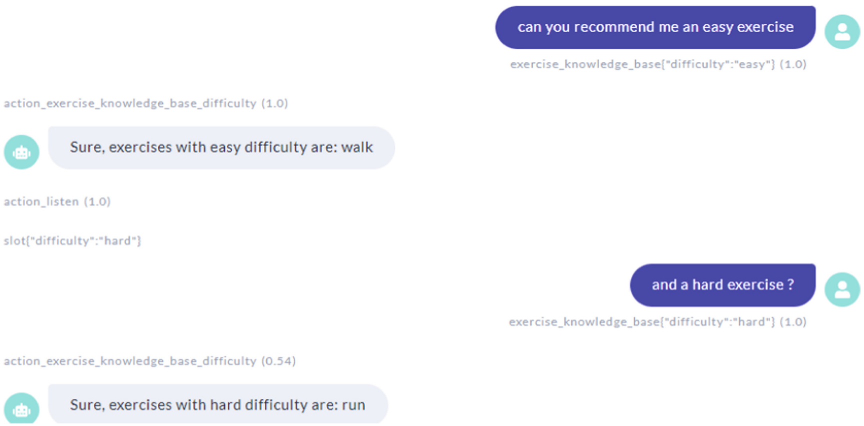

Fig. 5. A dialogue that recommends an exercise by difficulty.

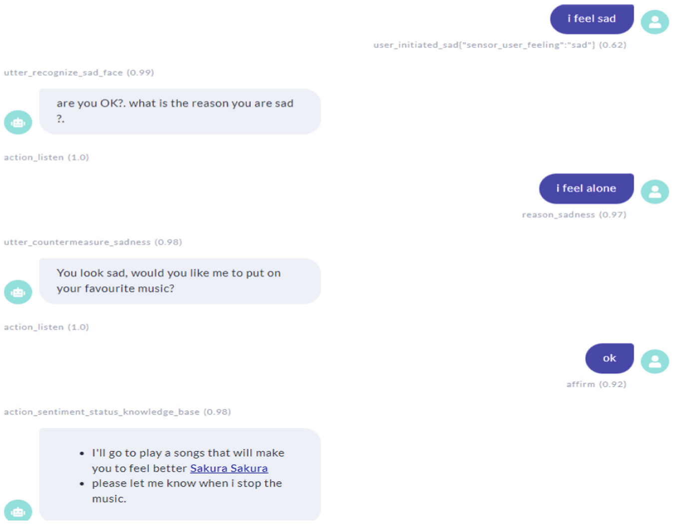

Fig. 6. A fragment of conversation that alleviates sadness.

6 Conclusions and Future Work

In this paper, we discussed a knowledgebase approach that integrates domain knowledge stored in knowledge graphs into system responses and actions. The work constructs knowledge graphs that describe the domain and use cases selected for the e-VITA project which aims to develop a coach to support elder people's active healthy living and daily activities. The system integrates content from the KG to process user utterances and to provide suitable recommendations and information to the user based on the domain knowledge, and the dialogues can be conducted using a web-based interface or by interacting with an embodied coach, the NAO robot. The e-VITA coach aims to cover several use cases, and we exemplified two example scenarios in this paper: first, the coach tries to provide a sympathetic response when it recognizes the user experiencing an emotional state, and the second one, where the coach provides information and recommendation to the user when it recognized the user wants to receive particular training recommendations.

The future work deals with sophisticating and extending knowledge graphs and their content with respect to the demands for natural and trustworthy dialogues. In particular, we will study solutions related to graph embedding to integrate knowledge graph reasoning into dialogue management. Important aspects also concern dialogue strategies and dialogue management so as to support friendly, contextually appropriate dialogues between the user and the coach. Subtle issues in emotional and empathic interaction need to be studied and discussed further for appropriate interaction design, and in order to take cultural and individual differences into account. Moreover, integration of several other domains in the system is important in order to increase the system functionality. We will also closely look at the technical of knowledge base construction and integration of heterogenous sensor data in the system, given availability of various environmental sensors in the smart home environment. Finally, ethical and privacy issues are important, and they will be studied and taken into account in the project as a whole.

Extensive user studies in living lab condition are planned to take place in the project context, with the dialogue system being integrated into various interface agents. Such evaluations have been problematic due to the Covid pandemic, but preliminary experiments with the first system version could be conducted, resulting in useful feedback on the dialogue strategies and technical implementation, showing that the system has potential to be developed into a useful application.

Acknowledgments. We thank our project partners in Japan and EU for discussions. This work was supported by the Strategic Information and Communications R&D Promotion Programme (SCOPE) of Ministry of Internal Affairs and Communications (MIC), Grant no. JPJ000595.

References

1. Hogan, A., et al.: Knowledge graphs. Synthesis Lectures on Data, Semantics, and Knowledge, vol. 12, pp. 1–257. Morgan & Claypool Publishers (2021)
2. Ji, S., Pan, S., Cambria, E., Marttinen, P., Philip, S.Y.: A survey on knowledge graphs: representation, acquisition, and applications. IEEE Trans. Neural Networks Learn. Syst. IEEE **33**(2), 494–514 (2021)
3. Robinson, I., Webber, J., Eifrem, E.: Graph DataBases, 2nd edn. O'Reilly Media (2015)
4. Bocklisch, T., Faulkner, J., Pawlowski, N., Nichol, A.: Rasa: open source language understanding and dialogue management. arXiv:1712.05181 (2017)
5. Tuan, Y.-L., Chen, Y.-N., Lee, H.-Y.: DyKgChat: benchmarking dialogue generation grounding on dynamic knowledge graphs. arXiv:1910.00610 (2019)

6. Zhang, H., Liu, Z., Xiong, C., Liu, Z.: Grounded conversation generation as guided traverses in commonsense knowledge graphs. arXiv:1911.02707 (2019)
7. Ma, Y., Crook, P.A., Sarikaya, R., Fosler-Lussier, E.: Knowledge graph inference for spoken dialog systems. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5346–5350 (2015)
8. Wu, S., Li, Y., Zhang, D., Zhou, Y., Wu, Z.: Diverse and informative dialogue generation with context-specific commonsense knowledge awareness. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 5811–5820 (2020)
9. Jung, J., Son, B., Lyu, S.: Attnio: knowledge graph exploration with in-and-out attention flow for knowledge-grounded dialogue. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 3484–3497 (2020)
10. Wilcock, G., Jokinen, K.: Conversational AI and knowledge graphs for social robot interaction. late-breaking reports. In: The 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI-2022) (2022)
11. Furhat Robotics Homepage. https://furhatrobotics.com//. Accessed 07 Feb 2022
12. Katsutoshi, Y., et al. (eds.): JSAI 2020. AISC, vol. 1357. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-73113-7
13. Martinho, D., Carneiro, J., Novais, P., Neves, J., Corchado, J., Marreiros, G.: A conceptual approach to enhance the well-being of elderly people. In: Moura Oliveira, P., Novais, P., Reis, L. P. (eds.) EPIA 2019. LNCS (LNAI), vol. 11805, pp. 50–61. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-30244-3_5
14. Menezes, P., Rocha, R.P.: Promotion of active ageing through interactive artificial agents in a smart environment. SN Appl. Sci. **3**(5), 1–15 (2021). https://doi.org/10.1007/s42452-021-04567-8
15. spaCy Homepage. https://spacy.io/. Accessed 04 Feb 2022
16. Bunk, T., Varshneya, D., Vlasov, V., Nichol, A.: Diet: lightweight language understanding for dialogue systems. arXiv:2004.09936 (2020)
17. Vlasov, V., Mosig, J. E., Nichol, A.: Dialogue transformers. arXiv:1910.00486 (2019)
18. Neo4j Homepage. https://neo4j.com/. Accessed 7 Feb 2022
19. Francis, N., et al.: Cypher: an evolving query language for property graphs. In: Proceedings of the 2018 International Conference on Management of Data, pp. 1433–1445 (2018)
20. RASA X Homepage. https://rasa.com/docs/rasa-x/. Accessed 7 Feb 2022
21. NAO Homepage. https://www.softbankrobotics.com/emea/en/nao. Accessed 7 Feb 2022