# Knowledge-centered conversational agents with a drive to learn

**Selene Báez Santamaría**
Vrije Universiteit Amsterdam
s.baezsantamaria@vu.nl

## Abstract

We create an adaptive conversational agent that assesses the quality of its knowledge and is driven to become more knowledgeable. Unlike agents with predefined tasks, ours can leverage people as diverse sources to meet its knowledge needs. We test the agent in social contexts, where personal and subjective information can be obtained through dialogue. We provide the agent both with generic methods for assessing its knowledge quality (e.g. correctness, completeness, redundancy, interconnectedness, and diversity), as well as with generic capabilities to improve its knowledge by leveraging external sources. We demonstrate that the agent can learn effective policies to acquire the knowledge needed by assessing the efficiency of these capabilities during interaction. Our framework enables on-the-fly learning, offering a dynamic and adaptive approach to shaping conversational interactions.

## 1 Introduction

Machines were initially designed as tools to help people with heavy or repetitive tasks. Over time, machines have become more advanced to the point where a proportion can now perform tasks independently. This challenges the societal perception of machines as passive tools and shifts it to consider them active participants performing the task in collaboration with people (Durante et al., 2024; Deng et al., 2023a). Within a Hybrid Intelligence framework (Akata et al., 2020), people and machine may collaborate as part of a team that is more effective than each individually.

With increased autonomy within such collaborative contexts, machines are more likely to encounter unforeseen and complex problems signalled by negative feedback, failure to make decisions, or unsuccessful actions (Kocoń et al., 2023). To tackle these issues, agents must have the ability to identify problems and evaluate knowledge conditions like missing information, uncertainty, misunderstandings and conflicts. Addressing unexpected problems requires adaptability, in the form of leveraging sources of knowledge and information effectively to resolve them.

We therefore propose the concept of *generic* and *knowledge-centered* agents that 1) can estimate the quality of their current knowledge, in terms of how sufficient it is to service certain needs and 2) have the capacity to actively consult sources of knowledge to become more *knowledgeable*. Unlike traditional task-oriented dialogue (TOD) agents (comparison shown in Figure 1), knowledgeable agents can autonomously determine what they know and do not know, what is the epistemic status of what they know, what they need to learn, and how to acquire that target knowledge.

This thesis proposal focuses on dialogue as the way an agent learns and adapts in social context. A flexible learning agent is able to acquire and modify current knowledge through natural language instead of solely relying on (structured) data. This adaptability allows the agent to navigate a dynamic knowledge landscape, making open-domain communication a versatile and practical approach. Focusing on machines as social agents capable of qualifying knowledge shared during human-machine conversations, we position this research in the broader context of conversational agents and specifically within the HI framework in which agents and people are expected to form teams.

## 2 Related work

**Dialogue systems** Task-oriented dialogue (TOD) systems (represented as service dialogue systems in Figure 1, left) are designed for specific service tasks, relying on supervised training with user input, dialogue state, and context such as history or user profiles (Mesnil et al., 2014; Mensio et al., 2018; Zhang et al., 2019). Reinforcement learn-



*Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 4: Student Research Workshop)*, pages 83–92
June 18, 2024 

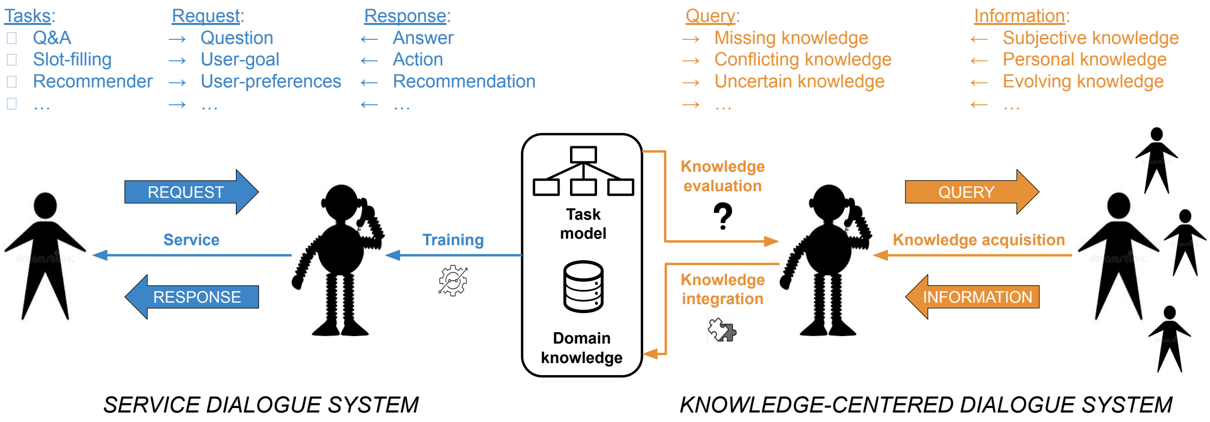


Figure 1: Comparison of dialogue systems: On the left, the conventional service dialogue system (S-DS) undergoes training on static data to subsequently service the needs of users. On the right, a *knowledge-centered* dialogue system (KC-DS) evaluates its knowledge base and actively participates in dialogue with users to acquire targeted knowledge, which is then integrated into its dynamic knowledge base.

ing (RL) further optimizes these agents for diverse and functionally correct responses through user feedback (Liu et al., 2017; Gao et al., 2018; Lippe et al., 2020). For open conversational agents, evaluating the dialogue state and formulating an adequate response to transit to the next, preferably better state, is more challenging (Shum et al., 2018). Furthermore, in open-domain dialogue (OOD) settings, conversational agents must also be equipped with various conversational skills like engagement, knowledge, and empathy to thrive in different social interactions and keep people engaged (Smith et al., 2020). Still, RL may improve the performance of systems but does not adapt the service it was designed for.

**Adaptive conversational agents** Several efforts have focused on making dialogue systems flexible to a broader range of use cases, focusing on different domains (Qian and Yu, 2019; Wen et al., 2016; Le et al., 2020; Qian and Yu, 2019), different tasks (Young et al., 2022; Chen et al., 2022; Deng et al., 2023c), or different users (Yang et al., 2021). However, adapting to entirely new tasks poses challenges, requiring generalizability or costly acquisition of domain knowledge. For instance, slot-filling actions lack adaptability in slot types or value ranges (Ni et al., 2023), and recommender systems are constrained by static knowledge (Liu et al., 2021b). Thus, the support that an agent can provide is inherently limited.
Making systems more adaptive has been generally studied with different techniques. Meta-learning is a data-driven approach that focuses on exposing models to various learning scenarios, so they can extract patterns that can be applied towards novel tasks (Hospedales et al., 2021). Never-ending learning focuses on developing systems that improve continuously as they encounter new data or tasks (Mitchell et al., 2018). While these acknowledge the shortcomings of static or limited knowledge, these techniques still rely on passive learning, where the system is exposed to certain situations instead of actively searching and prioritizing the learning of specific information. To conclude: these approaches do not determine a need to learn.

**Knowledge-grounded conversational agents** Knowledge-grounded conversational systems utilize knowledge sources for the retrieval of factual information. These sources can be unstructured texts or domain-specific triples (Xu et al., 2020). The dialogue task is conventionally modelled by taking a user utterance as input, selecting relevant knowledge items from a database, and verbalizing them in accordance with a dialogue history (Kim et al., 2023).
This process is typically unidirectional (Deng et al., 2023b), starting from the user (expressing a request) to the agent and then from the agent (providing static information) back to the user to satisfy the user need (left side of Figure 1). However, this unidirectional perspective neglects the reciprocal nature of information exchange. Agents can also find themselves in a state of uncertainty or lack of information, prompting a need to seek clarification or additional details from the user. In collaborative settings, a bidirectional flow is essential (adding the right side of Figure1), initiating from the agent's

need for information and extending to the user, who may provide the agent with new information.

As information flows from the user to the agent, understanding the status of the user (Liu et al., 2021a) and the conveyed knowledge is critical. While factual knowledge is well-handled, personal and opinion-based knowledge is more complex, gaining relevance in long-term interactions. Moreover, successful collaboration requires that agents learn to independently a) judge knowledge sources in terms of capability, expertise, and trust and b) judge specific perspectives in terms of diversity, bias and distribution within populations. Traditional knowledge-grounded agents often lack such an epistemic dimension in their representations.

Our research centres on developing a flexible agent capable of navigating uncertainties across various tasks. We consider adaptation within specific social and collaborative contexts, which requires real-time assessment of the agent's learning needs and human input. This entails adaptation for individual cases while remaining aware of the situational dependency of acquired knowledge when considering new scenarios. The primary objective is to enhance the agent's ability to acquire and process knowledge, extending beyond traditional factual knowledge to include social understanding.This expansion aligns with the evolving nature of human-machine interactions, where social dynamics play a vital role in fostering collaboration.

## 3 Knowledge-centered conversational agents

This thesis proposal tackles the research question: *"How can conversational agents be equipped to adapt in social collaborative settings by acknowledging and addressing knowledge limitations?"*. In the next subsections, we address three dimensions of focus and provide specific sub-questions, methodologies and preliminary results.

### 3.1 Knowledge integration

**Episodic memory for conversational agents** The role of memory in conversation is directly related to creating and retrieving shared memories. Beyond the social dimension of human-machine interaction, knowledge-centered agents benefit from having a memory since keeping track of their own knowledge also enables them to evaluate:

1. Its knowledge state: *What do I know?*
2. Its knowledge needs: *What do I need to know?*
3. Knowledge sources: *Who knows about this and can be trusted?*
4. Knowledge changes: *What things change, and which ones stay the same?*

The importance of memory poses the question of *"How can conversational agents be equipped with the ability to aggregate knowledge over time?"*. In our approach, we use graph technologies and the W3C web standard RDF[1] to model the knowledge that dialogue agents acquire through conversations. We design episodic Knowledge Graphs (eKG) to represent an agent's accumulated episodic experiences. Through this, we bridge gaps between disconnected individual interactions and model the cumulative knowledge of conversational agents across interactions (Báez Santamaría et al., 2021).

**Adaptability of knowledge** A significant limitation of current dialogue systems is that they follow the Closed World assumption (Hustadt et al., 1994), thus overly relying on the world model and current information they have and considering it static and complete. We challenge this and propose to follow an Open World assumption, where information not explicitly stated is considered unknown rather than false or out of scope. Furthermore, this outlook is better suited to address the concept of unknown unknowns, or simply put: "We don't know what we don't know".

Computationally representing this shift in assumptions brings the question of *"How to create a generic model of the world (T-Box) that can be adapted and extended during real-time interaction with a user?"*. As preliminary work, we create a social ontology that sufficiently covers essential concepts for human-machine interaction (e.g. a person's name, place of origin, occupation, interest) and thus enables basic communication for a KC agent. The agent, however, does not depend on this ontology to perform tasks or talk to a user but instead is able to extract information from what the user says and incorporate entities and their relations into the agent's knowledge base. Ideally, entities are typed, either by exploiting the interoperability with Linked Open Data (LOD) (Bauer and Kaltenböck, 2011) resources or by asking for further details from users in dialogue. In that case, these types are ingested as new classes of the on-

[1] Resource Description Framework: https://www.w3.org/RDF/

tology, while learned information is used to expand and enrich the class' description and object properties. This is a promising avenue to explore ontology learning through open-domain communication (Vossen et al., 2019a).

**Relativity of knowledge** Open-domain communication involves non-factual information like opinions, beliefs, and perspectives (Báez Santamaría et al., 2023) . To effectively perform in these scenarios, conversational agents must handle information from a social angle, recognizing the importance of acquiring diverse knowledge from different sources, each with its own biases and complementary views. Processing this type of information may involve reaching a consensus within a community, identifying areas of disagreement or diversity of perspectives, and recognizing that some perspectives are dynamic and evolve over time. This complexity mirrors human cognition, relying heavily on the Theory of Mind (ToM) (Wimmer and Perner, 1983), allowing the attribution of mental states to oneself and others for comprehending social interactions and implications.

The complexity of non-factual information brings forward the question of *"How to model and represent epistemic aspects of knowledge (A-Box) as a Theory of Mind?"*. For this, we choose to use the GRaSP (Fokkens et al., 2017) ontology to represent mentions and perspectives. `MENTIONS` differentiate between an `INSTANCE` in the world (e.g `Gabriela`), and a reference to it (e.g. `Gaby`, `the mother of Karla`, or `my aunt`). Each of these mentions is linked to a `SOURCE` and was expressed with a specific `ATTRIBUTION` that qualifies the information received according to the source's perspective (i.e. denial/confirmation, sentiment, emotion, and certainty). This approach enables the agent to represent social aspects of knowledge during human-machine interactions and also to reason over its epistemic status (Vossen et al., 2018, 2019c).

### 3.2 Knowledge evaluation

**Quality of knowledge** Beyond the accumulation of knowledge, it is important to evaluate the quality of the gathered knowledge. Specifically we want to quantitatively and qualitatively evaluate specific dimensions, such as correctness, completeness, redundancy, interconnectedness, and diversity

This results in the question of *"How can the quality of the accumulated knowledge be measured?"*. We propose to exploit the eKG representation to measure structural and semantic graph aspects at three levels: as a mathematical object, as an RDF knowledge representation object, and as an episodic memory. To test this multidimensional evaluation framework, we performed an exploratory analysis to search for correlations between these metrics and specific quality dimensions of the knowledge accumulated. We demonstrate that the framework can be used to evaluate any conversation, among which human-human, agent-human and agent-agent, by assessing the characteristics and the quality of the information and perspectives that are exchanged between the interlocutors. Furthermore, the eKG representation allows not only the evaluation of knowledge as a static object but also a comparison over time, thus assessing its potential improvements or deterioration (Báez Santamaría et al., 2022).

**Drives to improve knowledge** The previous section dealt with evaluating the knowledge gathered as a whole. While this is important, it might be more meaningful to identify specific areas of knowledge that are of low quality and might be crucial to improve. These areas, encompassing aspects like what is unknown, what is new, or what beliefs are uncertain, serve as the agent's specific objectives in relation to its current informational state. In scenarios where resources like time, energy, money, or accessibility to knowledge sources are limited, prioritizing targeted knowledge areas might lead to more promising avenues of improvement.

As such, this leads to the question of *"How can knowledge quality be related to specific knowledge drives?"*. As an approach we propose to exploit the intrinsic reasoning capabilities of RDF and OWL (McGuinness et al., 2004) to detect abstract graph patterns that may signal poor knowledge quality. We produce a set of eight[2] SPARQL (Harris and Seaborne, 21) queries that identify areas of the eKG where knowledge might be deficient or unreliable. These queries focus on gaps, analogies, conflicts, overlaps and novelty aspects of the accumulated knowledge.

Gaps and analogies are defined by the ontologies included, capturing what can be known, what is typical or what is expected. These aspects might behave similarly to slot-filling approaches and relate to a pre-defined world. In contrast, conflicts, overlaps, and novelties are determined by the stored

[2]The process by which these particular queries were created is generic and can produce additional drives.

information thus far and relate to specific epistemic aspects such as correctness, interconnectedness or redundancy. Each identified graph pattern can be translated into an agent's utterance, thus enabling the transition between specific knowledge states through dialogue (Vossen et al., 2019b).

### 3.3 Knowledge acquisition

**Instruments to improve knowledge** We have focused so far on the knowledge management aspect of dialogue, excluding the discussion of the capability to communicate through natural language. Both Natural Language Understanding (NLU) and Natural Language Generation (NLG) are crucial components in this regard. Given the chosen technologies, NLU requires implementing Information Extraction (IE) to transform natural language into RDF triples (Martinez-Rodriguez et al., 2020), while NLG verbalizes and summarizes knowledge subgraphs. Both of these are active areas of research on their own with considerable achievements.

Yet, in the specific context of this research, the question remains of *"How can social knowledge-centered agents be provided with the communicative skills to pursue their knowledge drives?"*. To answer this question, we developed specific triple extraction models to cover the large linguistic variation present in open-domain dialogue, emphasizing the extraction of perspective values such as polarity, certainty, sentiment, emotion, and temporality[3]. Similarly, we have invested effort into strengthening an agent's capability to express its knowledge state and drives transparently and concisely (Krause et al., 2023)[4].

As there is a strong dependency between the triple extractor tool employed and the eKg generated, we have explored various extraction approaches.In particular, we have implemented five triple extractors: 1. a tailored Context Free Grammar, 2. a spacy-based dependency parser, 3. an Open Information Extractor based on Standford's implementation (Angeli et al., 2015), 4. a fine–tuned multilingual BERT based model, and 5. a LLama3 prompting technique. It is important to note that the performance of these extractors impacts the graph's reasoning capabilities, as the granularity and meaningfulness of the nodes and relations will change. However, the impact on the graph-based comparative evaluations (either between agents or across time) is negligible, as the same biases of the tool are present across graphs.

**Strategies to improve knowledge** The evaluation framework established earlier produces an extensive repertoire of areas where knowledge can be targeted for improvement. Considering the specific setting of conversational agents, selecting one of these areas of knowledge will produce different conversations with a user, sometimes leading to valuable input, while others are less successful. Thus, the selection of the best or next area to focus on becomes an important one, potentially linked to dialogue management or planning.

The previous raises the question of *"How to learn effective strategies to exploit the communicative options of knowledge-centered agents for achieving their knowledge goals?"*. To explore this question, we experiment with RL to enable the agent to dynamically choose semantic patterns when responding to human cues. Preliminary evidence indicates that generic graph metrics as rewards elicit specific types of knowledge acquisition behaviour. For instance, metrics measuring the volume of knowledge, like `Total number of triples`, lead to an agent focused on addressing knowledge gaps, thus directly asking questions to the user around unknown values. Overall, this adaptive approach allows the agent to acquire knowledge through a conversation while also being flexible across different tasks, domains, and users.

### 3.4 Applications

Knowledge-centered agents can be applied to a wide range of conversational situations. Scenarios where non-factual or personal knowledge is predominant or where access to diverse knowledge sources is available could benefit the most. We focus on three specific domains: Diabetes Lifestyle Management (DLM) (de Boer et al., 2023), Reconstruction of Timelines and Personal Diaries and Counter-narrative creation for Hate Speech (HS) (Doğanç and Markov, 2023). In the context of DLM, the framework allows for the extraction of patient preferences to ensure a tailored and effective treatment plan (example on the Appendix, Figure 2) (Dudzik et al., 2024; Chen et al., 2024). For the Reconstruction of Timelines and Personal Diaries, attention is directed towards identifying temporal gaps between conversations to learn what happened since and what the user perspective

[3] https://github.com/leolani/cltl-knowledgeextraction

[4] https://github.com/leolani/cltl-languagegeneration

is (Vossen et al., 2024). Lastly, in the case of combating HS, the framework addresses the challenge of detecting reasoning faults and distinguishing differences of opinion from violations of modern values, such as the dehumanization of vulnerable groups. These cases have in common that an agent must actively get input from a user and critically evaluate the quality of the information received.

Finally, we demonstrate that the same framework can be deployed as a text-based chat system and as a multimodal robot. In either case, observations and experiences are captured in the eKG on which the agent can act using the proposed evaluative strategies to interact with its environment. This flexibility highlights our framework's potential to be integrated into various conversational modalities, offering a robust and adaptable solution across different interaction contexts (Baier et al., 2022).

## 4 Conclusion

We develop a framework for conversational agents designed to expand its knowledge for a better understanding of the world. The agent does not focus on servicing users in predefined tasks but instead focuses on knowledge that is lacking and needs to be acquired or verified from external sources.

Our approach is highly flexible, independent to any task-specific goals and capable of handling various dialogue domains without customization effort[5]. Our framework enables agents to modify task models on the fly and extend domain information, allowing for a dynamic and adaptive approach to shape conversational interactions.

### 4.1 Challenges

In real-world settings, the growth of these eKG can be rapid, thus presenting scalability challenges. We have identified two main challenges in particular. Firstly, querying these graphs in an efficient manner becomes crucial, demanding proper database management techniques such as ensuring correct indices and optimizing queries. Secondly, utilizing these graphs in neurosymbolic approaches involves storing these large graphs in memory for specific graph machine learning libraries, which can pose computational difficulties with very large graphs ($<$ 3 million triples).

To tackle these challenges, various strategies can be employed. One approach is to slice the graph over time, knowledge sources, or types of knowledge. These slices can be processed individually or stored separately. Another option is to summarize the graph (Čebirić et al., 2019), either through extractive or grouping methods, or to sample the graph (Hu and Lau, 2013) according to the needs of a given application. Overall, while none of these challenges are severe enough to make the proposed framework unfeasible, they do require careful planning and consideration.

[5] In certain cases, further training on the NLU/NLG modules might lead to improvements on the overall knowledge communication pipeline. However, this is not required.

## Acknowledgements

This research was funded by the Vrije Universiteit Amsterdam and the Netherlands Organisation for Scientific Research (NWO) via the *Spinoza* grant (SPI 63-260) awarded to Piek Vossen, and the *Hybrid Intelligence Centre* via the Zwaartekracht grant (024.004.022).

## References

Zeyneb Akata, Dan Balliet, Maarten de Rijke, Frank Dignum, Virginia Dignum, Guszti Eiben, Anstke Fokkens, Davide Grossi, Koen Hindriks, Holger Hoos, Hayley Hung, Catholijn Jonker, Christof Monz, Mark Neerincx, Frans Oliehoek, Henri Prakken, Stefan Schlobach, Linda van der Gaag, Frank van Harmelen, Herke van Hoof, Birna van Riemsdijk, Aimee van Wynsberghe, Rineke Verbrugge, Bart Verheij, Piek Vossen, and Max Welling. 2020. A research agenda for hybrid intelligence: Augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *IEEE Computer*, 53(08):18–28.

Gabor Angeli, Melvin Jose Johnson Premkumar, and Christopher D. Manning. 2015. Leveraging linguistic structure for open domain information extraction. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 344–354, Beijing, China. Association for Computational Linguistics.

Selene Báez Santamaría, Thomas Baier, Taewoon Kim, Lea Krause, Jaap Kruijt, and Piek Vossen. 2021. Emissor: A platform for capturing multimodal interactions as episodic memories and interpretations with situated scenario-based ontological references. In *1st Workshop on Multimodal Semantic Representations, MMSR 2021*, pages 56–77. Association for Computational Linguistics (ACL).

Selene Báez Santamaría, Lea Krause, Lucia Donatelli, and Piek Vossen. 2023. The role of personal perspectives in open-domain dialogue.

Selene Báez Santamaría, Piek Vossen, and Thomas Baier. 2022. Evaluating agent interactions through episodic knowledge graphs. In *Proceedings of the 1st Workshop on Customized Chat Grounding Persona and Knowledge@ COLING2022*, pages 15–28. Association for Computational Linguistics (ACL).

Thomas Baier, Selene Báez Santamaría, and Piek Vossen. 2022. A modular architecture for creating multimodal agents. *Computing Research Repository-arXiv*, 2206.

Florian Bauer and Martin Kaltenböck. 2011. Linked open data: The essentials. *Edition mono/monochrom, Vienna*, 710:21.

Šejla Čebirić, François Goasdoué, Haridimos Kondylakis, Dimitris Kotzinos, Ioana Manolescu, Georgia Troullinou, and Mussab Zneika. 2019. Summarizing semantic graphs: a survey. *The VLDB journal*, 28:295–327.

Pei-Yu Chen, Selene Báez Santamaría, Maaike H. T. de Boer, Floris de Hengst, Bart A. Kamphorst, Quirine Smit, Shihan Wang, and Johanna Wolff. 2024. Intelligent support systems for lifestyle change: Integrating dialogue, information extraction, and reasoning. In *HHAI 2024: Hybrid Human AI Systems for the Social Good: Proceedings of the Second International Conference on Hybrid Human-Artificial Intelligence*. IOS Press.

Zhiyu Chen, Bing Liu, Seungwhan Moon, Chinnadhurai Sankar, Paul Crook, and William Yang Wang. 2022. KETOD: Knowledge-enriched task-oriented dialogue. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 2581–2593, Seattle, United States. Association for Computational Linguistics.

Maaike HT de Boer, Jasper van der Waa, Sophie van Gent, Quirine TS Smit, Wouter Korteling, Robin M van Stokkum, and Mark Neerincx. 2023. A contextual hybrid intelligent system design for diabetes lifestyle management.

Yang Deng, Wenqiang Lei, Minlie Huang, and Tat-Seng Chua. 2023a. Goal awareness for conversational AI: Proactivity, non-collaborativity, and beyond. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 6: Tutorial Abstracts)*, pages 1–10, Toronto, Canada. Association for Computational Linguistics.

Yang Deng, Wenqiang Lei, Wai Lam, and Tat-Seng Chua. 2023b. A survey on proactive dialogue systems: Problems, methods, and prospects. *arXiv preprint arXiv:2305.02750*.

Yang Deng, Wenxuan Zhang, Weiwen Xu, Wenqiang Lei, Tat-Seng Chua, and Wai Lam. 2023c. A unified multi-task learning framework for multi-goal conversational recommender systems. *ACM Transactions on Information Systems*, 41(3):1–25.

Mekselina Doğanç and Ilia Markov. 2023. From generic to personalized: Investigating strategies for generating targeted counter narratives against hate speech. In *Proceedings of the 1st Workshop on CounterSpeech for Online Abuse (CS4OA)*, pages 1–12.

Bernd Dudzik, Jasper S. van der Waa, Pei-Yu Chen, Roel Dobbe, Iñigo M.R. de Troya, Roos Bakker, Maaike H. T. de Boer, Quirine Smit, Davide Dell'Anna, Emre Erdogan, Pınar Yolum, Shihan Wang, Selene Báez Santamaría, Lea Krause, and Bart A. Kamphorst. 2024. Viewpoint: Hybrid intelligence supports application development for diabetes lifestyle management. *Journal of Artificial Intelligence Research*.

Zane Durante, Bidipta Sarkar, Ran Gong, Rohan Taori, Yusuke Noda, Paul Tang, Ehsan Adeli, Shrinidhi Kowshika Lakshmikanth, Kevin Schulman, Arnold Milstein, et al. 2024. An interactive agent foundation model. *arXiv preprint arXiv:2402.05929*.

Antske Fokkens, Piek Vossen, Marco Rospocher, Rinke Hoekstra, and Willem Robert van Hage. 2017. GRaSP: Grounded representation and source perspective. In *Proceedings of the Workshop Knowledge Resources for the Socio-Economic Sciences and Humanities associated with RANLP 2017*, pages 19–25, Varna. INCOMA Inc.

Jianfeng Gao, Michel Galley, and Lihong Li. 2018. Neural approaches to conversational ai. In *The 41st international ACM SIGIR conference on research & development in information retrieval*, pages 1371–1374.

Steve Harris and Andy Seaborne. 21. March 2013. sparql 1.1 query language. w3c recommendation.

Timothy Hospedales, Antreas Antoniou, Paul Micaelli, and Amos Storkey. 2021. Meta-learning in neural networks: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(9):5149–5169.

Pili Hu and Wing Cheong Lau. 2013. A survey and taxonomy of graph sampling. *arXiv preprint arXiv:1308.5865*.

Ullrich Hustadt et al. 1994. Do we need the closed world assumption in knowledge representation? In *KRDB*, volume 1. Citeseer.

Seokhwan Kim, Spandana Gella, Chao Zhao, Di Jin, Alexandros Papangelis, Behnam Hedayatnia, Yang Liu, and Dilek Z Hakkani-Tur. 2023. Task-oriented conversational modeling with subjective knowledge track in DSTC11. In *Proceedings of The Eleventh Dialog System Technology Challenge*, pages 274–281, Prague, Czech Republic. Association for Computational Linguistics.

Jan Kocoń, Igor Cichecki, Oliwier Kaszyca, Mateusz Kochanek, Dominika Szydło, Joanna Baran, Julita Bielaniewicz, Marcin Gruza, Arkadiusz Janz, Kamil Kanclerz, et al. 2023. Chatgpt: Jack of all trades, master of none. *Information Fusion*, page 101861.

Lea Krause, Selene Báez Santamaría, Michiel van der Meer, and Urja Khurana. 2023. Leveraging few-shot data augmentation and waterfall prompting for response generation. In *Proceedings of The Eleventh Dialog System Technology Challenge*, pages 193–205.

Hung Le, Doyen Sahoo, Chenghao Liu, Nancy Chen, and Steven C.H. Hoi. 2020. UniConv: A unified conversational neural architecture for multi-domain task-oriented dialogues. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1860–1877, Online. Association for Computational Linguistics.

Phillip Lippe, Pengjie Ren, Hinda Haned, Bart Voorn, and Maarten de Rijke. 2020. Diversifying task-oriented dialogue response generation with prototype guided paraphrasing. *CoRR*, abs/2008.03391.

Bing Liu, Gokhan Tur, Dilek Hakkani-Tur, Pararth Shah, and Larry Heck. 2017. End-to-end optimization of task-oriented dialogue model with deep reinforcement learning. *arXiv preprint arXiv:1711.10712*.

Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang, and Minlie Huang. 2021a. Towards emotional support dialog systems. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 3469–3483, Online. Association for Computational Linguistics.

Yi Liu, Bohan Li, Yalei Zang, Aoran Li, and Hongzhi Yin. 2021b. A knowledge-aware recommender with attention-enhanced dynamic convolutional network. In *Proceedings of the 30th ACM international conference on information & knowledge management*, pages 1079–1088.

Jose L Martinez-Rodriguez, Aidan Hogan, and Ivan Lopez-Arevalo. 2020. Information extraction meets the semantic web: a survey. *Semantic Web*, 11(2):255–335.

Deborah L McGuinness, Frank Van Harmelen, et al. 2004. Owl web ontology language overview. *W3C recommendation*, 10(10):2004.

Martino Mensio, Giuseppe Rizzo, and Maurizio Morisio. 2018. Multi-turn qa: A rnn contextual approach to intent classification for goal-oriented systems. In *Companion Proceedings of the The Web Conference 2018*, pages 1075–1080.

Grégoire Mesnil, Yann Dauphin, Kaisheng Yao, Yoshua Bengio, Li Deng, Dilek Hakkani-Tur, Xiaodong He, Larry Heck, Gokhan Tur, Dong Yu, et al. 2014. Using recurrent neural networks for slot filling in spoken language understanding. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(3):530–539.

Tom Mitchell, William Cohen, Estevam Hruschka, Partha Talukdar, Bishan Yang, Justin Betteridge, Andrew Carlson, Bhavana Dalvi, Matt Gardner, Bryan Kisiel, et al. 2018. Never-ending learning. *Communications of the ACM*, 61(5):103–115.

Jinjie Ni, Tom Young, Vlad Pandelea, Fuzhao Xue, and Erik Cambria. 2023. Recent advances in deep learning based dialogue systems: A systematic survey. *Artificial intelligence review*, 56(4):3055–3155.

Kun Qian and Zhou Yu. 2019. Domain adaptive dialog generation via meta learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2639–2649, Florence, Italy. Association for Computational Linguistics.

Heung-Yeung Shum, Xiao-dong He, and Di Li. 2018. From eliza to xiaoice: challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, 19(1):10–26.

Eric Michael Smith, Mary Williamson, Kurt Shuster, Jason Weston, and Y-Lan Boureau. 2020. Can you put it all together: Evaluating conversational agents' ability to blend skills. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2021–2030, Online. Association for Computational Linguistics.

Piek Vossen, Selene Báez Santamaría, and Thomas Baier. 2024. A conversational agent for structured diary construction enabling monitoring of functioning well-being. In *HHAI 2024: Hybrid Human AI Systems for the Social Good: Proceedings of the Second International Conference on Hybrid Human-Artificial Intelligence*. IOS Press.

Piek Vossen, Selene Báez Santamaría, Lenka Bajčetić, Suzana Bašić, and Bram Kraaijeveld. 2019a. Leolani: A robot that communicates and learns about the shared world. In *2019 ISWC Satellite Tracks (Posters and Demonstrations, Industry, and Outrageous Ideas), ISWC 2019-Satellites*, pages 181–184. CEUR-WS.

Piek Vossen, Selene Báez Santamaría, Lenka Bajcetić, Suzana Bašić, and Bram Kraaijeveld. 2019b. A communicative robot to learn about us and the world. In *2019 Annual International Conference on Computational Linguistics and Intellectual Technologies, Dialogue 2019*, pages 728–743.

Piek Vossen, Selene Báez Santamaría, Lenka Bajcetić, and Bram Kraaijeveld. 2018. Leolani: A reference machine with a theory of mind for social communication. In *21st International Conference on Text, Speech, and Dialogue, TSD 2018*, pages 15–25. Springer/Verlag.

Piek Vossen, Lenka Bajčetić, Selene Báez Santamaría, Suzana Bašić, and Bram Kraaijeveld. 2019c. Modelling context awareness for a situated semantic agent. In *11th International and Interdisciplinary Conference on Modeling and Using Context, CONTEXT 2019*, pages 238–252. Springer.

Tsung-Hsien Wen, Milica Gašić, Nikola Mrkšić, Lina M. Rojas-Barahona, Pei-Hao Su, David Vandyke, and Steve Young. 2016. Multi-domain neural network language generation for spoken dialogue systems. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 120–129, San Diego, California. Association for Computational Linguistics.

Heinz Wimmer and Josef Perner. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1):103–128.

Hongcai Xu, Junpeng Bao, and Junqing Wang. 2020. Knowledge-graph based proactive dialogue generation with improved meta-learning. In *Proceedings of the 2020 2nd International Conference on Image Processing and Machine Vision*, pages 40–46.

Runzhe Yang, Jingxiao Chen, and Karthik Narasimhan. 2021. Improving dialog systems for negotiation with personality modeling. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 681–693, Online. Association for Computational Linguistics.

Tom Young, Frank Xing, Vlad Pandelea, Jinjie Ni, and Erik Cambria. 2022. Fusing task-oriented and open-domain dialogues in conversational agents. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 11622–11629.

Zheng Zhang, Minlie Huang, Zhongzhou Zhao, Feng Ji, Haiqing Chen, and Xiaoyan Zhu. 2019. Memory-augmented dialogue management for task-oriented dialogue systems. *ACM Transactions on Information Systems (TOIS)*, 37(3):1–30.

# A Appendix

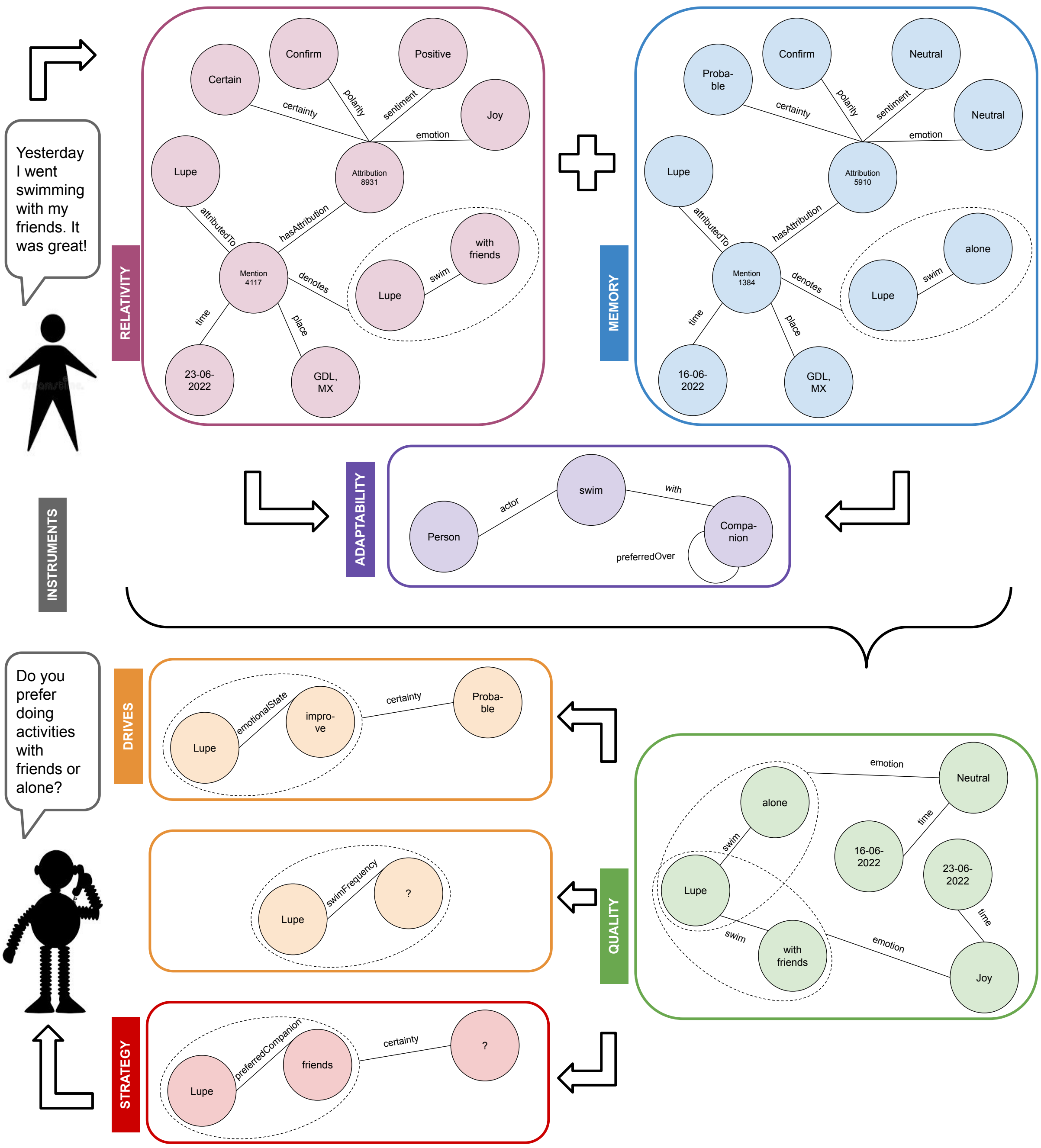


Figure 2: Example of dialogue in a Diabetes Lifestyle Management. A patient, Lupe, reports that she has done an activity with friends, expressing joy. This information gets incorporated into the memory, where information has been previously stored regarding a similar activity a week before, expressed in a neutral emotion. At the same time, this new information updates the T-Box, registering that activities may be performed with different companions, and some companions might be preferred over others. The accumulated information is assessed as a whole, in this case particularly focusing on differences between interactions. Furthermore, several areas of knowledge arise for potential improvement, including 1) improve certainty over Lupe's improved emotional state, 2) acquire information regarding the frequency of Lupe's activity, and 3) Lupe's preferences for performing activities with company. The latter is selected to continue the dialogue, and the information is expressed in natural language.