The 1st RobotDial Workshop on Dialogue Models for Human-Robot Interaction (Joint Workshop of RobotDial and LaCATODA)
Workshop in conjunction with IJCAI-PRICAI 2020. January 8 2020, online.

Dialogue Processing and System Involvement in Multimodal Task Dialogues

Ivana Kruijff-Korbayová
Multilinguality and Language Technology Lab
DFKI, Saarbrücken
ivana.kruijff@dfki.de

Kristiina Jokinen
AI Research Center
AIST Tokyo Waterfront
kristiina.jokinen@aist.go.jp

Abstract

We compare dialogue processing needs for several types of applications involving human-human and human-system collaboration and communication in various scenarios for a range of tasks. The tasks in these scenarios include collaboration which requires communication for information exchange and for task management. However, the necessary dialogue processing in the corresponding applications differs depending on the level of system involvement in the task(s) at hand, and in this paper we propose to put them on a continuum corresponding to the depth/coverage of the dialogue processing that is required. We provide a different and complementary perspective on ways of comparing dialogue processing than is common when looking at dialogue as either goal/task-oriented or chit-chat, or when considering different approaches to dialogue management, such as finite-state transducers or template-filling. We explore the possibility of gradually scaling dialogue processing complexity while sharing development resources for different applications along the continuum within one application domain.

1 Introduction

When working on several projects developing dialogue systems for different applications in various domains, it is clear that they have various needs and requirements concerning the complexity of the dialogue processing, but also have similar types and forms of dialogues which are common across the application domains. In this paper we explore the idea that systems for very different applications and domains have similar dialogue processing needs depending on the level of system involvement in the task(s) at hand. By dialogue processing needs we mean, for example, how deeply the system needs to understand the semantic content of what is being communicated, or how many modalities it needs to integrate into utterance processing. We compare dialogue processing needs for several classes of applications in various scenarios for a range of tasks. Our observations are based on various example scenarios including face-to-face customer service, such as at the airport or in a retail shop; elderly care interactions and healthcare support; collaborative assembly; disaster response teamwork. These scenarios involve various and very different tasks, such as collaboratively finding a travel plan, e.g., after a flight cancellation; determining and locating products according to a customer's needs; providing personal care, such as feeding; locating and handing over parts or tools; exploring a disaster site to search for casualties or to assess risks. In broad terms, collaboration on these tasks requires communication for information exchange (requesting/providing information) and for task management (assigning/assuming responsibilities). The classes of dialogue processing applications we consider include monitoring; performance feedback; task assistance; task execution. We observe that they differ in the level of system involvement in the task(s) at hand and in the interaction pertaining to doing the task(s). At one end of the spectrum are applications where a system observes a human doing a task, e.g., providing customer service to a client. The system is not actively involved in doing the task or in the interaction; it monitors the human's performance, documents and possibly assesses it. At the other end of the spectrum are applications where the system itself performs the task and conducts the corresponding interaction. Accordingly, we put the application classes on a continuum corresponding to the level of involvement of the system in the dialogue. We propose that this continuum also corresponds to the depth/coverage of the required dialogue processing, irrespective of (or: orthogonally to) the task at hand.

We provide a different and complementary perspective on ways of comparing dialogue processing than is common when looking at dialogue as either goal/task-oriented or (social) chit-chat, or considering different approaches to dialogue modelling, such as finite-state transducers, template-filling, plan-based reasoning, etc. We aim to provide a fresh perspective on the possibility of gradually scaling dialogue processing complexity while sharing development resources for the different applications along the continuum within one application domain, such as air travel planning, retail shop product information, assistance on a particular daily routine, assembly of a specific object, situation awareness gathering for disaster response, etc. We also try to identify domain-independent dialogue processing needs or aspects corresponding to the application classes.

The goal of the paper is to explain the proposed continuum and to initiate an exploration of the dialogue processing requirements for developing automated systems, which range from system support for human-human interactions to autonomous robot-human interactions.

The following figure shows the continuum. System involvement in the task and in the interaction increases from left to right, and as we propose so does the complexity of the required dialogue processing.

Figure 1: Continuum of system involvement and dialogue processing needs

We seek to find the parameters that characterize the interactions and also could be used as metrics to evaluate the system performance in a given task. This will allow us to adapt the system to different domains and scale the system up (and down) with respect to the requirements of the interaction. Generality of the system can thus be measured with respect to the parameters that define the interactions, and the interactions can also be compared with respect to these parameters. We intend the parameters to be on a general architectural and processing level describing the cognitive requirements for communication and language-based interaction rather than specific programming concepts or technical modules.

We also make a terminological distinction. When talking about a *task*, we refer to a certain sequence of actions that is needed to fulfil the goal of the interaction. For instance, a task can be to assist a human in daily life, e.g. getting exercise or taking medicine, or to find out some information, learn a new skill, or assist in a disaster situation. When we talk about *activity*, however, we refer to individual communicative acts and physical actions that underlie the completion of the task. For instance, dialogue activities consist of informing the partner of something, asking questions, taking turns and giving feedback. On the physical level, activities include moving one's body, gazing, gesturing, or manipulating objects in the world. Activities can be combined, e.g. in order to perform the task of assisting a human in their medicine routine, an interactive robot agent needs to perform the activities of informing the user of the task (steps) and giving feedback after the user has taken the medicine, as well as the activities of using gaze to monitor the situation and gesturing e.g. to offer the pillbox. It is important to notice that from a human point of view, activities can be considered multimodal. From a processing point of view, when we consider the different modalities as inputs (via human perception or system input sensors), the activities are linked to the media that they are expressed in, and after the fusion level where the information from the input modalities is integrated into one response action, we can talk about multimodal activities. We are primarily concerned about the communication, although there is, of course, always the communication about/concerning the task and the task itself.

The paper is organized as follows: We first describe the various scenarios that we consider and exemplify the tasks in Section 2. Then we describe each class of dialogue processing application along the continuum in turn and discuss the dialogue processing needs in Section 3. We conclude and indicate future work directions in Section 4.

2 Scenarios and Tasks

In this section we briefly introduce the various scenarios that are included in our consideration.

2.1 Customer Service

An example scenario is customer service at an airport (Fukuda et al., 2020, Nishimura et al. 2020), where the aim of the interaction (carried out by a human or a robot) is to help customers to solve problems, such as lost luggage or flight cancellation. Since it is a mixed-initiative scenario, it is possible that the customer leads the problem solving whereas the service worker just responds to information requests; or the service worker takes an active part in co-solving the customer's problem, gathers relevant information and proposes possible solutions which they may revise and refine together.

Customer service in a retail shop, such as a grocery or convenience store, is another case we consider. It is similar to the above in that a service worker assists a customer in solving a problem, such as the customer is not familiar with the available products; cannot find a product; or a product is not available and the customer needs advice on a suitable substitute.

Example interaction.

U: I need some salad.

R: There are different types available. Any particular type in mind?

U: Green salad

R: We have baby leave salad, Cesar's salad, and mix veg salad

U: What does mix veg salad contain?

R: It has cabbage, carrots, tomatoes, cucumber, and pumpkin seeds

U: Ok. I'll take that one.

R: Fine. I'll put it in the shopping cart for you. Anything else?

As we will discuss below, at the low end of the continuum (monitoring/assessment) a human is performing the customer service, while a system observes the interaction. At the high end of the continuum (task assistance and execution) the service is performed by a human and a system/robot in collaboration, or by an autonomous robot.

2.2 Elderly care and Healthcare

We consider elderly care and health-care scenarios, where a person is monitored and assisted by healthcare professionals or (partly) robots, who can provide help as needed, and especially in cases where the person does not know how to proceed or the situation seems to have changed so that action is needed. In rehabilitation situations, the goal is to assist the patient with daily routines and exercises, while in healthcare monitoring situations, the goal is to help the patient maintain a healthy lifestyle by monitoring their physiological data and providing appropriate recommendations. The new aspects in comparison to customer service are the multiple simultaneous goals, e.g., food intake and patient being active or self-sufficient. The first example interaction is from the domain of healthcare, in particular self-management support for young diabetic patients (Belpaeme at al. 2013; Kruijff-Korbayová et al. 2015). Here, a system interacts with a young patient to elicit information regarding their level of self-management.

R: Do you check glycemia yourself at home?

U: Yes, yes. I'm doing it all by myself. My parents don't check me practically never.

R: Very good. My I ask more?

U: Sure. Say all you want.

R: Do you realize when your blood sugar drops?

U: Yes, always.

The second example interaction is from the domain of elderly care. It also exemplifies an instruction and education application, where a knowledgeable expert or a more experienced colleague instructs or guides an inexperienced worker (apprentice), e.g. instructing a newcomer on a manual assembly task or coaching a novice care-giver on basic care-giving actions. In both scenarios, knowledge of a stepwise action plan of what actions to take, how to proceed, and how to finish off the action is needed. The teacher can also physically show how to do the task . Instances of this scenario differ w.r.t. whether an automated agent (robot) is involved and whether it is in the role of an expert or an apprentice. The case where the apprentice is an automated agent fall under task execution in our grouping, discussed below. The main problem is to have the agent learn the physical actions so as to be able to produce them in the appropriate context as needed. Here we focus on the cases where an automated system has the expert role, because that is where the levels of involvement can vary. Task completion requires communication so as to provide timely feedback or advice to the human apprentice, and it also assumes ability to observe the situation, identify problematic cases, and proactively suggest helpful information.

U: Hello

R: Hi. How can I help you today?

U: Well. I would need to know the right way to move a patient.

R: Ok. Do you mean to move the patient to another place, or improve the patient's position?

U: The latter. Improve the patient's position and lift them up

R: Ok, change the patient's position. Follow my instructions. Do you want an overview or stepwise presentation?

U: An overview please.

R: Ok, first move the patient sideways. Then gently lift the patient up. Check that the patient sits properly and can support themselves. Ask if the position feels good. Then you are done.

U: Ok. Thanks.

Similarly to the customer service, elderly care and health care applications can span the full continuum: at the low end of the continuum (monitoring/assessment) a human is performing the care, while a system observes the interaction. At the high end of the continuum (task assistance and execution) the care is performed by a human and a system/robot in collaboration, or by a robot.

2.3 Situation Assessment Teamwork

The final scenarios we consider involve teamwork for situation information gathering in disaster response, such as a team of firefighters using robots for reconnaissance in the aftermath of an incident, e.g., an earthquake or an explosion at an industrial site (Kruijff et al. 2014, Anakina and Kruijff-Korbayová, 2019). The team members have various roles: mission commander (MC), team leader (TL), robot operators (OP). The team explores the site, searching for persons, hazard sources, fires and other relevant points of interest. The MC leads the mission, the TL leads the human-robot team. They request situation information from the OPs, who report back with updates and can also share photos taken by the robot camera, as in the example interaction:

OP: Andreas, Markus from Andreas, come in.

TL: Andreas, come in.

OP: On first floor in the smoke found a barrel, green, labelled as environmentally hazardous material.

TL: Yeah, can you [unintelligible] whether anything is leaking?

OP: Yeah. It is a 200 litre barrel, whether anything is leaking I cannot currently tell.

TL: [EHM] Any thermal emission?

OP: No thermal emission.

TL: Okay. Priority on continuing person search. Andreas from Markus, priority on continuing person search.

Also here the possible applications span the entire continuum as we will discuss below. At the low end of the continuum a human team is performing the situation assessment, possibly using robots as (teleoperated) tools. At the high end of the continuum a decision-support system assists, and/or (partially) autonomous robots are agents in the team.

3 The Continuum of Dialogue Processing

In this section we explain the classes of applications on the continuum and the requirements on dialogue processing.

3.1 Monitoring

In the monitoring and assessment class we consider applications where the dialogue processing system is external to a human-human interaction, i.e., it passively observes and overhears, but does not participate as a partner in it.

A monitoring application supports a human performing a task by documenting the situation and the interaction, and processing data. For example, many healthcare diagnostic systems are monitoring the patient's state, and many multimodal systems (Kinect, voice and video recorder, biosensors, heart-beat, breathing controller) can be used to monitor the human behaviour in interactive situations for training purposes (e.g. how to serve customers, how to lift elderly people smoothly). (Kasper 2016) describes a monitoring application in the domain of robot-assisted disaster response, which creates mission reports by capturing system messages and automatically transcribing spoken communication among the team members.

Monitoring can require no dialogue processing at all, just capturing/recording. The minimal level of processing is speech recognition and/or detection of non-verbal signs. In our definition of the continuum, monitoring systems are tools to provide useful information. They do not make decisions, interpretation of the data is left for the humans. A monitoring system does not assess performance or provide performance feedback, which would make it a performance feedback or task assistance application, discussed below.

3.2 Assessment

A performance assessment application is like monitoring in that the dialogue processing system is also external to a human-human interaction, it only passively observes and overhears it. It processes the data to assess performance.

For performance assessment there needs to be some way to measure success of the communication and/or the task. For example in the customer service scenario, success is defined by the customer being happy/satisfied. This is some combination of being satisfied with the interaction itself and/or the solution, i.e., next possible course of action. An additional aspect of performance evaluation may be the quality of the solution, e.g., efficiency of the flight connection taking the customer's constraints into account, or usefulness of the product information. Giving the customer a sales recommendation may be an additional criterion of performance in the retail shop scenario.

The elderly care and healthcare scenarios are very similar to the customer service ones from the viewpoint of performance assessment. Success is again defined by some combination of the assisted person being satisfied with the interaction and the assistance provided. In addition, the carer's ability to listen to the client, providing appropriate feedback and especially encouraging them to continue speaking by showing listening capability is crucial.

We are not aware of studies applying dialogue processing to performance assessment of teamwork in disaster response scenarios, in particular situation awareness gathering. These scenarios however have many similarities with military teamwork, and there has been a lot of human-factors research on assessing the performance of military teams. Some of this work studied verbal communication and found that analysis of communication sequences contributes to the understanding of effective crew processes, for example in situation assessment command & control tasks performed in simulation (Bowers et al. 1998). Results obtained on manually annotated data were also reproduced in studies using automatic analysis of the semantic content of team communication and automatic verbal behaviour labelling. For example, (Martin and Foltz 2004) report that teams that tend to state more facts and acknowledge other team members more tend to perform better; those that express more uncertainty and need to make more responses to each other tend to perform worse. These results should be directly applicable to teamwork in robot-assisted disaster response.

Is task expertise needed for performance assessment? Obviously, it is needed for assessing task performance. It is also needed or at least useful for a deep understanding of what is going on in the communication at the content level. Nevertheless, the military teamwork performance studies mentioned above obtained good results, on-a-par with human judgments, with only superficial semantic analysis and no additional task knowledge. This example shows that even with very simple surface-oriented modelling one can already build a useful application to support humans on a task.

In order to provide performance assessment a system needs to process only those selected features of dialogue, which are correlated with performance in the given scenario. For many scenarios which are already done by humans there exist performance assessment studies and guidelines. This is not to say that detecting the features automatically is always easy. Sometimes it makes sense to use surrogate features that lend themselves to automatic detection more easily, if they provide a good-enough approximation. Since dialogue act sequences, such as question-answer pairs, are relevant, dialogue acts need to be recognised. Dialogue processing for performance assessment in face-to-face service encounters needs to take into account multimodal interaction features, including linguistic and paralinguistic cues. Finally, since emotions play an important role in some scenarios, features such as intonation, facial expressions, hand and body posture need to be recognised.

3.2 Performance Feedback

In this class of applications, a system provides the performance assessment discussed above to a human doing the respective task and conducting the corresponding interaction, such as customer service, client/patient care or situation information gathering.

Performance feedback is the simplest form of training assistance. Many interesting questions concern the delivery of performance assessment/feedback, so as to motivate and encourage a learner. We do not discuss this here.

From the viewpoint of dialogue processing, the same is required as discussed above for performance assessment. The performance feedback can be provided either at the end or during the task/interaction session. Performance feedback provided during a session (training or real), entails a certain amount of involvement in the session, thus a notch away from pure passive external monitoring/assessment. As such it may help to improve the performance within the session. It may also hurt it if delivered poorly and/or at the wrong moment(s). If performance feedback is delivered during a session then its timing is also important, which means that additional understanding of the dynamics of the interaction is needed. The system needs to know when to deliver the assessment, e.g., have a model of turn management and possible task structure. Dialogue technology can also be used to communicate the feedback to the human.

Beyond delivering just performance assessment as the simplest training assistance, a system can provide tips/hints how to improve performance. Using the above example from the military domain, tips to improve performance through better communication would include recommendations to present facts and acknowledge (more often).

Providing further levels of training assistance involving task-level hints beyond the performance assessment, is a kind of task assistance application on our continuum, and is discussed below. In the next steps of the continuum dialogue processing needs go deeper into actual understanding.

3.3 Task Assistance

In a task assistance application, the system supports the human(s) doing the task, including communicating about it. The human remains in control, decides how to conduct the interaction and how to proceed on the task. The system provides (various kinds of) information relevant to making the decisions and performing the task. This may involve background knowledge and/or accurate and up to date information for the human to make decisions. The system may autonomously notify the human of changes in certain monitored values (cf. Section 3.1), of abnormalities or issues that can be potentially dangerous. Even more elaborate assistance involves suggesting solutions and/or pointing out problems. Regarding the interaction, the system may alert about what might be embarrassing, offensive, bad manners, or otherwise infringing on privacy and ethical issues. Crucially, a system can provide useful support, even if it does not perform the task itself.

Task assistance may support a human in performing the task on the given occasion, or it may help increase their competence in performing the task in the future. This reflects the concept of training on a job, or workplace learning, which is emerging in Industry 4.0. (Kravčík 2019). It is also consistent with the Japanese view of Society 5.0 which emphasizes the technology being used for human well-being and assistance, and the symbiosis of humans and robots in solving tasks. Even further steps towards more sophisticated training require a simulation of the task and/or the communication partner, all the way to training in virtual reality. We do not discuss these here.

When the human is performing a physical task, such as feeding a client in the care scenario or gathering situation information in the disaster response scenario, task assistance may go beyond providing information in that a robot assists the human by supporting actions, e.g. by handing over utensils or stabilizing the camera, respectively. This is on the borderline to (collaborative) task execution, which will be discussed below.

Task assistance requires more/deeper understanding of the interaction than performance assessment, although also partial understanding may be enough, such as the main semantic units of utterances. The required depth and completeness of understanding depends on the targeted type/level of assistance. The important point w.r.t. to the continuum of dialogue processing is, what is the minimum of interaction understanding and/or task competence that is necessary to be able to provide a certain (required) level of useful support. The system also needs to determine the timing of the assistance, following the principle “right information at the right time”.

3.4 Task Execution

Finally, we consider as task execution the class of applications where it is the system performing the task, alone or in collaboration with a human, and communicating about it, making decisions partially or fully autonomously. In a shared control mode both the system and the human can make the decisions. Depending on the situation, the system may release and/or the human take over control. In the case of an autonomous system, the system is in control (even if a human supervisor may still have the possibility to override the system’s decision).

Task execution clearly requires the deepest understanding of the interaction and the task, or the most accurate models trained on task execution data. Context-aware dialogue management is important in order to make human-robot interaction more natural, see argumentation e.g. in (Jokinen 2018), (Wilcock & Jokinen 2020). This presupposes not only understanding of what the user, e.g., the customer, client/patient or team leader in our example scenarios, says but also understanding the context and being able to provide appropriate feedback to the user. Such systems need to have deeper processing capability and larger knowledge, to be able to understand the user needs for communication, which are very different in the different tasks and scenarios under consideration.

Task execution applications are what most human-robot interaction system development focuses on, because autonomous robotic systems are seen as the Holy Grail. What we aim to draw attention to by discussing the system involvement and dialogue processing continuum is the fact that also more limited dialogue and task understanding can suffice and serve well to support humans performing tasks and interaction in various scenarios.

Moreover, the dialogue processing competence can be extended stepwise, from the least demanding applications, i.e., monitoring and assessment, to task assistance and task execution. An example of such a gradual extension can be found in (Cooke et al. 2016), who describe a synthetic teammate to perform situation assessment in a command and control task in a simulated environment. The synthetic teammate has components for language comprehension and generation, dialogue management, task behaviour and a situation component that provides context, which were developed following up on the performance assessment work in the military domain reviewed above in Section 3.1.

5 Conclusion and Future Work

We proposed to consider dialogue processing applications which are aimed at supporting humans or collaborating with them on a task, such as robotic agents, along a continuum in terms of the system's involvement in the interaction and the execution of the task. The minimum involvement end of the proposed continuum corresponds to monitoring and performance assessment applications, where the system is a passive observer. Involvement increases in applications providing performance feedback and task assistance, and culminates in task execution applications.

Our claim is that the dialogue processing functionality, such as dialogue interpretation, needed in order to provide useful support for humans on a task, is lowest at the minimum involvement end of the continuum, and increases with increasing system involvement. In other words, useful human-support systems can be developed with relatively little dialogue processing, and the amount and quality/difficulty of dialogue processing can be gradually extended for increasing involvement. In this sense the continuum can be seen as providing a system development roadmap, where the development of a particular system for human support in any given domain/scenario can proceed from the simplest application to more complex ones, allowing dialogue processing resources to be reused and incrementally extended.

We presented initial observations concerning the dialogue processing needs along the continuum for several example scenarios and tasks from projects we have been working on. The proposed continuum provides a first outline of various dialogue processing applications and a first classification of the dialogue resources to be reused and extended when developing applications that enable and support dialogue interaction in the given scenarios.

One of the important issues nowadays is evaluation of the systems so as to compare various strategies and system performances and be able to adapt the system to the needs and requirements of the users. As human evaluation with interviews and questionnaires is costly and time-consuming, much work is also put into standardisation methods so as to automatically and quickly extend and modify dialogue systems to new domains and new topic areas. This is relevant especially in prototyping information-providing systems such as QA systems and chatbots. For instance, Deriu et al. (2020) survey evaluation methods for this purpose and differentiate dialogue systems into three classes: task-oriented, conversational agents, and question-answering dialogue systems. Our research is related to their work, but we focus on the interaction related to activities between humans and robots, rather than verbal interaction as such. In other words, the focus is on the involvement of the human partner as well as the necessary technology to enable cooperative communicative activities.

In future work we plan to elaborate on the dialogue processing needs along the proposed continuum in more detail for selected scenarios and tasks.

Acknowledgments

We would like to thank the members of our research teams for fruitful discussions. The study is based on the results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO), Japan, past EU-funded projects Aliz-E (ICT-248116), PAL (Horizon2020-43783) and TRADR (IST-609763), as well as the current A-DRZ project funded by the German Ministry of Education and Research (I3N14856).

References

Anakina, T. and Kruijff-Korbayová, I. 2019. Dialogue act classification in team communication for robot assisted disaster response. Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue, Stockholm, Sweden, September 11-13, 2019.

Belpaeme, T., Baxter, P., Read, R., Wood, R., Cuayáhuitl, H., Kiefer, B., Racioppa, S., Kruijff-Korbayová, I. et al. *Multimodal Child-Robot Interaction: Building Social Bonds*. Journal of Human-Robot Interaction, *1*(2), 33-53. 2013.

Bohus, D. Horvitz, E. 2009. Models for multiparty engagement in open-world dialog. Proceedings of SIG-DIAL '09, pp. 225–234.

Deriu, J., Rodrigo, A., Otegi, A, Echegoyen, G., Rosset, S., Agirre, E., Cieliebak, M. (2020). Survey on evaluation methods for dialogue systems. *Artificial Intelligence Rev*iew. https://doi.org/10.1007/s10462-020-09866-x

Fukuda, K., Vizcarra, J., Nishimura, S. : Massive Semantic Video Annotation in High-end Customer Service - Example in Airline Service Value Assessment. In: Proceedings of HCI International 2020. (2020)

Jokinen K., Nishimura S., Watanabe K., Nishimura T. (2019) Human-Robot Dialogues for Explaining Activities. In: D'Haro L., Banchs R., Li H. (eds) 9th International Workshop on Spoken Dialogue System Technology. Lecture Notes in Electrical Engineering, vol 579. Springer, Singapore, pp. 239-252.

Jokinen, K., Fukuda, K., Iino, N., Nishimura, S., Nishimura, T., Oota, Y., Watanabe, K., Yoshida, Y. (2019). Privacy and Sensor Information in the Interactive Service Applications for Elder People. Proceedings of 7th National Conference of Serviceology, Tokyo.

Jokinen K. (2018) Dialogue Models for Socially Intelligent Robots. In: Ge S. et al. (eds) Social Robotics. ICSR 2018. Lecture Notes in Computer Science, vol 11357. Springer, Cham, pp. 127-138

Jokinen, K., Wilcock. G. 2014. Multimodal Open-Domain Conversations with the Nao Robot, In J. Mariani, S. Rosset, M. Garnier-Rizet and L. Devillers (eds.), Natural Interaction with Robots, Knowbots and Smartphones: Putting Spoken Dialogue Systems into Practice. Springer New York, pp. 213–224.

Kasper, W. (2016) Team monitoring and reporting for robot-assisted USAR missions. In *International Symposium on Safety, Security, and Rescue Robotics (SSRR)* (2016), IEEE, pp. 246–251.

Kruijff, G.J., Kruijff-Korbayová, I, Keshavdas, S., Larochelle, B., Janíček, M., Colas, F., Liu, M., Pomerleau, F., Siegwart, R., Neerincx, M., Looije, R., Smets, N., Mioch, T., van Diggelen, J., Pirri, F., Gianni, M., Ferri, F., Menna, M., Worst, R., Linder, T., Tretyakov, V., Surmann, H., Svoboda, T., Reinstein, M., Zimmermann, K., Petříček, T., Hlaváč, V. *Designing, developing, and deploying systems to support human-robot teams in disaster response*. Advanced Robotics, 28 (23), Pages 1547-1570, Taylor & Francis, London, 12/2014.

Kruijff-Korbayová, I., Oleari, E., Bagherzadhalimi, A., Sacchitelli, F., Kiefer, B., Racioppa, S., Pozzi, C., Sanna, A. (2015) Young Users' Perception of a Social Robot Displaying Familiarity and Eliciting Disclosure. In: Proc. of ICSR 2015. Social Robotics. Lecture Notes in Computer Science, vol. 9388, pp. 380-389.

Kravčík, M. 2019. Adaptive Workplace Learning Assistance. In Augstein, M., Herder, E. & Wörndl, W. *Personalized Human-Computer Interaction* . Chapter 11, Pages 283-302. De Gruyter Oldenbourg.

Lala, D., Wilcock, G., Jokinen, K., Kawahara, T. ERICA and WikiTalk. Proceedings of 28th International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao. pp. 6533–6535.

Nishimura, S., Oota, Y., Fukuda, K.: Ontology construction for annotating skill and situation of airline servceies to multi-modal data. In: Proceedings of HCI International 2020. pp. 1--15 (2020)