

Conversational Agents for Intelligent Buildings

Weronika Sieńska, Nancie Gunson, Christopher Walsh,
Christian Dondrup, and Oliver Lemon

School of Mathematical and Computer Sciences,

Heriot-Watt University,

Edinburgh, EH14 4AS, UK

{w.sieinska, n.gunson, c.dondrup, o.lemon}@hw.ac.uk
c.walsh.1@research.gla.ac.uk

Abstract

We will demonstrate a deployed conversational AI system that acts as a host of a smart-building on a university campus. The system combines open-domain social conversation with task-based conversation regarding navigation in the building, live resource updates (e.g. available computers) and events in the building. We are able to demonstrate the system on several platforms: Google Home devices, Android phones, and a Furhat robot.

1 Introduction

The combination of social chat and task-oriented dialogue has been gaining more and more popularity as a research topic (Papaioannou et al., 2017c; Pecune et al., 2018; Khashe et al., 2019). In this paper, we describe a social bot called Alana¹ and how it has been modified to provide task-based assistance in an intelligent building (called the GRID) at the Heriot-Watt University campus in Edinburgh. Alana was first developed for the Amazon Alexa Challenge in 2017 (Papaioannou et al., 2017b,a) by the Heriot-Watt University team and then improved for the same competition in 2018 (Curry et al., 2018). The team reached the finals in both years. Now Alana successfully serves as a system core for other conversational AI projects (Foster et al., 2019).

In the GRID project, several new functionalities have been added to the original Alana system which include providing the users with information about:

- the GRID building itself (e.g. facilities, rooms, construction date, opening times),
- location of rooms and directions to them,
- events happening in the building,
- computers available for use – updated live.

¹See <http://www.alanaai.com>

Currently, our intelligent assistant is available for users on several Google Home Mini devices distributed in the GRID – a large university building with multiple types of users ranging from students to staff, and visitors from business/industry. It is also available on Android phones via Google Actions as part of the Google Assistant. The system is reconfigurable for other buildings, via a graph representation of locations and their connectivity. It connects to live information about available resources such as computers and to an event calendar.

2 System Architecture

Figure 1 presents the architecture of the system.

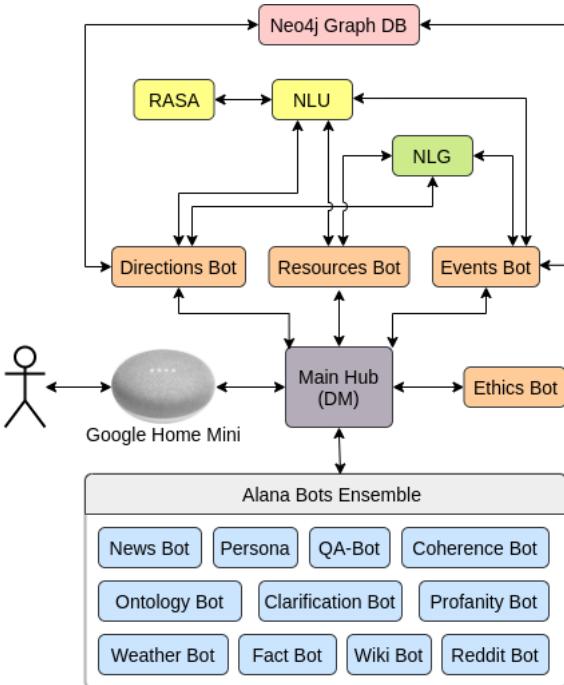


Figure 1: System architecture.

The Alana system is an ensemble of several different conversational bots that can all potentially

produce a reply to the user’s utterance. Each bot uses different information resources to produce its reply. Example resources are: Wikipedia, Reddit, many different News feeds, a database of interesting facts etc. Additionally, there are also conversational bots that drive the dialogue in case it has stalled, deal with profanities, handle clarifications, or express the views, likes, and dislikes of a virtual Persona. The decision regarding which bot’s reply is selected to be verbalised is handled by the Dialogue Manager (DM).

ASR/TTS In the GRID project, the audio stream is handled using the Google Speech API and the system is therefore also available as an action on Google Assistant on Android phones.

NLU In the Alana system, users’ utterances are parsed using a complex NLU pipeline, described in detail in (Curry et al., 2018), consisting of steps such as Named Entity Recognition, Noun Phrase extraction, co-reference and ellipsis resolution, and a combination of regex-based and deep-learning-based intent recognition. In the GRID project, an additional NLU module has been implemented for building-specific enquiries which uses the RASA² framework. In the Persona bot we use AIML patterns for rapid reconfigurability and control.

NLG The NLG strategy depends on the different conversational bots. It ranges from the use of complex and carefully designed templates to automatically summarised news and Wikipedia articles (Curry et al., 2018).

DM In every dialogue turn each of the bots attempts to produce a response. Which response will be uttered to the user is determined by a selection strategy which is defined by a bot priority list and can also be learned from data (Shalymnov et al., 2018). If several bots produce a response in a given turn, the response of the bot with the highest priority is selected. The Directions, Events, and Resources bots are placed at the top of the priority list in order to increase the chance of their responses being selected because all of them provide information regarding the GRID building. The Coherence Bot has the lowest priority since its role is to handle turns when none of the other bots produce a response so that the conversation can be driven forward.

²See <https://rasa.com/>

The Neo4j Graph Database Figure 2 presents the graph representation of how the rooms are connected with each other in the GRID building. To port the system to different buildings, this representation would be changed to reflect the locations and their spatial connectivity.

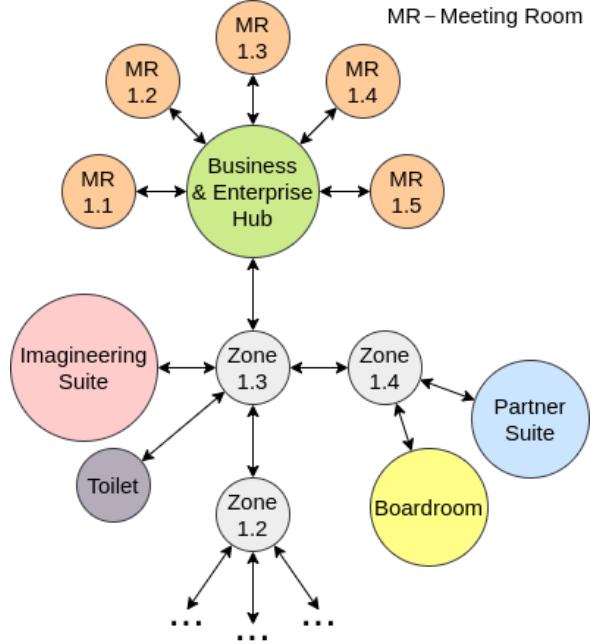


Figure 2: Graph representation of spatial connectivity – example for the GRID building.

3 Example Dialogues

General information about the GRID building is encoded in Persona’s AIML templates. It is this bot that responds to questions whose answers are generally fixed over time – for example about opening hours and where one can buy a cup of coffee. For examples of responses generated by Persona, see Dialogue 1.

The Ethics Bot has been implemented to obtain users’ consent for storing and processing transcriptions of their conversations. The request for user consent is omitted in the example dialogues.

The Directions Bot handles requests for locations of rooms and directions to them. It connects to the Neo4j graph database where it retrieves the shortest paths. The Events Bot responds whenever events information is requested for either a specific room or the building in general. The information is stored in the Neo4j database as well where the events are assigned to particular rooms. It is retrieved from a university timetable system and updated every day.

A: Welcome to the GRID building! How can I help you today?

U: What does GRID stand for?

A: (persona) GRID stands for Global, Research, Innovation and Discovery. Is there anything else I can help you with?

U: What time does this building close?

A: (persona) This building never closes. It is open 24/7. Did you know that you can have a chat with me about music? Also, if you're looking for a specific room, you can ask me how to get there.

U: Actually, I really need a cup of coffee. Do you know where I can get one?

A: (persona) You can have a cup of coffee at the cafe on the ground floor. It looks like a tuk tuk. Can I help you with anything else?

U: What is a tuk tuk?

A: (wiki.bot) I think that an auto rickshaw is a motorized version of the pulled rickshaw or cycle rickshaw. Most have three wheels and do not tilt. They are known by many terms in various countries including auto, baby taxi, bajaj, chand gari, lapa, tuk-tuk, 3wheel or tukxi.

U: Wow, that's interesting. Thanks, goodbye!

A: It was lovely talking to you! Goodbye!

Dialogue 1: Questions about the GRID building. U – User, and A – Alana system.

The role of the Resources Bot is to provide the user with information about how many computers are currently free in the GRID building's labs. The bot scrapes the data live from a university website. Dialogue 2 presents example responses produced by the Directions, Events, and Resources Bots.

Dialogue 3 shows an example of social chat one can have with our system. Alana offers conversation about various topics such as music, movies, and sports. It provides current news as well.

4 Future Work

4.1 Deployment on a Robot

We can already deploy Alana on a Furhat³ robot during special events in the GRID building. We believe that having Alana deployed on a robot makes it even more attractive and encourages more people to interact with it.

This will also be useful for a future project which will aim to further extend Alana's functionalities so that for example, it can become a virtual receptionist in the National Robotarium (a smart research building for Human-Robot Interaction) which is soon to be built at the Heriot-Watt University campus in Edinburgh.

³See <https://www.furhatrobotics.com/>

A: Hello, and welcome to the GRID building! What can I assist you with?

U: Hi! What can you do?

A: (persona) If you're looking for a room, I can tell you where it is. I can also tell you what events are happening in this building and where you can find a free computer.

U: Are there any available computers in the Flex Lab?

A: (resources.bot) Sorry, there is Flex Lab 1 and Flex Lab 2 in this building. Which one do you mean?

U: Number 2.

A: (resources.bot) I'm afraid there are no free computers in the Flex Lab 2 at the moment, but I know that the room Digital Lab has the most free computers, it has 97 PCs available.

U: Where is the Digital Lab?

A: (directions.bot) The Digital Lab is on the first floor. Would you like directions?

U: Yes, please.

A: (directions.bot) From the entrance, head across the atrium to the right and take the accessible lift to the first floor, then turn right, the door to the Digital Lab will be in front of you. Don't forget that you can ask me about events happening in this building.

U: Ok, is there anything on today?

A: (events.bot) There are 2 events taking place in the Boardroom today. An event called Negotiate Training is scheduled from 10:30 until 12:30. Another event called Virtual Reality Workshop is taking place from 13:15 until 19:15. Is there anything else I can help you with?

U: No, I'm good. Thanks, goodbye!

A: It was lovely talking to you! Goodbye!

Dialogue 2: Task-based dialogue. U – User, and A – Alana system.

A: Welcome to the GRID building! How can I help you today?

U: Let's talk about music.

A: (coherence.bot) I'm into music. I'm really into rap and rock. What about you?

U: I like Bob Dylan a lot.

A: (news.bot) A friend of mine told me about this news headline from MTV News about Bob Dylan. Timothée Chalamet will play Bob Dylan in his next Oscar worthy role. Do you want to hear some more about this?

U: Who is Timothée Chalamet?

A: (wiki.bot) As far as I know, Timothée Hal Chalamet is an American actor.

U: What movies do you like?

A: (persona) My favourite movie is Star Wars. I love the two androids in that film. Which movie do you like?

U: The new Star Trek movies are pretty good.

A: (news.bot) Related to Star Trek, I heard this news headline from a friend, 'Star Trek: Picard': Everything you need to watch before the new CBS series. Do you want to learn more?

U: No, thanks. I've gotta go. Bye!

A: It was lovely talking to you! Goodbye!

Dialogue 3: Open-domain chat about music, movies, and news. U – User, and A – Alana system.

4.2 Graphical App

Furthermore, we collaborate with Speech Graphics⁴ which is a company that provides facial animation for speaking game characters from audio input. Together we aim to create a mobile app with Alana and its animated avatar which will be available to students and academics at Heriot-Watt University in Edinburgh. Figure 3 presents two of the avatars available for Alana. We believe that the interaction with the graphical app will be more appealing for users than talking to the Google Home devices.



Figure 3: Example Speech Graphics avatars.

4.3 Evaluation

We are conducting experiments where we compare two versions of the developed system. One of them is the full version of the Alana-GRID system implemented in this project and the other is Alana deprived of its open-domain conversational skills i.e. only capable of providing information about the GRID building which the user requests. Our hypothesis is that open-domain social chat adds value to virtual assistants and makes it more pleasurable and engaging to talk to them.

Acknowledgments

This work has been partially funded by an EPSRC Impact Acceleration Award (IAA) and by the EU H2020 program under grant agreement no. 871245 (SPRING)⁵.

References

Amanda Cercas Curry, Ioannis Papaioannou, Alessandro Suglia, Shubham Agarwal, Igor Shalyminov,

Xinnuo Xu, Ondřej Dušek, Arash Eshghi, Ioannis Konstas, Verena Rieser, et al. 2018. Alana v2: Entertaining and informative open-domain social dialogue using ontologies and entity linking. *Alexa Prize Proceedings*.

Mary Ellen Foster, Bart Craenen, Amol Deshmukh, Oliver Lemon, Emanuele Bastianelli, Christian Dondrup, Ioannis Papaioannou, Andrea Vanzo, Jean-Marc Odobez, Olivier Canévet, Yuanzhouhan Cao, Weipeng He, Angel Martínez-González, Petr Motlicek, Rémy Siegfried, Rachid Alami, Kathleen Belhassen, Guilhem Buisan, Aurélie Clodic, Amandine Mayima, Yoan Sallami, Guillaume Sarthou, Phani-Teja Singamaneni, Jules Waldhart, Alexandre Mazel, Maxime Caniot, Marketta Niemelä, Päivi Heikkilä, Hanna Lammi, and Antti Tammela. 2019. MuMMER: Socially Intelligent Human-Robot Interaction in public spaces.

Saba Khashe, Gale Lucas, Burcin Becerik-Gerber, and Jonathan Gratch. 2019. Establishing social dialog between buildings and their users. *International Journal of Human-Computer Interaction*, 35(17):1545–1556.

Ioannis Papaioannou, Amanda Cercas Curry, Jose Part, Igor Shalyminov, Xu Xinnuo, Yanchao Yu, Ondřej Dušek, Verena Rieser, and Oliver Lemon. 2017a. An ensemble model with ranking for social dialogue. NIPS 2017 Conversational AI Workshop; Conference date: 08-12-2017 Through 08-12-2017.

Ioannis Papaioannou, Amanda Cercas Curry, Jose L. Part, Igor Shalyminov, Xinnuo Xu, Yanchao Yu, Ondřej Dušek, Verena Rieser, et al. 2017b. Alana: Social dialogue using an ensemble model and a ranker trained on user feedback. *Alexa Prize Proceedings*.

Ioannis Papaioannou, Christian Dondrup, Jekaterina Novikova, and Oliver Lemon. 2017c. Hybrid chat and task dialogue for more engaging HRI using reinforcement learning. In *26th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2017, Lisbon, Portugal, August 28 - Sept. 1, 2017*, pages 593–598. IEEE.

Florian Pecune, Jingya Chen, Yoichi Matsuyama, and Justine Cassell. 2018. Field trial analysis of socially aware robot assistant. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS ’18, page 1241–1249, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

Igor Shalyminov, Ondřej Dušek, and Oliver Lemon. 2018. Neural response ranking for social conversation: A data-efficient approach. In *Proceedings of the 2018 EMNLP Workshop SCAI: The 2nd International Workshop on Search-Oriented Conversational AI*, Brussels, Belgium. Association for Computational Linguistics.

⁴See <https://www.speech-graphics.com/>

⁵See <http://spring-h2020.eu/>