Towards a Versatile Intelligent Conversational Agent as Personal Assistant for Migrants

Leo Wanner[1,2](✉), Matthias Klusch[3], Athanasios Mavropoulos[4], Emmanuel Jamin[5], Víctor Marín Puchades[5], Gerard Casamayor[2], Jan Černocký[6], Steffi Davey[7], Mónica Domínguez[2], Ekaterina Egorova[6], Jens Grivolla[2], Gloria Elena Jaramillo Rojas[3], Anastasios Karakostas[4], Dimos Ntioudis[4], Pavel Pecina[8], Oleksandr Sobko[5], Stefanos Vrochidis[4], and Lena Wertmann[9]

[1] Catalan Institute for Research and Advanced Studies, Barcelona, Spain
[2] Pompeu Fabra University, Barcelona, Spain
leo.wanner@upf.edu
[3] Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI), Saarbrücken, Germany
[4] Centre for Research and Technology Hellas, Thessaloniki, Greece
[5] Everis Inc., Madrid, Spain
[6] Brno University of Technology, Brno, Czech Republic
[7] Centric, Sheffield Hallam University, Sheffield, UK
[8] Charles University Prague, Prague, Czech Republic
[9] Nurogames GmbH, Cologne, Germany

Abstract. We present a knowledge-driven multilingual conversational agent (referred to as "MyWelcome Agent") that acts as personal assistant for migrants in the contexts of their reception and integration. In order to also account for tasks that go beyond communication and require advanced service coordination skills, the architecture of the proposed platform separates the dialogue management service from the agent behavior including the service selection and planning. The involvement of genuine agent planning strategies in the design of personal assistants, which tend to be limited to dialogue management tasks, makes the proposed agent significantly more versatile and intelligent. To ensure high quality verbal interaction, we draw upon state-of-the-art multilingual spoken language understanding and generation technologies.

Keywords: Conversational agent · Agent service coordination · Dialogue management · Ontologies · Multilingual

Supported by the European Commission in the framework of its Horizon 2020 R&D Program under the contract number 870930.

F. Dignum et al. (Eds.): PAAMS 2021, LNAI 12946, pp. 316–327, 2021.
https://doi.org/10.1007/978-3-030-85739-4_26

1 Introduction

With Siri, Cortana, Alexa, Google Assistant, etc. the concept of a virtual conversational assistant reached the general public, and with it also rose the demands for more intelligence, versatility, and a broader coverage of background information. However, data-driven models underlying assistants like those mentioned above are naturally limited in terms of the scope of their memory, interpretative intelligence and cognition. Thus, although one can ask Siri about the weather prediction, one will not obtain a satisfactory answer when the inquiry concerns the location "where my mother lives" or even just whether "it will get warmer". Notwithstanding, for some contexts, such skills are highly desirable, in particular, when the assistant is supposed to be personalized, i.e., be aware of the profile, needs and capabilities of a specific user. Such contexts, include, e.g., assistance in healthcare or education, interaction with elderly, or support of migrants during their reception and integration in the host country.

Knowledge-driven conversational agents are in this sense an alternative to data-driven assistants. The design of a state-of-the-art knowledge driven conversational agent usually foresees a dialogue manager (DM) as the central knowledge-processing module, which accesses and reasons over an underlying ontology to plan the dialogue moves, possibly taking into account the context and the profile of the addressee; cf., e.g., [6,12,21,22]. Recently, neural reinforcement learning-based DMs proved to achieve an impressively good performance in well-defined, limited contextual setups such as restaurant reservation [24], travel booking [23], movie ticket purchase [11], etc. Still, while they cope well with the management of the dialogue history and the belief states, planning of dialogue moves, control of the coherence of the generated discourse, etc., they are not designed to carry out tasks such as, e.g., retrieval of useful information based on the profile of the user, identification of the closest office for residence application submission, or assessment of health data and determination of the adequate reaction. And they are even less designed to interact with each other. To cover these tasks, techniques for multi-agent coordination such as coalition formation and clustering, semantic service selection and composition planning can be used.

In what follows, we present work in progress on Embodied Conversational Agents (ECA) in the context of the WELCOME Project, henceforth referred to as "MyWelcome agent", whose design foresees the separation of the agent behavior and dialogue management tasks in two different modules. The agents are intended to act as personalized virtual companions of migrants or refugees (henceforth *Third Country Nationals*, TCNs) in specific contexts related to reception and integration in the host country.

The reception, administration and integration of TCNs imply a variety of tasks that can be roughly grouped as follows:

(1) registration and guidance through the administrative procedures and documentation needed for registration as well as provision of information on basic healthcare, legal, and social services;

(2) language courses and civic and social education;
(3) cultural and societal integration;
(4) coaching in daily life contexts, including, e.g., filling out a form or submission of an application; job interview training, legal consultation, etc.

Personalized embodied conversational agents that are capable of covering these tasks must show the following characteristics: **(i)** dispose of targeted background knowledge in a number of areas (including, among others, host country-specific migrant reception and integration procedures, prominent local cultural, civic, and social facts, everyday life contexts, etc.) as well as of the personal profile of a user (TCN, in our case); **(ii)** command situational awareness in interaction with the user; **(iii)** understand and interpret the verbal statements of the user; **(iv)** plan the actions to be undertaken by the agent in order to act respectively react appropriately to the statement of the user under given situational circumstances and the dialogue to be conducted by the agent; and **(v)** communicate with the user in the language of their preference.

Consider a fragment of a sample conversation as targeted by the MyWelcome agent:

U: *Hello, I would like to register for the First Reception Service.*
A: *Hello! What is your name?*
U: *Karim*
A: *Very good Karim! I need some basic information from you. Since when are you in the country and what is your current address?*
U: *I am still staying with friends.*
A: *In this case, I need your friends' address.*
U: *It is Lepant, 142 in Terrassa.*
A: *And since when do you live there?*
...

In the next section, we outline the design of the MyWelcome agent. Section 3 presents the corresponding knowledge and data models and the agent services, and Sect. 4 the language technologies. Section 5, finally, summarizes the current state of our work and sketches its ongoing extension.

2 Design of the MyWelcome Agent

Each user is assigned their own personalized MyWelcome agent realized in terms of a mobile cloud application. Figure 1 shows the global MyWelcome agent platform architecture, which is grounded in the concept of services.[1] The backbone of the platform is the *Dispatcher* service, which channels the communication between the different platform services. Due to the challenging system reaction time restrictions on the one side and a required 1:n control flow efficiency on the

[1] We use the notion of *service* in order to highlight the functionality perspective on the individual software modules or components of the platform.

other side, a Dispatcher is preferred to a central Bus service or direct interaction between the services. From the perspective of their function, the individual platform services can be grouped into *message understanding* services, which include language identification (LIS), automatic speech recognition (ASR), language analysis (LAS); *message interpretation and reaction* services, which consist of one or multiple personal MyWelcome agents for agent-driven semantic service coordination (ADSC) and a dialogue management service (DMS); and *message generation* services, which are composed of the natural language generation (NLGS) and text-to-speech (TTS) services. For multi-agent coordination tasks, the MyWelcome agents communicate directly with each other via their endpoints, without leveraging the Dispatcher service.

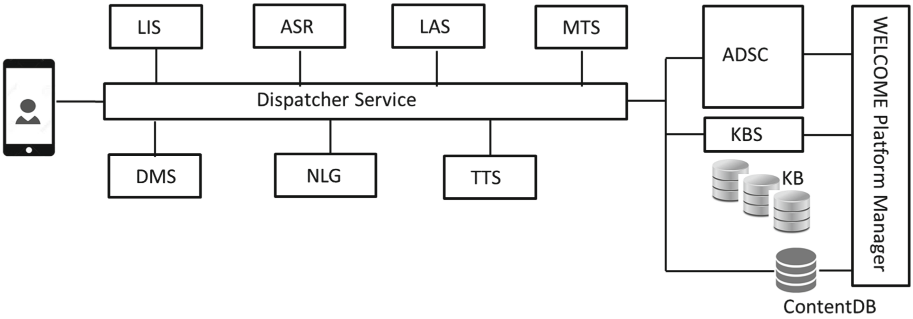


Fig. 1. Architecture of the MyWelcome agent platform.

Figure 2 displays the data flow between the individual MyWelcome services for the first two turns of our dialogue in Sect. 1. The user speaks to the avatar that embodies the MyWelcome agent in the mobile MyWelcome Application. The spoken turn is passed to the Dispatcher, which dispatches its analysis by the LID, ASR and LAS services. The outcome of the analysis, which consists of a predicate argument structure and the corresponding speech act (i.e., *communicative intent* of the speaker), is mapped onto an OWL representation and introduced into the local knowledge repository of the ADSC, where it is used for context-aware semantic service selection and planning. The output of the ADSC is passed via the Dispatcher to the DMS, which decides on the next dialogue move. The syntactic structure and the exact wording of the move is synthesized by the NLG service, spoken by the TTS service and played on the MyWelcome application.

3 Knowledge and Data Models and Agent Services

The quality of a knowledge-driven personalized conversational agent decisively depends on the coverage of its ontologies and the extent to which the personal features of the users are captured. In this section, we first present the knowledge and data models drawn upon by the agent and then the realization of the agent interaction services.

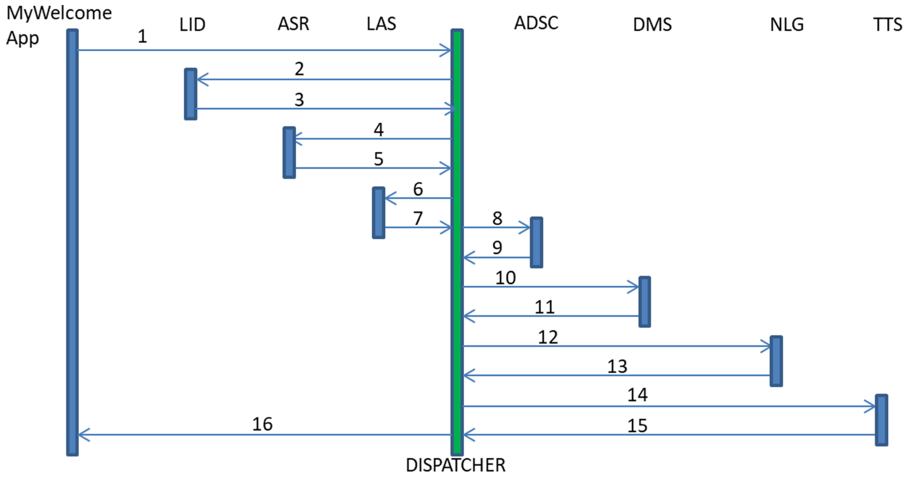


Fig. 2. Illustration of the data flow (the arrows indicate the data exchange between the individual modules and the Dispatcher in the order marked by the numbers); see description in Sect. 2.

3.1 Knowledge and Data Models of the MyWelcome Agent

The MyWelcome agent knowledge models leverage different types of information: (i) background information on migrant reception and integration policies and procedures, language learning curricula, social services, etc.; (ii) user-specific data, initially provided during user registration and subsequently distilled from the natural language interaction of the TCN with the MyWelcome agent; (iii) temporal properties of user resources, dialogue strategies and dialogue history, cultural integration activities, language learning advances, etc., and (iv) information obtained via reasoning and decision making on the available information.

To ensure GDPR-compliance with respect to the maintenance of personal TCN data, we distinguish between local and global agent knowledge repositories, which are realized as separate partitions of a semantic graph DB. Each MyWelcome agent is assigned its own **Local agent repository** realized as a tripartite RDF triple store: the *Local Agent Knowledge Repository* (LAKR), which manages the personal data of its "TCN master", the *Local Agent Repository* (LAR), which keeps the internal states of the agent, and the Local Service Repository (LSR), which maintains the services of the agent. Each triple store can be accessed only by its owner agent.[2] The population and management of local repositories are handled by the Knowledge Management Service (KMS), which is a service responsible for initially converting external knowledge into RDF triple store-compliant representations, and then mapping them to the respective ontologies. The **Global Knowledge Repositories** contain RDF triple stores

[2] The LAKRs can be also accessed via secure interface by the TCN whose data it contains and the responsible authority.

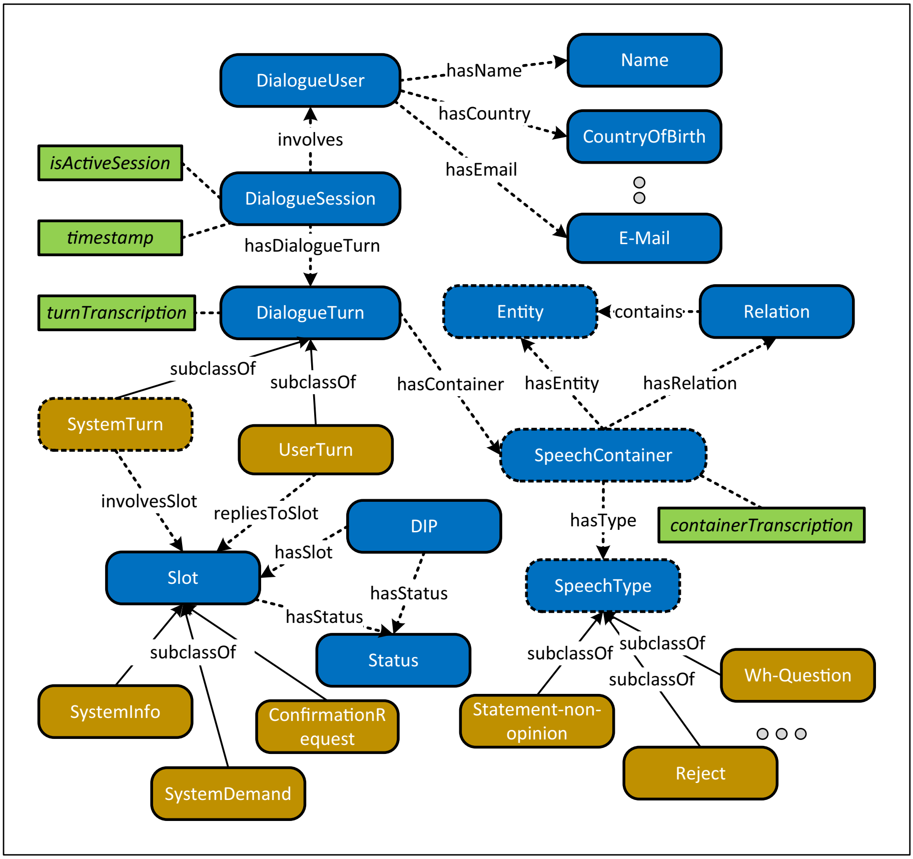


Fig. 3. Abstract representation of the WELCOME ontology.

that concern all agents, again separated into three partitions: the *WELCOME Domain Knowledge Repository* (WDRK), which contains the content of the scenarios covered by the agents, the *Semantic Service Descriptions* (WSR), and the *WELCOME Agent Repository* (WAR) with the ids of all created agents. A KMS-like service called *Knowledge Base Service* (KBS) facilitates the conversion into RDF and introduction of new information into WDKR or LAKR via the Platform Manager and initiates the LSR of a newly created agent. In addition to the knowledge repositories, a **content database** is included into the knowledge/data model of the MyWelcome Application. The content DB contains, e.g., sentence templates used for the agent move generation, relevant textual material pointed to in the generated moves, etc.

A fragment of the ontology, which forms the backbone of LAKR is shown in Fig. 3. It depicts the vocabulary that supports the dialogue with the user to inquire information on the profile of the user. The fact that a user exists in the ontology is modeled in terms of an instance of the class `DialogueUser`.

If the name of the user is unknown, the agent will ask for it. The content of the question is modeled as an instance of the class DialogueTurn, which is part of DialogueSession:

```
:session1 a :DialogueSession ;
    :hasDialogueTurn :systemTurn1 ;
    :involvesDialogueUser :TCN .
```

The agent turn consists of a timestamp and a speech act container, which is associated with a dialogue input slot obtainName:

```
:systemTurn1 a :DialogueTurn ;
    :hasSpeechActContainer :speechAct1 ;
    :timestamp "2021-04-28T10:46:50.623+03:00"^^xsd:dateTime .

:speechAct1 a :SpeechActContainer ;
    :involvesSlot :obtainName .

:obtainName a :SystemDemand ;
    :hasInputRDFContents :Unknown ;
    :hasOntologyType :Name ;
    :confidenceScore "0"^^xsd:integer ;
    :hasNumberAttempts "0"^^xsd:integer ;
    :hasStatus :Pending ;
    :isOptional "yes" .
```

The user replies to the agent's question by providing their name (e.g., *Karim*). The user's response is also modeled as an instance of DialogueTurn and consists of a timestamp, a transcription and a speech act container. The LAS service detects that there is an entity of type Person in the user's speech act. Cf. the corresponding codification in ontological terms:

```
:session1 a :DialogueSession ;
:hasDialogueTurn :systemTurn1 ;
:hasDialogueTurn :userTurn1 ;
:involvesDialogueUser :TCN .

:userTurn1 a :DialogueTurn ;
    :hasTurnTranscription "Karim"^^xsd:string ;
    :hasSpeechActContainer :speechAct2 ;
    :timestamp "2021-04-28T10:47:00.300+03:00"^^xsd:dateTime ;
    :prevTurn :systemTurn1 .

:speechAct2 a :SpeechActContainer ;
    :hasContainerTranscription "Karim"^^xsd:string ;
    :hasDetectedEntity :entity1 .
    :repliesToSlot :obtainName .

:entity1 a :DetectedEntity ;
    :hasEntityType "Person"^^xsd:String .
```

Finally, the KMS infers that this particular entity of the user's speech act is an expected response for the particular question, creates a new instance of type `Name` that is associated with the user's profile and updates the status of the dialogue input slot to *Completed*.

```
:TCN a :DialogueUser ;
    :hasName :Karim .
:Karim a :Name .
```

3.2 Agent Interaction Services and Dialogue Management

The TCN shall be intelligently assisted in the execution of the different procedures (referred to as *social services*) associated to the reception and integration, such as first reception, language learning, asylum application, etc. The actions of the agent in the context of social services are determined by two different modules. The first of them (ADSC) plans all "interaction" moves of the agent and coordinates its behavior; the second (DMS) plans the verbalization of the communication-oriented moves. In what follows, both are briefly introduced.

The behavior of an agent is determined by its *Behavior Trees* (BTs), which are conditioned by the facts stored in its LAKR. The facts are accessed via SPARQL queries attached to the nodes of the tree. The agent core, which is implemented in an *Access Java Agent Nucleus* (AJAN) server[3], encodes targeted sub-BTs that determine how to react to speech acts (cf. Sect. 4) that request specific social services. To react to a request, the agent core invokes its internal *Semantic Service Computing* (SSC), which identifies the corresponding service in its *Local Service Repository* (LSR). The LSR encodes the representation of the services in OWL-S 1.1. In this context, two different and complementary strategies are implemented inside the SSC for semantic service coordination: service selection and service composition.

To select a service, the agent core launches a service request that contains all relevant information (speech act and facts) from the user's move (passed by LAS to the KMS of the agent). The request is taken up by the *iSeM matchmaker* [9], which retrieves top-k semantically relevant services from the LSR.

If no relevant service is found in the LSR, the agent core invokes the SSC to call its semantic service composition planner to satisfy the given request of a service. This planner works as an offline state-based action planner [4,5]. Its action plan corresponds to the desired service plan; the initial state is a set of facts in OWL extracted by the KMS from the LAKR (its fact base), which describes the current state of the world; the goal state is a set of facts in OWL that shall persist after performing the plan. If a service composition matches the goal, the planner solves the given semantic service composition problem. If no service composition matches the goal, the agent asks DMS to inform the user that the service request cannot be satisfied and to propose to check the information available in the Frequently Asked Questions also provided in the MyWelcome

[3] https://asr.dfki.de/ajan/.

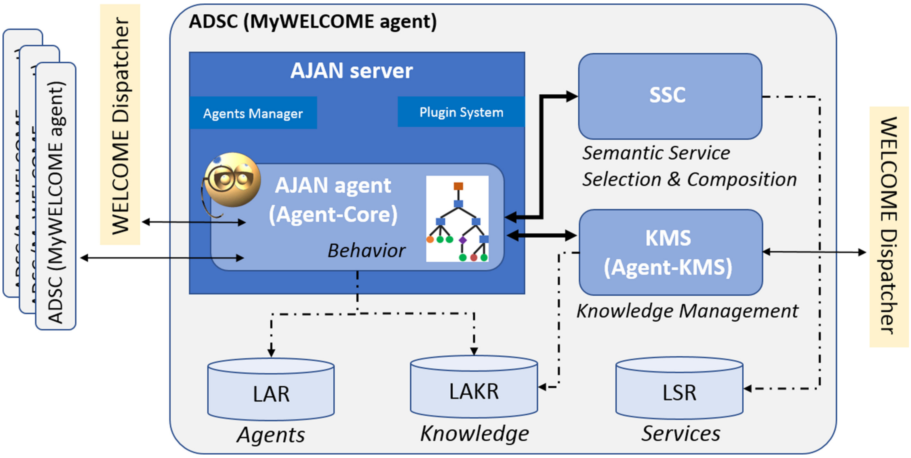


Fig. 4. Architecture of the MyWelcome agent (ADSC).

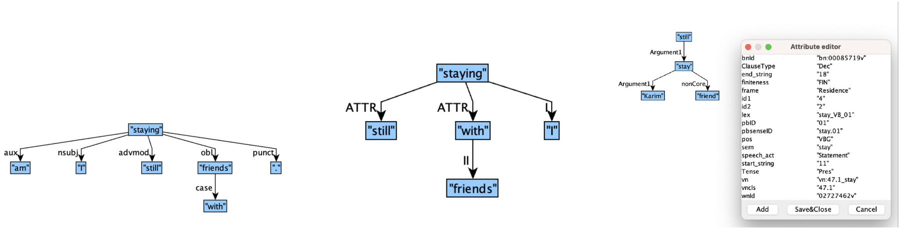
Application. On the other hand, the agent-core performs the atomic or composed service by collecting or providing information from/to the user. To create the dialogue strategies, the agent-core and DMS interact with each other based on a specific data structure called *Dialogue Input Package* (DIP). A DIP is created by the agent-core describing what the particular information that needs to be asked or communicated (slots in the DIP) is, and some meta-information such as the number of times that a piece of information has been sent to the user, and the type of the information (System Info or System Demand). DMS, together with the message generation service, decides what information (slot) to communicate to the user and how to do it. The current version of the DMS is based on the KRISTINA dialogue manager [21]. Its next version will incorporate a versatile neural network-based repair strategy of the dialogue faulted by ASR errors or misleading user information.

4 Language Technology Services

The MyWelcome agent is projected to be multilingual, speaking Syrian (Levantine) and Moroccan (Darija) Arabic, Catalan, English, German, Greek, and Spanish. In its current state, it is tested in English

Spoken Language Understanding. The spoken language understanding technologies cover language identification (LI), automatic speech recognition (ASR), language analysis (LAS), and machine translation (MT). **The LI service** is based on [16], which utilizes a robust generative concept of i-vectors, i.e., utterance embeddings, based on generative models. Acoustic features, which serve as input for i-vector training, are multiligually trained on stacked bottleneck features [2]. After the extraction of i-vectors, a Gaussian Linear Classifier is applied. **The ASR module** is based on the kaldi toolkit [17] and [8]. **The LAS**

module consists of the surface language analysis and deep language analysis modules. The distinction between surface and deep analysis is made in order to ensure word sense disambiguation and entity identification and linking at the deep side of the analysis. For surface language analysis, [19] is used. For word sense disambiguation and entity linking, we use BabelFy [14]; for entity (concept) identification, we adapt [18]. The relations that hold between the concepts are identified applying rule-based grammars to the results of these analysis submodules; the grammars are implemented in the graph transduction framework [1]. LAS outputs a predicate argument structure, which is mapped by the KMS onto an RDF-triple structure in the local knowledge repository of the agent, and the speech act (e.g., ‘suggest’, ‘commit’, ‘complain’, etc.) that characterizes the analyzed statement of the user. The speech act classification is done with https://github.com/bhavitvyamalik/DialogTag. For illustration, Fig. 5 shows some sample structures as provided by LAS. **The MT service** [7] ensures that the agent is able to converse with users who speak a language not covered by the language analysis/production modules.



bnId	"bn:00085719v"
ClauseType	"Dec"
end_string	"18"
finiteness	"FIN"
frame	"Residence"
id1	"4"
id2	"2"
lex	"stay_VB_01"
pbID	"01"
pbsenseID	"stay.01"
pos	"VBG"
sem	"stay"
speech_act	"Statement"
start_string	"11"
Tense	"Pres"
vn	"vn:47.1_stay"
vncls	"47.1"
wnId	"02727462v"

Fig. 5. Sample language analysis structures for “I am still staying with friends”.

Spoken language synthesis. The spoken language synthesis technologies comprise multilingual natural language generation (NLG) and text-to-speech synthesis (TTS). The NLG module is an extension of the multilayer FoG generation module [13], which uses the same types of structures as depicted in Fig. 5. We adapt FoG for dialogue generation and use of sentence templates for scenarios in which the reactions of the agent differ only in terms of provided data. For TTS, we use a flexible multilingual service that comprises different off-the-shelf Tacotron-based TTS applications [20]. Coqui.ai[4] pre-trained models are used for English, German, and Spanish. For Arabic, Catalan, and Greek, smaller off-the-shelf models [3,10,15] are currently being worked on.

5 Conclusions and Ongoing Work

We presented the first prototypical implementation of a personalized knowledge-based ECA. In contrast to the overwhelming majority of the state-of-the-art

[4] https://coqui.ai/.

ECAs, its interaction core consists of an agent planning service, which plans the overall interaction with the user (including actions not related to communication) and a dialogue management service, which plans the dialogue moves of the agent. This separation ensures that the agent is capable of also performing actions that are not directly related to communication – a prerequisite of a genuine personal assistant, which is expected to be not only communicative, but also intelligent. The first assessment of the functionality of the prototype by TCNs, NGOs and governmental institutions indicates that the information provided by the MyWelcome agent is useful, supports the TCNs in their concerns and alleviates the workload of NGO workers and officers. Formal evaluation trials are planned to obtain a more detailed picture on the performance of the agent. The future efforts will target the consolidation of the modules of MyWelcome agent and the extension of the topics in which the agent can support the TCNs.

Acknowledgements. Many thanks to our colleagues from the NGOs PRAKSIS and CARITAS, Generalitat de Catalunya, Mind Refuge, KEMEA, and ISocial Foundation for their guidance and constant constructive feedback!

References

1. Bohnet, B., Wanner, L.: Open source graph transducer interpreter and grammar development environment. In: Proceedings of LREC, pp. 211–218 (2010)
2. Fér, R., Matějka, P., Grézl, F., Plchot, O., Veselý, K., Černocký, J.: Multilingually trained bottleneck features in spoken language recognition. Comput. Speech Lang. **2017**(46), 252–267 (2017)
3. Halabi, N.: Modern standard Arabic phonetics for speech synthesis. Ph.D. Diss. (2016)
4. Helmert, M.: The fast downward planning system. J. Artif. Int. Res. **26**(1), 191–246 (2006)
5. Helmert, M., Röger, G., Karpas, E.: Fast downward stone soup: a baseline for building planner portfolios. In: ICAPS 2011 Workshop on Planning and Learning, pp. 28–35 (2011)
6. Janowski, K., Ritschel, H., Lugrin, B., André, E.: Sozial interagierende Roboter in der Pflege, pp. 63–87. Springer Fachmedien Wiesbaden, Wiesbaden (2018). https://doi.org/10.1007/978-3-658-22698-5_4
7. Junczys-Dowmunt, M., et al.: Marian: Fast Neural Machine Translation in C++. In: Proceedings of ACL 2018, System Demonstrations, pp. 116–121 (2018)
8. Karafiat, M., Karthick, B., Szoke, I., Vydana, H., Benes, K., Černocký, J.: BUT OpenSAT 2020 speech recognition system. In: Proceedings of the InterSpeech 2021 (2021)
9. Klusch, M., Kapahnke, P.: The ISEM matchmaker: a flexible approach for adaptive hybrid semantic service selection. J. Web Seman. **15**, 1–14 (2012)
10. Külebi, B., Öktem, A., Peiró-Lilja, A., Pascual, S., Farrús, M.: CATOTRON – a neural text-to-speech system in Catalan. In: Proceedings of the Interspeech, pp. 490–491 (2020)
11. Li, X., Chen, Y.N., Li, L., Gao, J., Celikyilmaz, A.: End-to-end task-completion neural dialogue systems. In: Proceedings of the IJCNLP, pp. 733–743 (2017)

12. Mencía, B.L., Pardo, D.D., Trapote, A.H., Gómez, L.A.H.: Embodied Conversational Agents in interactive applications for children with special educational needs. In: Griol Barres, D., Callejas Carrión, Z., Delgado, R.L.C. (eds.) Technologies for Inclusive Education: Beyond Traditional Integration Approaches, pp. 59–88. IGI Global, Hershey (2013)
13. Mille, S., Dasioupoulou, S., Wanner, L.: A portable grammar-based NLG system for verbalization of structured data. In: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, pp. 1054–1056 (2019)
14. Moro, A., Raganato, A., Navigli, R.: Entity linking meets word sense disambiguation: a unified approach. Trans. Assoc. Comput. Linguist. (TACL) **2**, 231–244 (2014)
15. Park, K., Mulc, T.: CSS10: a collection of single speaker speech datasets for 10 languages (2019)
16. Plchot, O., et al.: Analysis of BUT-PT Submission for NIST LRE 2017. In: Proceedings of Odyssey 2018 the Speaker and Language Recognition WS, pp. 47–53 (2018). https://www.fit.vut.cz/research/publication/11762
17. Povey, D., et al.: The Kaldi speech recognition toolkit. In: IEEE 2011 Workshop on Automatic Speech Recognition and Understanding, December 2011
18. Shvets, A., Wanner, L.: Concept extraction using pointer–generator networks and distant supervision for data augmentation. In: Keet, C.M., Dumontier, M. (eds.) EKAW 2020. LNCS (LNAI), vol. 12387, pp. 120–135. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-61244-3_8
19. Straka, M.: UDPipe 2.0 prototype at CoNLL 2018 UD shared task. In: Proceedings of the CoNLL 2018, pp. 197–207 (2018)
20. Wang, Y., et al.: Tacotron: Towards end-to-end speech synthesis (2017)
21. Wanner, L., et al.: KRISTINA: a knowledge-based virtual conversation agent. In: Demazeau, Y., Davidsson, P., Bajo, J., Vale, Z. (eds.) PAAMS 2017. LNCS (LNAI), vol. 10349, pp. 284–295. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59930-4_23
22. Wargnier, P., Carletti, G., Laurent-Corniquet, Y., Benveniste, S., Jouvelot, P., Rigaud, A.S.: Field evaluation with cognitively-impaired older adults of attention management in the Embodied Conversational Agent Louise. In: Proceedings of the IEEE International Conference on Serious Games and Applications for Health, pp. 1–8 (2016)
23. Wei, W., Le, Q., Dai, A., Li, L.J.: AirDialogue: an environment for goal-oriented dialogue research. In: Proceedings of the EMNLP, pp. 3844–3854 (2018)
24. Weisz, G., Budzianowski, P., Su, P.H., Gašić, M.: Sample efficient deep reinforcement learning for dialogue systems with large action spaces. IEEE/ACM Trans. Audio Speech Lang. Process. **26**(11), 2083–2097 (2018)