Potentials of Chatbot Technologies for Higher Education: A Systematic Review

Stefan Ullmann
University of Hohenheim, stefan.ullmann@uni-hohenheim.de

Mareike Schoop
University of Hohenheim, schoop@uni-hohenheim.de

Potentials of Chatbot Technologies for Higher Education: A Systematic Review

Stefan Ullmann, Mareike Schoop
Information Systems Group 1,
University of Hohenheim,
70599 Stuttgart, Germany
{stefan.ullmann, schoop}@uni-hohenheim.de

Abstract

Chatbots are used in different areas such as customer service, healthcare and education. The potential for improving outcomes and processes in education is high but differs for different types of chatbots. As universities want to provide excellent teaching, it is important to find the chatbot technologies with the greatest possible benefit. This paper presents a systematic review of chatbot technologies in five application areas. For each application area, the ten most cited publications are analysed and a possible categorisation scheme for chatbot technologies is derived. Furthermore, it is investigated which chatbot technology types are used and their suitability for higher education is analysed. The results show that chatbots can be categorised using five categories derived from the 50 publications. A total of 14 different types of chatbot technologies are found in the five areas. Nine of them are suitable for use in higher education.

Keywords: artificial intelligence, chatbot, chatbot technologies, categorisation, conversational agent, higher education, natural language processing, systematic review

1.0 Introduction

Universities want to provide excellent teaching and a high-quality learning environment but often lack a clear vision to pursue these goals (Fook and Sidhu 2015). In recent years, mental health and wellbeing have become increasingly important for the future of teaching and learning in higher education (Horizon Report 2020).

Adaptive technologies and the use of artificial intelligence (AI) and machine learning techniques in education applications are considered as some of the emerging technologies with the greatest impact on teaching and learning in higher education (Horizon Report 2020). Chatbots are AI systems that understand human language by Natural Language Processing (NLP). They can respond like intelligent entities and the conversation can be via text or speech (Khanna et al. 2015). There is a growing interest and use of chatbots, also called conversational agents, in several areas including customer service, healthcare and business in general (Adamopoulou and Moussiades 2020a; Wang et al. 2021).

In education, chatbots have started to be applied, but they are still at an early stage of implementation. The potential to improve the learning processes and outcomes through their usage is high (Winkler and Söllner 2018). Different types of chatbots have different influences on learning outcomes and processes. Therefore, it is important to identify and apply the most useful chatbots for a particular educational scenario and to focus research on their added value (Wang et al. 2021; Winkler and Söllner 2018). As different technologies are used in different fields (Adamopoulou and Moussiades 2020a; Hussain et al. 2019), it should be analysed which of the diverse chatbot technologies from different application fields can be used in higher education to have the greatest possible benefit. Consequently, the three research questions of the present paper are the following:

- RQ1: How can chatbot technologies be categorised?
- RQ2: What types of technologies do chatbots use?
- RQ3: What chatbot technologies are suitable for higher education?

To answer these questions, a keyword search is conducted in three of the most important databases for chatbot technologies (Dyba et al. 2007; Kitchenham and Charters 2007) leading to categories for chatbot technologies. Subsequently, the publications found are assigned to the derived categories and each unique category combination then defines a chatbot technology type. Finally, it is discussed which of these chatbot technology types are suitable for the use in higher education.

The remainder of this work is structured as follows: Section two provides the theoretical background containing an overview of chatbot technologies and the requirements for chatbots in higher education. Section three describes the applied methodology and the conducted literature search. Section four introduces the categorisation for chatbots and the different types of chatbot technologies. In section five the suitability of the different types is discussed. Finally, a summary of contributions and an outlook concludes this paper.

2.0 Theoretical Background

The following section describes general chatbot technologies and their main components, as well as important terminology. Furthermore, the requirements for chatbots in higher education are provided in the form of seven principles of good practice and use cases extracted from the literature. These requirements form the basis for assessing the suitability of chatbot technologies in higher education.

2.1 Overview of Chatbot Technologies

Chatbots, also referred to as conversational agents or chatterbots, are conversational interfaces that enable users to interact with software using natural human language (Neumann et al. 2021; Wang et al. 2021). They are able to communicate with the user in natural language via audio or messaging methods (Thorat and Jadhav 2020). This definition includes all kinds of software that enable humans to have a conversation with a computer and, therefore, contains dialogue systems, conversational agents, personal assistants and voice-control interfaces (Serban et al. 2017).

The first chatbot was called Eliza and used pattern matching algorithms and sentence reconstruction to simulate a psychotherapist. It was created in 1964 and even without in-depth knowledge or NLP influenced the development of many other bots (Reshmi and Balakrishnan 2018; Weizenbaum 1983; Winkler and Söllner 2018).

The main components of a chatbot are the user interface, the user messages analysis component, the dialogue management component, the backend, and the response generation component (Adamopoulou and Moussiades 2020a; Adamopoulou and Moussiades 2020b).

The process of the chatbot usually starts with receiving a user input (usually speech or text) via the user interface (e.g. “Where is Room HS11?”) (Thorat and Jadhav 2020). After the chatbot receives the user input, it is forwarded to the user message analysis component. This component analyses the user’s intention (e.g. “Show location”) and existing entities in the input (Adamopoulou and Moussiades 2020b).

The dialogue management component is used to enable the chatbot to remember and update the context of the conversation; it thus saves the actual intent and the entities identified by the user message analysis component (Kucherbaev et al. 2018). To fulfil the user's intention, the chatbot needs appropriate information which is retrieved from the backend, its data source, through Application Programming Interface (API) calls or database queries (Adamopoulou and Moussiades 2020b; Kucherbaev et al. 2018).

In the response generation component, responses are generated using a rule-based, retrieval-based or generative-based method. In some chatbots, multiple approaches are used (Adamopoulou and Moussiades 2020b). Rule-based approaches search their knowledge base, which consist of pattern and templates, to generate the response, hence these systems do not create new responses. When the chatbot receives an input which matches a pattern, the chatbot will response with the message stored in the template. Therefore, such chatbots make no grammatical errors and give quick responses, but are not capable of handling unknown conditions or inputs. They usually only take the last user input into account and their responses can be repetitive, but they are suitable for closed domain communications (Adamopoulou and Moussiades 2020b; Wu et al. 2017). Creating the knowledge base can often be tedious and time-consuming (Wallace 2009).

Retrieval-based approaches select their response from a data source, which is a repository of responses. They select the best answer using algorithms, machine learning and neural networks. The machine learning and the neural networks are used to score the stored responses and to select the most likely response from the dataset (Adamopoulou and Moussiades 2020b; Ramesh et al. 2017; Wu et al. 2016). They can also use a Named-Entity Recognition to classify entities in the user input. Named-Entity Recognition processes documents, identifies entities e.g. companies or places, and classifies these in a second step in predefined categories (Mansouri et al. 2008). Retrieval-based technologies are more flexible than the rule-based approaches, as they are able to query and analyse resources using APIs (Hien et al. 2018). They do not create new responses and make no grammatical errors (Ramesh et al. 2017)

Generative approaches do not select their answer from a set, but take the input word by word and generate matching answers (Thorat and Jadhav 2020). They can apply techniques that are used for machine translation, but instead of translating the input in another language, they translate the input into an output. This can be done using different machine learning techniques. In general, such generative approaches are “smarter” and make the conversation more human-like but need a huge amount of training data and they can make grammatical errors. They are not ideal for closed domain communication but are useful for informal open domain conversations (Adamopoulou and Moussiades 2020b; Ramesh et al. 2017). Furthermore, there are hybrid approaches which combine different response generation methods (Lokman and Ameedeen 2019)

To be able to use voice input or give voice output, speech recognition and speech synthesis can be used. Speech recognition is the process of creating a sequence of words from a speech signal (Gaikwad et al. 2010). Today, speech-to-text methods are widespread and using these technologies has become quite easy (Shakhovska et al. 2019). Speech synthesis is the generation of speech by a computer and is also known as text-to-speech, i.e. the process of transferring text into speech (Dharwadkar and Deshpande 2018).

2.2 Seven Principles of Good Practice in Higher Education

Chickering and Gamson (1987) introduce seven principles of good practice in undergraduate teaching. To realise the full potential of communication and information technologies for higher education, they should be in line with the seven principles (Ehrmann and Chickering 1996). These seven principals are described in the following. Good practice *encourages contact between students and faculty* inside and outside the classroom, as this is the most important factor in student motivation and engagement. Further, it *facilitates cooperation and develops reciprocity among students*, as good learning is collaborative and social. One way to encourage cooperation is to assign students to learning groups of five to seven students. The third principle is to *encourage active learning* as students need to integrate what they have learned into themselves. (Chickering and Gamson 1987).

Additionally, good practice *gives prompt feedback* as students need appropriate feedback on their performance to benefit from the courses (Chickering and Gamson 1987). To be effective, feedback should be non-judgmental, supportive, timely and specific (Shute 2008). The fifth principle of good practice is to *emphasise time on task*, as there is no substitute for time on task and good time management is critical for students. To *communicate high expectations* is the sixth principle, as having high expectations of students can become a self-fulfilling prophecy. To *respect diverse talents and ways of learning* is the last of the seven principles of good practice, as students need to learn in ways that are appropriate for them (Chickering and Gamson 1987).

2.3 Chatbot Usage in Higher Education

There are several use cases for chatbots in higher education. One use case is answering questions on the course content, via e-mails. In case the chatbot has fitting responses, it should answer and let the instructor reply otherwise. The chatbot should learn from the instructor's responses (Buyrukoğlu and Yılmaz 2021).

Helping students learn factual knowledge is another use case. The chatbot should be dialogue-based, adaptive and support students in learning factual knowledge in different subjects. For this purpose, the chatbot gives answer options, recognises the correctness of answers and should be able to give positive reinforcement feedback (*"That's right!"*). It should also have a casual chat mode which provides fun facts or jokes (Ruan et al. 2019).

Chatbots can be used to replace the manual customer service of a university. Therefore, a chatbot should be able to answer questions related to the university, navigate users to the registration page, show the map of the university and upcoming events. Additionally, it should provide feedback regarding the university and be able to make appointments (Ahmad et al. 2019).

Helping students to select specialized courses out of several choices is another use case for chatbots. Further the chatbot should be able to analyse and provide possible course selections that can be adapted by the student. At the end, the student should receive a list of personalised recommendations (Chun Ho et al. 2018).

Chatbots are also used in the admission service domain, to ensure the applicant satisfaction, to support the work of admission stuff, and to reduce the time applicants must wait for information. The chatbot should be able to store new data if unable to answer a question (Agus Santoso et al. 2018).

A further use case is providing the users with health-promoting strategies based on cognitive behavioural therapy, positive psychology, and mindfulness techniques. The purpose is to improve well-being through conversational dialogue and audio-visual learning material. The chatbot is not a substitute for professional help of a therapist (Gabrielli et al. 2021).

Another use case is the support with synchronous meaningful learner's interaction in online courses as this is conducive to learner's persistence. The chatbot should give humorous feedback and needs a social aspect, as social interaction between chatbot and the learner can enhance the feeling of being with others. (Baylor and Kim 2009). It should be personalised for each learner as this has a positive impact on learner performance and would facilitate interaction in a more natural way (Dabbagh and Kitsantas 2012; Song et al. 2017).

3.0 Methodology

To fulfil the purpose and answer the specific research questions, a systematic literature review with a sequential approach was conducted which consisted of the definition of the scope, the literature search and the analysis and synthesis of the literature (vom Brocke et al. 2009).

3.1 The Scope Definition and Literature Search

As figure 1 shows, the approach of this work is to derive categories of chatbot technologies applied in different domains, to identify different technologies using these categories and to analyse their suitability for higher education usage.

To this end, a keyword search was conducted with often used synonyms for "chatbot", selected from other review publications on the topic (Io and Lee 2017; Winkler and Söllner 2018; Paschoal et al. 2020). After reviewing the set of search results, from application areas found in the literature, the following five were found to be most relevant to this work: *Education*, *Healthcare*, *Customer Service*, *Business* and *Psychology*. These results are broadly consistent with the findings of Wang et al. (2021).

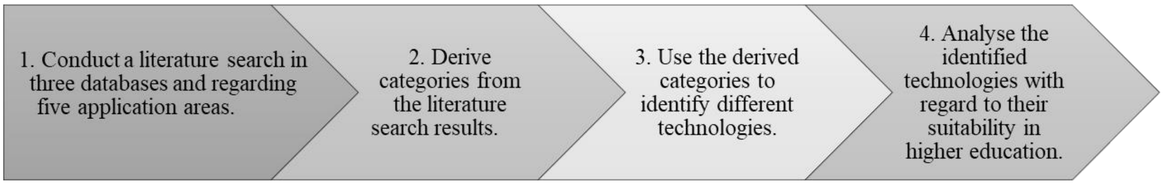

Figure 1. Four Step Approach of this Work.

Most of the publications related to chatbots and conversational agents are in computer science (Io and Lee 2017). According to Dyba et al (2007) and Kitchenham and Charters (2007), the following three databases can be considered as some of the most relevant sources for computer science and were thus be used: ACM Digital library, IEEExplore, Scopus (Kitchenham and Charters 2007; Dyba et al. 2007).

The terms *chatbot*, *chat bot*, *conversational agent*, *dialogue system*, *chatterbot* were searched in the abstract, title and keywords. The search was repeated for each of the five application areas and the ten most cited publications of each area were selected.

Highly cited publications in a field can identify essential clues and can be representative for latest changes (Cheng et al. 2020; Lai 2020). Therefore, a representative coverage is provided, by including the most cited publications in the most important application areas of chatbots.

Due to the rapid development in the fields of NLP, neural networks, machine learning (Winkler and Söllner 2018) and the limited number of articles selected in the different application domains, the search was limited to publications from 2015 onwards. Further the results were limited to available publications in English or German.

3.2 Literature Selection Process

When the search was conducted, the result of each query in each searched database was ordered by the number of citations. Then the ten most cited articles for each of the five application areas in all databases are selected. For the selection, the title and abstract were checked for the inclusion and exclusion criteria. Finally, the full text was read to ensure the paper fits the requirements. The search was conducted on 21 July 2021.

In addition to the general criteria described earlier, such as language and availability, it is important that the publications are relevant to this work. For the purposes of this paper, publications were included which:

- describe a system to communicate with a user in natural language via audio or messaging methods (Thorat and Jadhav 2020);
- describe the used chatbot technologies on a functional level;
- describe the context that the chatbot technology is used in.

Consequently, publications were excluded which:

- present chatbot technologies without describing their application;
- describe chatbot technologies only through their used chatbot framework;
- reviews and meta-analyses are excluded.

The 50 most cited publications that fit the inclusion and exclusion criteria are analysed in terms of the technologies they use at the functional level.

4.0 Search Results and Categorisation of Chatbot Technologies

During the literature search, 335 publications were scanned, 268 in the Scopus database, 42 in the ACM database, and 25 in the IEEExplore database. For the analysis of chatbot technologies. Seven papers were selected from ACM with an average of 10 citations, one publication from IEEExplore with three citations and 42 papers were selected from the Scopus database with an average of 23 citations. Hence, a total of 50 publications were selected, ten from each area of application. An overview of the papers is given in Appendix 1. Information on the total number of hits per database and area can be found in the round brackets, the number of selected publications from each database in square brackets, and the selected publications per search term and database on the arrows in Figure 2.

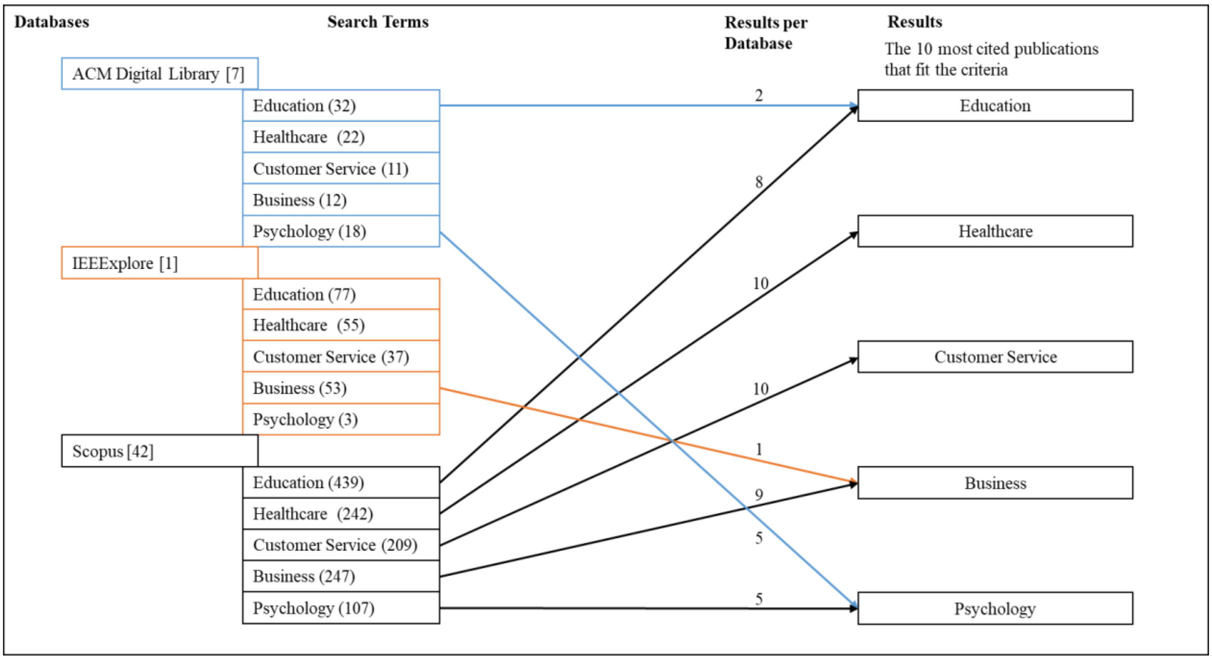

Figure 2. Overview of Search Results for each Search Term in all three Databases.

4.1 Categories of Chatbot Technologies

The results of the literature review are now used in the following to answer RQ1, *"How chatbot technologies can be categorised?"*. Chatbot categorisations can be based on several criteria or design techniques (Ramesh et al. 2017). The 50 publications that resulted from the search were analysed and combined with knowledge from the existing chatbot literature to derive a possible categorisation of chatbot technologies. In addition to being able to derive them from the literature, the categories are chosen to have meaningful implications for the use of the technologies in higher education.

It is possible to categorise chatbots according to their response generation method as *rule-based*, *retrieval-based* and *generative* chatbots (Hien et al. 2018). Generating responses is one main function of chatbots and each of the methods has its own advantages and disadvantages for their use in higher education.

Rule-based systems do not create new responses; they use pre-written conversational patterns to replay. They choose their response based on a predefined set of rules. Therefore, they make no grammatical errors and give fast responses, but are not capable of handling unknown conditions. Most of them only take the last user input into account and their responses can be repetitive (c.f. section 2.1.5).

Retrieval-based chatbots also select their output from a dataset and, therefore, make no grammatical errors. They use machine learning and neural networks to give the stored responses scores and to select the most likely response from the dataset. Other *retrieval-*

based approaches use algorithms, while simple ones select response based on keywords in the input. *Retrieval-based* chatbots can use Named-Entity Recognition to classify entities in the user input (c.f. section 2.1).

Generative chatbots synthesise the replay using neural networks and various machine learning techniques. They apply techniques that are usually used for machine translation, but instead of translating the user input in a different language, they translate it into a response. The advantages of generative chatbots are more intelligent responses and more human-like conversations. Disadvantages are the large amount of training data needed and the possibility of making grammatical errors. This makes them not ideal for closed domain conversations, but useful for informal conversations in open domains (c.f. section 2.1).

Some chatbots are based on more than one of these categories or use different additional types of response generation. They can be called *hybrids*, as they combine different technologies for the response generation e.g. *rule-based*, and *retrieval-based* techniques (Bharti et al. 2020).

Another category is the duration of the interaction. Depending on the duration of the interaction, the chatbot must use different technologies and is useful for different use cases in higher education. *Short-term interactions* in this case describe interactions that end after the chatbot has completed a certain task or after the interaction is ended. The next interaction between the user and the chatbot starts from scratch. *Long-term interaction* technologies must be able to use any long-term data and distinguish between different users. The chatbot should be able to personalise its responses or interactions based on previous interactions. Long-term data is needed to personalise responses or interactions, and user identification is important to find the right data or interactions for a particular user.

Human-assisted technologies can give consistent, detailed and individual answers, while fully automated systems may provide irrelevant and unsatisfactory answers (Buyrukoğlu and Yılmaz 2021). Some *non-human-assisted* technologies only interact with the user if they have an appropriate response (Xue et al. 2018). In addition to the differences in responses, the technologies also differ, as human-assisted technologies need a way to interact with the human assisting them, e.g. through a special interface. Most of the found chatbots work autonomously and are *non-human-assisted* while responding to user input. *Human assistance* is always given, when a human supports the chatbot in some way in its interaction process with the user.

The mode of interaction is one possibility to categorise chatbots, it can be divided into *voice*-based or *text*-based chatbots (Hussain et al. 2019). But not all chatbots that are able to understand *voice*-based input also respond with *voice*-based output. Therefore, the categorisation needs to be more detailed and is split in this work into output and user input. To be able to understand *voice*-based input, the chatbot needs technologies that differ from those of chatbots using only *text*-input. Additionally, to *voice*-based input and *text*-based input, a few chatbot technologies are found that can use *video*-based input or *buttons* as an input option.

Text- and *voice*-based outputs combined lead to higher memorisation of information compared to purely *text*-based outputs (Kock 2016). In addition, a *voice*-based output can create a warmer relationship between the user and the chatbot (Novielli et al. 2010). Due to these differences, the output is an important category. The response generation methods described in section 2.1 describe ways to generate *text*-based output. To be able to give *voice*-based output, chatbots need to apply technologies that allow them to generate speech out of the text. Further they can use techniques to make the *voice* experience more interactive and robust (Bharti et al. 2020).

In all of the categories (e.g. Output), each chatbot technology falls into at least one of the criteria (e.g., Text). Figure 3 shows the resulting categorisation.

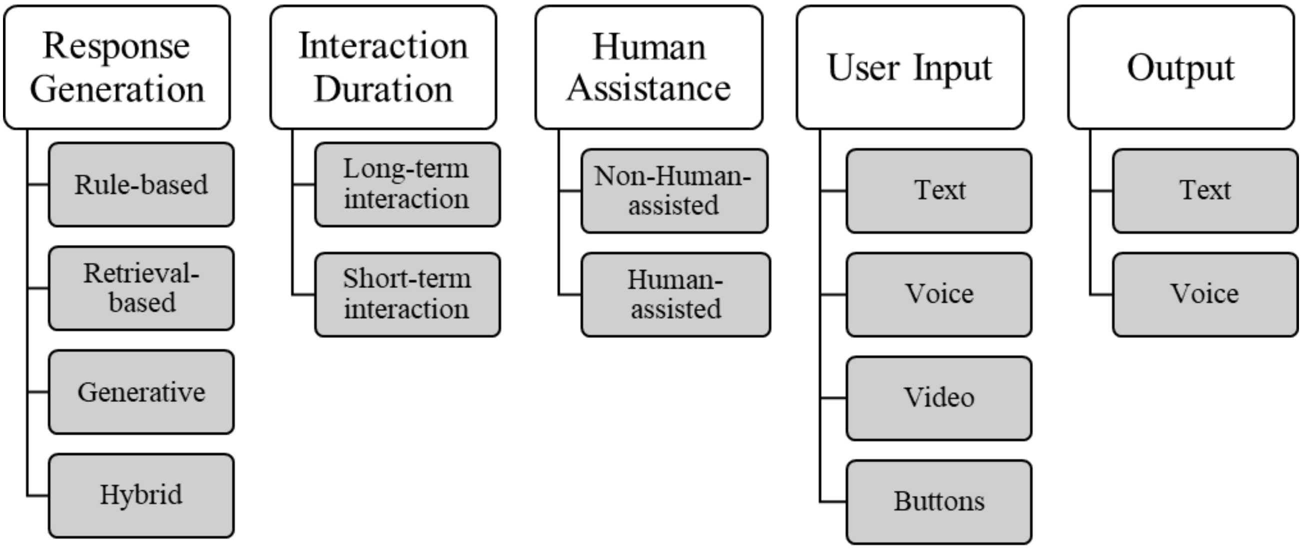

Figure 3. Categorisation Scheme for Chatbot Technologies.

4.2 Different Types of Chatbot Technologies

Following the categorisation of chatbot technologies, the 50 publications found in the five application areas education, healthcare, customer service, business, and psychology were analysed in terms of their used types of chatbot technologies. For this purpose, each described chatbot technology in each publication is assigned to the corresponding categories. Each unique combination of categories then forms a type of chatbot technology. E.g., technology type 13 in table 1 is a *hybrid*, *short-term*, *non-human-assisted* technology with *text in- and output*, because of its ability of using *voice-based in- and outputs*, 14 is defined as a further type of chatbot technology. The result of assigning the found publications to the corresponding categories are the different chatbot technologies shown in table 1. Therefore, the result of this process answers RQ2, “*What types of technologies do chatbots use?*”.

As table 1 shows, a total of 14 types of chatbot technologies were found in all five searched domains. Most of them use rule-based response generation, which is consistent with other chatbot literature (Liu and Mei 2020).

	Response Generation				Interaction Duration		Human Assistance		User Input				Output	
	Rule-based	Retrieval-based	Generativ	Hybrid	Short-term	Long-term	Not Human assisted	Human assited	Text	Voice	Video	Buttons	Text	Voice
1.	x					x		x	x	x			x	
2.	x				x		x		x		x		x	
3.	x					x	x		x				x	
4.	x				x		x		x			x	x	
5.	x				x		x		x				x	
6.	x				x		x		x	x			x	x
7.		x			x			x	x				x	
8.		x			x		x		x				x	
9.			x		x		x		x	x			x	
10.			x			x	x		x	x	x		x	
11.			x		x		x		x				x	
12.				x		x	x		x				x	
13.				x	x		x		x				x	
14.				x	x		x		x	x			x	x

Table 1. Overview of Different Types of Chatbot Technologies

Only four of the found technologies are able to generate their responses or interact based on previous interactions with the user. Rule-based and retrieval-based technologies are not able to generate new responses that are only based on the context of the conversations. In contrast, generative response generation technologies can create responses that fit the given context and information. The rule-based approaches are only able to select a predefined question that asks e.g. for knowledge that a learner has not talked about (Tegos et al. 2015), adapt their interaction, to give a pre-written response if someone has not been interacting in a group chat for a while (Wang et al. 2018) or they can start a sub dialogue with scaffolding questions, explanations and feedback if an incorrect input is given (Winkler et al. 2020). The hybrid response generation technology, that can have long-term interactions, is able to select recommendations on the long-term preferences of a user (Sun and Zhang 2018). The short-term hybrid technologies are using a retrieval-based approach, in case of no appropriate response from the rule-based approach (Bharti et al. 2020). Conversations are stored in a database and unanswered question can be used to train the chatbot (Satu et al. 2015).

Most of the chatbots are autonomous and do not use any human assistance. The rule-based response generation technology is assisted by a human by giving the human the possibility to regulate the behaviour of the chatbot, for example by setting rules (Tegos et al. 2015). The human-assisted retrieval-based chatbot technology is able to decide whether there is a fitting response in his response dataset; if not, the request is passed on to a human specialist for answering the query. Correct responses can be used as training data for the chatbot (Paikens et al. 2020). No generative response generation technology is human-assisted. The generative technology is not able to decide whether the response that is given is helpful for the user, as it translates inputs into output. Therefore, the chatbot is not able to hand a question it is not able to answer in a helpful way to a human.

As chatbots are generally able to communicate with the user in natural language via audio or messaging methods (cf. section 2.1), it is no surprise that all found chatbot technologies can use text-based input and some of them additionally use voice-based input for the interaction with the user. The same is true for the output. The fact that five chatbots are able to understand voice-based input but only two give voice-based output could be due to the reason that the usage of speech-to-text methods are widespread, and their application has become quite easy (c.f. section 2.1).

Only two of the found technologies use video-based input (Oh et al. 2017). The rule-based technology uses the video-based input, only to decide if the user is hurt or not. This information is then used to decide whether a pre-defined response for a hurt user should be provided or the normal pre-defined dialogue flow should be followed (Carranza et al. 2019). The only technology that uses buttons as user input uses rule-based technology to generate answers, but also includes machine learning. The machine learning part is only used to check the correctness of user answers for given questions, which is why the technology is not categorised as a hybrid technology (Ruan et al. 2019). That a rule-based technology uses buttons is not surprising, since an answer can be given based on a rule when the user presses a button.

5.0 Suitable Chatbot Technologies for Higher Education

Previously, the results of the literature review were analysed for the different types of chatbot technologies used, which are summarised in table 1. The 14 types of technologies in the figure are analysed in the following for their suitability for higher education. For this purpose, it is discussed how the different technologies could be helpful to fulfil the requirements described in section 2.2 and 2.3.

5.1 Discussion of the Suitability for Higher Education

For providing users with health-promoting strategies (cf. section 2.3), a rule-based approach is suitable, as the different strategies can be stored as pre-written responses and can be selected by rules. A chatbot technology using video-based user input can be suitable (2), as the strategies could be selected based on the mood of the person. The audio tracks and media files can be made available via URLs, in the pre-written strategies. As the technology should not be a substitute for a professional therapist, human assistance is not required (c.f. section 2.3). The strategies can be selected based on the video input and the actual emotions of the user (c.f. section 4.2); thus no long-term interaction technology is needed.

To develop reciprocity and cooperation among students, a principle of good practice (c.f. section 2.2), a rule-based, long-term interaction technology could be used (3). The long-term technology, that is able to analyse the amount of communication in a group chat and send a reminder to users who have not interacted for an extended period of time (c.f. section 4.2), can help to encourage students to collaborate.

The chatbot technology (4) is suitable for the use case of factual knowledge learning, where the chatbot technology should be able to recognise the correctness of an answer, give feedback and further include a casual chat mode to provide jokes or fun facts. The casual chatbot mode can be selected using buttons and the jokes and fun facts could be provided random via keywords e.g. “joke”. For some types of feedback, (4) is also a suitable technology. The buttons enable the user to easily ask for further feedback or explanations and the rule-based approach selects pre-written answers (c.f. section 2.3 and 4.2).

To support learners with meaningful synchronous interactions in online courses, the chatbot should provide a feeling of being with others, and be able to personalise interaction with the learners. Further it should be able to give humorous feedback if needed (c.f. section 2.3). For this use case a technology with voice-based output is suitable, as it can create a warmer relationship between the user and the chatbot, and therefore could help providing a feeling of being with others. Further a combination of text-and voice-based output can help to better memorise information (cf. section 2.1). A rule-based approach can be able to personalise interactions as scaffolding questions can be asked and explanations can be given if the learner needs them (c.f. section 4.2). Further it can provide humours positive feedback, depending on the pre-defined dialogues. Therefore, chatbot technology (6) is suitable for this purpose.

Promoting student-faculty contact is the first principle of good practice (c.f. section 2.3). For this purpose, human-assisted technologies are suitable, as autonomously working chatbots are not able to connect students directly with faculty members. Especially the described retrieval-based response generation technology with human assistance is suitable (7). Frequently asked questions can be answered by the chatbot technology autonomously; therefore, faculty members could have more time to connect with students and answer questions that the chatbot technology automatically hands over to them. Because of the retrieval-based approach, the chatbot technology is likely to give correct responses without grammatical errors. As correct responses can be used to train the chatbot, it can learn from responses the instructor makes (c.f. section 2.1 and 4.2). Therefore, the described technology is also suitable for the question answering on course content, where the instructor should be able to give answers, if the chatbot has no fitting response and the chatbot technology should be able to learn from responses the instructor makes (c.f. section 2.3).

Giving prompt feedback is the fourth principle of good practice (c.f. section 2.2). The previously described technology (7) is suitable, if detailed and individual feedback is needed. Retrieval-based response generation technologies are more flexible than rule-based ones and can use Named-Entity Recognition to identify missing aspects, to give a more precise individual feedback. In addition, human assistance can help to provide more consistent, detailed, and personalised responses (c.f. section 4.1). As the feedback may not be as prompt as an autonomously working approach, the decision which technology is more suitable depends on what kind of feedback the chatbot technology is used for. All autonomously working technologies can give prompt feedback if they are designed to give feedback at all. As the feedback should be correct, non-judgemental, supportive, positive, and specific, approaches that select their feedback from a pre-written database are more suitable (c.f. section 2.2).

To replace the manual customer service of a university, the chatbot should be able provide information about activities inside the campus, upcoming events, feedback of the university and to answer questions related to the university. Furthermore, it should be able to navigate the users to the registration page and show a map of the university (c.f. section 2.3). Hence, a generative response technology is not suitable as it is a closed domain task (c.f. section 2.1). Most information about activities, events, and the university is often already given on different websites of the university. Therefore, a retrieval-based response technology is most suitable, as it is able to query and analyse resources using APIs (c.f. section 2.1). For a rule-based approach, changing the events and activities would be a large amount of work, as they must be saved in the pre-written responses in the knowledge base. To replace the manual customer service, a technology, that is not human-assisted is appropriate. A long-term technology is not required to answer the questions, as they do not need a personalisation beyond the actual interaction. Therefore, a retrieval-based response generation, short-term interaction technology without human assistance is suitable (8).

To help students select specialised courses from several offers and give them a list of personalised recommendations is another use case of chatbot technologies (c.f. section 2.3). For this purpose, a hybrid long-term interaction technology with text-based out- and input is suitable (12). The technology can combine a recommender system with a rule-based question answering approach. It is suitable for this use case, as it is able to ask question for missing facets, that are important for the recommendation. Therefore, the chatbot technology can ask specific questions for the course selection. As the system is able to give recommendations based on the long-term preferences (c.f. section 4.2), already made decisions with regard to the course selection could be integrated to give a personalised recommendation list.

Another use case for chatbots is the admission service domain, to reduce the time students must wait for asked information. Additionally, should the chatbot be able to store new data if it is not able to answer a question (c.f. section 2.3). For this purpose, a hybrid approach would be suitable, with text-based input and output, without human assistance and short-term interaction (13). The same technology, with text-and voice-based input and output would also be suitable (14). A combination of a rule-based and retrieval-based response generation is suitable, as the rule-based approach can be used to give responses to frequently asked questions and the retrieval-based approach is used if the rule-based approach is not able to answer a question. Therefore, a fitting rule does not have to be designed for each question, and the already existing data can be used for the retrieval-based response generation. Additionally, the system can store unanswered questions and be trained with them (c.f. section 4.2). Since the use case does not require the chatbot to personalise feedback, provide information about previous dialogues, or pass unanswered responses to an admission agent, the autonomous short-term interaction technologies are suitable.

5.2 Summary and Interpretation of Suitable Chatbot Technologies

Of the 14 different technologies, that were introduced in section 4.2, nine were found to be suitable for the requirements of higher education. There are long- and short-term interaction technologies, technologies that use human assistance and autonomously working chatbots.

From the rule-based approaches, there are suitable non-human-assisted, short-term interaction technologies, which use different inputs and outputs (4,6,2). These technologies are mainly used to provide different kinds of information. Further there is a suitable rule-based long-term technology (3). Rule-based response generation technologies are especially suitable, when the information does not change fast or the dialogue is pre-defined.

Both retrieval-based response generation technologies are suitable for at least one use case in higher education. They are applicable, if the information is already existing and often changing, as they are more flexible and able to use resources from API request. Further they are useful in situations, where the response should be based on more complex information.

All three hybrid chatbot technologies are suitable for the usage in higher education. Due to the use of different response generation technologies, some of these are suitable for very specific purposes, such as giving suggestions for course selection, for which technology (12) is particularly suitable because of its ability to integrate previous decisions. Some hybrid technologies meanwhile are particularly suitable in other cases by combining the advantages of rule-based and retrieval-based response generation and are able to be trained with data from unanswerable queries. The rule-based approach can be used to cover standard queries. Overall, the ability to use voice-based input is not as critical as other factors of the chatbot technology. This is due to the fact, that they are wide-spread, and their use in a system has become quite easy (c.f. section 2.1).

The fact that all three found generative chatbot technologies are not suitable can be explained by the use cases described - they are all closed domain tasks. This list of suitable technologies is not exhaustive, as their suitability is checked with the requirements in higher education presented in section two. Especially the use cases, which present an excerpt from the literature are not exhaustive. This is not only due to the reason, that it is out of the scope of this work, to find all possible use cases but also, to the fact, that the implementation of chatbots is in an early stage in education. Hence there is still room for applications of these technologies in education and therefore, it is possible, that chatbot technologies, that have not been found suitable for a specific use case, could be used in a, until now unknown, use case.

6.0 Conclusion and Outlook

To find suitable chatbot technologies for the requirements in higher education, three research questions are formulated and subsequently answered. The proposed categorisation categorises chatbot technologies in terms of their *response generation method*, their *interaction duration* and their *human assistance*. Further an interaction mode category was found in the literature which is split into *user input* and *output* in this work, as not every chatbot that is able to work with voice-based user input is also able to give voice-based output.

A total of 14 different types of chatbot technologies were found. Most of them use rule-based response generation and are short-term interaction technologies. Only four technologies are capable of long-term interactions and only two technologies use human assistance. All found chatbots can work with text-based user input and five also with voice-based. All technologies give text-based output and two can also give voice-based output.

In analysing their suitability for higher education, nine of the 14 different types were found to be suitable. Most of the rule-based response generation technologies and every retrieval-based or hybrid technology are suitable for the use in higher education.

Especially, the technology with retrieval-based response generation and human assistance is useful, as it is able to answer complex questions. It can give more consistent, detailed and individual output, as the responses are only given, when the chatbot is able to give a fitting response. In all other cases, the question is handed to a human, which answers the question in a precise way and uses this answer to train the chatbot.

Such chatbot is suitable to give prompt feedback on complex questions, while a technology with a rule-based response generation and without human assistance is more suitable in a simple question-answer scenario. Therefore, the work has shown, that depending on the use case and the complexity of a task, different technologies are suitable.

No technology with a generative response generation was found to be suitable for the requirements of higher education as these technologies translate their input into an output. This makes them a good choice for informal open domain tasks, but not for closed domain tasks. As there are currently no requirements and use cases with informal open domain tasks, they are not suitable for the requirements in higher education.

Therefore, the technologies that have the greatest possible benefit in higher education are, depending on the task, technologies with rule-, retrieval-based or hybrid response generation. The task determines whether the highest benefit is given by a long-term or short-term interaction technology, whether human assistance is required and what types of input and output they should be able to work with.

The current work has some limitations. Only publications between 2015 and 2021 were examined, so older chatbot technologies are not included in this work. This problem could be addressed by conducting a backward search. This would also help with the fact that only three databases were searched. As the analysis of technologies was made on a functional level, there was no deeper technical analysis which excluded questions on data security as an example. This is an important goal for a next research iteration

Future research should also explore more use cases in higher education and analyse the, as the use of chatbots is still at an early stage and the benefits they can bring are great.

References

Aalipour, G., Kumar, P., Aditham, S., Nguyen, T., Sood, A., 2018. *Applications of Sequence to Sequence Models for Technical Support Automation.* 2018 IEEE International Conference on Big Data (Big Data), 10.12.2018 - 13.12.2018 2018 Seattle, WA, USA IEEE, 4861–4869.

Adamopoulou, E., Moussiades, L., 2020a. *An Overview of Chatbot Technology*. In: Maglogiannis, I., Iliadis, L., Pimenidis, E. (Eds.). Artificial Intelligence Applications and Innovations. Cham: Springer International Publishing, 373–383.

Adamopoulou, E., Moussiades, L., 2020b. *Chatbots: History, technology, and applications.* Machine Learning with Applications 2 (2020).

Agus Santoso, H., Anisa Sri Winarsih, N., Mulyanto, E., Wilujeng saraswati, G., Enggar Sukmana, S., Rustad, S., Syaifur Rohman, M., Nugraha, A., Firdausillah, F., 2018. *Dinus Intelligent Assistance (DINA) Chatbot for University Admission Services.* 2018 International Seminar on Application for Technology of Information and Communication, 21.09.2018 - 22.09.2018 2018 Semarang IEEE, 417–423.

Ahmad, N.A., Zainal, A., Baharum, Z., 2019. *UNISEL Bot: Designing Simple Chatbot System for University FAQs.* International Journal of Innovative Technology and Exploring Engineering, 9 (2), 4689–4693.

Alambo, A., Gaur, M., Lokala, U., Kursuncu, U., Thirunarayan, K., Gyrard, A., Sheth, A., Welton, R.S., Pathak, J., 2019. *Question Answering for Suicide Risk Assessment Using Reddit.* 2019 IEEE 13th International Conference on Semantic Computing (ICSC) 2019 Newport Beach, CA, USA IEEE, 468–473.

Ali, M.R., Razavi, S.Z., Langevin, R., Al Mamun, A., Kane, B., Rawassizadeh, R., Schubert, L.K., Hoque, E., 2020. *A Virtual Conversational Agent for Teens with Autism Spectrum Disorder: Experimental Results and Design Lessons.* Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, 20 10 2020 22 10 2020 2020 Virtual Event Scotland UK. New York, NY, USA: ACM, 1–8.

Battineni, G., Chintalapudi, N., Amenta, F., 2020. *AI Chatbot Design during an Epidemic Like the Novel Coronavirus*.

Baylor, A.L., Kim, S., 2009. *Designing nonverbal communication for pedagogical agents: When less is more.* Computers in Human Behavior, 25 (2), 450–457.

Benotti, L., Martinez, M.C., Schapachnik, F., 2018. *A Tool for Introducing Computer Science with Automatic Formative Assessment.* IEEE Transactions on Learning Technologies, 11 (2), 179–192.

Bharti, U., Bajaj, D., Batra, H., Lalit, S., Lalit, S., Gangwani, A., 2020. *Medbot: Conversational Artificial Intelligence Powered Chatbot for Delivering Tele-Health after COVID-19.* 2020 5th International Conference on Communication and Electronics Systems (ICCES), 10.06.2020 - 12.06.2020 2020 COIMBATORE, India IEEE, 870–875.

Bhawiyuga, A., Fauzi, A.M., Pramukantoro, E.S., Yahya, W., 2017. *Design of E-Commerce Chat Robot for Automatically Answering Customer Question.* Piscataway, NJ: IEEE.

Buyrukoğlu, S., Yılmaz, Y., 2021. *A Novel Semi-Automated Chatbot Model: Providing Consistent Response of Students Email in Higher Education based on Case-Based Reasoning and Latent Semantic Analysis*.

Carranza, K.A., Manalili, J., Bugtai, N.T., Baldovino, R.G., 2019. *Expression Tracking with OpenCV Deep Learning for a Development of Emotionally Aware Chatbots.* Piscataway, NJ: IEEE.

Chickering, A.W., Gamson, Z.F., 1987. *Seven Principles for Good Practice in Undergraduate Education*.

Chun Ho, C., Lee, H.L., Lo, W.K., Lui, K.F.A., 2018. Developing a Chatbot for College Student Programme Advisement. 2018 International Symposium on Educational Technology (ISET), 31.07.2018 - 02.08.2018 2018 Osaka, Japan IEEE, 52–56.

Cui, L., Huang, S., Wei, F., Tan, C., Duan, C., Zhou, M., 2017. *SuperAgent: A Customer Service Chatbot for E-commerce Websites.* In: Bansal, M., Ji, H. (Eds.). Proceedings of ACL 2017, System Demonstrations 2017 Vancouver, Canada. Stroudsburg, PA, USA: Association for Computational Linguistics, 97–102.

Dabbagh, N., Kitsantas, A., 2012. *Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning.* The Internet and Higher Education, 15 (1), 3–8.

Delahunty, F., Wood, I.D., Arcan, M., 2018. *First Insights on a Passive Major Depressive Disorder Prediction System with Incorporated Conversational Chatbot.* Irish Conference on Artificial Intelligence and Cognitive Science 2018.

Dharwadkar, R., Deshpande, N.A., 2018. *A Medical ChatBot*.

Dyba, T., Dingsoyr, T., Hanssen, G.K., 2007. *Applying Systematic Reviews to Diverse Study Types: An Experience Report.* First International Symposium on Empirical Software Engineering and Measurement (ESEM 2007), 20.09.2007 - 21.09.2007 2007 Madrid, Spain IEEE, 225–234.

Ehrmann, S., Chickering, A.W., 1996. *Implementing the Seven Principles: Technology as Lever*.

Fiore, D., Baldauf, M., Thiel, C., 2019. *"Forgot Your Password Again?" - Acceptance and User Experience of a Chatbot for In-Company IT Support.* In: Paternò, F., Jacucci, G., Rohs, M., Santoro, C. (Eds.). Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia, 26 11 2019 29 11 2019 2019 Pisa Italy. New York, NY, USA: ACM, 1–11.

Fook, C.Y., Sidhu, G.K., 2015. *Investigating Learning Challenges Faced by Students in Higher Education.* Procedia - Social and Behavioral Sciences, 186, 604–612.

Gabrielli, S., Rizzi, S., Bassi, G., Carbone, S., Maimone, R., Marchesoni, M., Forti, S., 2021. *Engagement and Effectiveness of a Healthy-Coping Intervention via Chatbot for University Students During the COVID-19 Pandemic: Mixed Methods Proof-of-Concept Study.* JMIR mHealth and uHealth, 9 (5).

Gaikwad, S.K., Gawali, B.W., Yannawar, P., 2010. *A Review on Speech Recognition Technique*.

Ghosha, S., Bhatia, S., Bhatia, A., 2018. *Quro: Facilitating User Symptom Check Using a Personalised Chatbot - Oriented Dialogue System.* Amsterdam, Washington DC: IOS Press.

Handoyo, E., Arfan, M., Soetrisno, Y.A.A., Somantri, M., Sofwan, A., Sinuraya, E.W., 2018. *Ticketing Chatbot Service using Serverless NLP Technology*.

Hayashi, Y., 2018. *The Power of a "Maverick" in Collaborative Problem Solving: An Experimental Investigation of Individual Perspective-Taking Within a Group.* Cognitive science, 42 Suppl 1, 69–104. 01.02.2018.

Hien, H.T., Cuong, P.-N., Nam, L.N.H., Le Nhung, H.T.K., Thang, L.D., 2018. *Intelligent Assistants in Higher-Education Environments.* In: Unknown (Ed.). Proceedings of the Ninth International Symposium on Information and Communication Technology - SoICT 2018, 06.12.2018 - 07.12.2018 2018 Danang City, Viet Nam. New York, New York, USA: ACM Press, 69–76.

Horizon Report, 2020. *2020 EDUCAUSE Horizon Report | Teaching and Learning Edition.* Available at: https://library.educause.edu/-/media/files/library/2020/3/2020_horizon_report_pdf.pdf?la=en&hash=08A92C17998E8113BCB15DCA7BA1F467F303BA80 [Last accessed at: 1-14-2022].

Huang, C.-Y.H., Ming-Chin Yang, Chin-Yu Huang, Yu-Jui Chen, Meng-Lin Wu, Kai-Wen Chen, 2018. *A Chatbot-supported Smart Wireless Interactive Healthcare System for Weight Control and Health Promotion.* Piscataway, NJ: IEEE.

Hussain, S., Ameri Sianaki, O., Ababneh, N., 2019. *A Survey on Conversational Agents/Chatbots Classification and Design Techniques.* In: Barolli, L., Takizawa, M., Xhafa, F., Enokido, T. (Eds.). Web, Artificial Intelligence and Network Applications. Cham: Springer International Publishing, 946–956.

John, L., Potti, N., Patel, J.M., 2017. *Ava: From Data to Insights Through Conversation*.

Jusoh, S., 2018. *Intelligent Conversational Agent for Online Sales*.

Kadariya, D., Venkataramanan, R., Yip, H.Y., Kalra, M., Thirunarayanan, K., Sheth, A., 2019. *kBot: Knowledge-enabled Personalized Chatbot for Asthma Self-Management.* Proceedings of International Conference on Smart Computing (SMARTCOMP), 2019, 138–143.

Khanna, A., Pandey, B., Vashishta, K., Kalia, K., Pradeepkumar, B., Das, T., 2015. *A Study of Today's A.I. through Chatbots and Rediscovery of Machine Intelligence.* International Journal of u- and e-Service, Science and Technology, 8 (7), 277–284.

Kitchenham, B., Charters, S., 2007. *Guidelines for performing systematic literature reviews in software engineering*.

Kock, W.D. de, 2016. *Speech versus text supported hints in learning to solve word problems.* Computers in Human Behavior, 57, 300–311.

Koetter, F., Blohm, M., Drawehn, J., Kochanowski, M., Goetzer, J., Graziotin, D., Wagner, S., 2019a. *Conversational Agents for Insurance Companies: From Theory to Practice*, 11978 (5), 338–362.

Koetter, F., Blohm, M., Kochanowski, M., Goetzer, J., Graziotin, D., Wagner, S., 2019b. *Motivations, Classification and Model Trial of Conversational Agents for Insurance Companies*, 19–30.

Kucherbaev, P., Bozzon, A., Houben, G.-J., 2018. *Human-Aided Bots.* IEEE Internet Computing, 22 (6), 36–43.

Lee, D., Oh, K.-J., Choi, H.-J., 2017. *The chatbot feels you - a counseling service using emotional response generation.* 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), 13.02.2017 - 16.02.2017 2017 Jeju Island, South Korea IEEE, 437–440.

Li, J., Zhou, M.X., Yang, H., Mark, G., 2017. *Confiding in and Listening to Virtual Agents: The Effect of Personality*. In: Papadopoulos, G.A., Kuflik, T., Chen, F., Duarte, C., Fu, W.-T. (Eds.). Proceedings of the 22nd International Conference on Intelligent User Interfaces, 13 03 2017 16 03 2017 2017 Limassol Cyprus. New York, NY, USA: ACM, 275–286.

Liu, B., Mei, C., 2020. *Lifelong Knowledge Learning in Rule-based Dialogue Systems* [Last accessed at: 1-14-2022]. Available at: https://arxiv.org/pdf/2011.09811.

Lokman, A.S., Ameedeen, M.A., 2019. *Modern Chatbot Systems: A Technical Review.* In: Arai, K., Bhatia, R., Kapoor, S. (Eds.). Proceedings of the Future Technologies Conference (FTC) 2018. Cham: Springer International Publishing, 1012–1023.

Ly, K.H., Ly, A.-M., Andersson, G., 2017. *A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods.* Internet interventions, 10, 39–46. 10.10.2017.

Maicher, K., Danforth, D., Price, A., Zimmerman, L., Wilcox, B., Liston, B., Cronau, H., Belknap, L., Ledford, C., Way, D., Post, D., Macerollo, A., Rizer, M., 2017. *Developing a Conversational Virtual Standardized Patient to Enable Students to Practice History-Taking Skills.* Simulation in healthcare: Journal of the Society for Simulation in Healthcare, 12 (2), 124–131.

Mansouri, A., Affendey, L.S., Mamat, A., 2008. *Named Entity Recognition Approaches.* IJCSNS International Journal of Computer Science and Network Security, 8 (2), 339–344.

Medeiros, L., Bosse, T., 2017. *An Empathic Agent that Alleviates Stress by Providing Support via Social Media.* Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems, 1634–1636.

Muangkammuen, P., Intiruk, N., Saikaew, K.R., 2018. *Automated Thai-FAQ Chatbot using RNN-LSTM.* Piscataway, NJ: IEEE.

Neumann, A.T., Arndt, T., Köbis, L., Meissner, R., Martin, A., Lange, P. de, Pengel, N., Klamma, R., Wollersheim, H.-W., 2021. *Chatbots as a Tool to Scale Mentoring Processes: Individually Supporting Self-Study in Higher Education.* Frontiers in Artificial Intelligence, 4.

Ni, L., Lu, C., Liu, N., Liu, J., 2017. *MANDY: Towards a Smart Primary Care Chatbot Application.* In: Chen, J., Theeramunkong, T., Supnithi, T., Tang, X. (Eds.). Knowledge and Systems Sciences. Singapore: Springer Singapore, 38–52.

Novielli, N., Rosis, F. de, Mazzotta, I., 2010. *User attitude towards an embodied conversational agent: Effects of the interaction mode.* Journal of Pragmatics, 42 (9), 2385–2397.

Oh, K.-J., Lee, D., Ko, B., Choi, H.-J., 2017. *A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation.* 2017 18th IEEE International Conference on Mobile Data Management (MDM), 29.05.2017 - 01.06.2017 2017 Daejeon, South Korea IEEE, 371–375.

Paikens, P., Znotiņš, A., Bārzdiņš, G., 2020. *Human-in-the-Loop Conversation Agent for Customer Service.* In: Métais, E., Meziane, F., Horacek, H., Cimiano, P. (Eds.). Natural Language Processing and Information Systems. Cham: Springer International Publishing, 277–284.

Pichponreay, L., Kim, J.-H., Choi, C.-H., Lee, K.-H., Cho, W.-S., 2016. *Smart answering Chatbot based on OCR and Overgenerating Transformations and Ranking.* 2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN), 05.07.2016 - 08.07.2016 2016 Vienna, Austria IEEE, 1002–1005.

Portela, M., Granell-Canut, C., 2017. *A new friend in our Smartphone? Observing Interactions with Chatbots in the search of emotional engagement.* In: González-Calleros, J.M. (Ed.). Proceedings of the XVIII International

Conference on Human Computer Interaction, 25 09 2017 27 09 2017 2017 Cancun Mexico. New York, NY, USA: ACM, 1–7.

Prabowo, Y.D., Warnars, Harco Leslie Hendric Spits, Budiharto, W., Kistijantoro, A.I., Heryadi, Y., Lukas, 2018. *Lstm And Simple Rnn Comparison In The Problem Of Sequence To Sequence On Conversation Data Using Bahasa Indonesia.* Piscataway, NJ, USA: IEEE.

Putri, F.P., Meidia, H., Gunawan, D., 2019. *Designing Intelligent Personalized Chatbot for Hotel Services.* Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence, 20 12 2019 22 12 2019 2019 Sanya China. New York, NY, USA: ACM, 468–472.

Ramesh, K., Ravishankaran, S., Joshi, A., Chandrasekaran, K., 2017. *A Survey of Design Techniques for Conversational Agents.* In: Kaushik, S., Gupta, D., Kharb, L., Chahal, D. (Eds.). Information, Communication and Computing Technology. Singapore: Springer Singapore, 336–350.

Ravi, R., 2018. *Intelligent Chatbot for Easy Web-Analytics Insights: 19-22 Sept. 2018.* Piscataway, NJ: IEEE.

Reshmi, S., Balakrishnan, K., 2018. *Empowering Chatbots with Business Intelligence by Big Data Integration.* International Journal of Advanced Research in Computer Science, 9 (1), 627–631.

Romero, O.J., Zhao, R., Cassell, J., 2017. *Cognitive-Inspired Conversational-Strategy Reasoner for Socially-Aware Agents.* In: Bacchus, F., Sierra, C. (Eds.). Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, 19.08.2017 - 26.08.2017 2017 Melbourne, Australia. California: International Joint Conferences on Artificial Intelligence Organization, 3807–3813.

Ruan, S., Jiang, L., Xu, J., Tham, B.J.-K., Qiu, Z., Zhu, Y., Murnane, E.L., Brunskill, E., Landay, J.A., 2019. *QuizBot: A Dialogue-based Adaptive Learning System for Factual Knowledge.* In: Brewster, S., Fitzpatrick, G., Cox, A., Kostakos, V. (Eds.). Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 04 05 2019 09 05 2019 2019 Glasgow Scotland Uk. New York, NY, USA: ACM, 1–13.

Satu, M.S., Parvez, M.H., Shamim-Al-Mamun, 2015. *An E-business Chatbot using AIML and LSA.* 2015 International Conference on Computer and Information Engineering (ICCIE), 26.11.2015 - 27.11.2015 2015 Rajshahi IEEE, 87–90.

Serban, I.V., Sankar, C., Germain, M., Zhang, S., Lin, Z., Subramanian, S., Kim, T., Pieper, M., Chandar, S., Ke, N.R., Rajeshwar, S., Brebisson, A.d., Sotelo, J.M.R., Suhubdy, D., Michalski, V., Nguyen, A., Pineau, J., Bengio, Y., 2017. *A Deep Reinforcement Learning Chatbot* [Last accessed at: 1-14-2022]. Available at: http://arxiv.org/pdf/1709.02349v2.

Shakhovska, N., Basystiuk, O., Shakhovska, K., 2019. *Development of the Speech-to-Text Chatbot Interface Based on Google API.* Available at: http://ceur-ws.org/Vol-2386/paper16.pdf [Last accessed at: 1-14-2022].

Sheikh, S.A.S., Tiwari, V., Singhal, S., 2019. *Generative model chatbot for Human Resource using Deep Learning.* Piscataway, NJ: IEEE.

Shute, V.J., 2008. *Focus on Formative Feedback.* Review of Educational Research, 78 (1), 153–189.

Singh, R., Paste, M., Shinde, N., Patel, H., Mishra, N., 2018. *Chatbot using TensorFlow for small Businesses.* 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 20.04.2018 - 21.04.2018 2018 Coimbatore IEEE, 1614–1619.

Song, D., Oh, E.Y., Rice, M., 2017. *Interacting with a Conversational Agent System for Educational Purposes in Online Courses.* Piscataway, NJ: IEEE.

Suleman, R.M., Mizoguchi, R., Ikeda, M., 2016. *A New Perspective of Negotiation-Based Dialog to Enhance Metacognitive Skills in the Context of Open Learner Models.* International Journal of Artificial Intelligence in Education, 26 (4), 1069–1115.

Sun, Y., Zhang, Y., 2018. *Conversational Recommender System.* In: Collins-Thompson, K., Mei, Q., Davison, B., Liu, Y., Yilmaz, E. (Eds.). The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, 08 07 2018 12 07 2018 2018 Ann Arbor MI USA. New York, NY, USA: ACM, 235–244.

Tegos, S., Demetriadis, S., Karakostas, A., 2015. *Promoting academically productive talk with conversational agent interventions in collaborative learning settings.* Computers & Education, 87, 309–325.

Thorat, S.A., Jadhav, V., 2020. *A Review on Implementation Issues of Rule-based Chatbot Systems.* SSRN Electronic Journal.

Vijayan, A., S Janmasree, C Keerthana, L Baby Syla, 2018. *A Framework for Intelligent Learning Assistant Platform based on Cognitive Computing for Children with Autism Spectrum Disorder.* Piscataway, NJ: IEEE.

Villegas-Ch, W., Arias-Navarrete, A., Palacios-Pacheco, X., 2020. *Proposal of an Architecture for the Integration of a Chatbot with Artificial Intelligence in a Smart Campus for the Improvement of Learning.* Sustainability, 12 (4), 1500.

Vinod Kumar Shukla, Amit Verma, 2019. *Enhancing LMS Experience through AIML Base and Retrieval Base Chatbot using R Language.* Piscataway, NJ: IEEE.

vom Brocke, J., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R., Cleven, A., 2009. *Reconstructing the Giant: on the Importance of Rigour in Documenting the Literature Search Process*.

Wallace, R.U., 2009. *The Anatomy of A.L.I.C.E.* BMJ, 1 (2787), 1218.

Wang, H., Zhang, Q., Ip, M., Fai Lau, J.T., 2018. *Social Media–based Conversational Agents for Health Management and Interventions.* Computer, 51 (8), 26–33.

Wang, J., Hwang, G.-H., Chang, C.-Y., 2021. *Directions of the 100 most cited chatbot-related human behavior research: A review of academic publications.* Computers and Education: Artificial Intelligence, 2, 100023.

Weizenbaum, J., 1983. *ELIZA — a computer program for the study of natural language communication between man and machine.* Communications of the ACM, 26 (1), 23–28.

Winkler, R., Hobert, S., Salovaara, A., Söllner, M., Leimeister, J.M., 2020. *Sara, the Lecturer: Improving Learning in Online Education with a Scaffolding-Based Conversational Agent.* In: Bernhaupt, R., Mueller, F.'., Verweij, D., Andres, J., McGrenere, J., Cockburn, A., Avellino, I., Goguey, A., Bjørn, P., Zhao, S., Samson, B.P., Kocielnik, R. (Eds.). Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 25 04 2020 30 04 2020 2020 Honolulu HI USA. New York, NY, USA: ACM, 1–14.

Winkler, R., Söllner, M., 2018. *Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis* 2018 Chicago, USA.

Wu, Y., Wu, W., Xing, C., Zhou, M., Li, Z., 2017. *Sequential Matching Network: A New Architecture for Multi-turn Response Selection in Retrieval-based Chatbots.* Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 496–505.

Xu, A., Liu, Z., Guo, Y., Sinha, V., Akkiraju, R., 2017. *A New Chatbot for Customer Service on Social Media.* In: Mark, G., Fussell, S., Lampe, C., schraefel, m., Hourcade, J.P., Appert, C., Wigdor, D. (Eds.). Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 06 05 2017 11 05 2017 2017 Denver Colorado USA. New York, NY, USA: ACM, 3506–3510.

Xue, Z., Ko, T.-Y., Yuchen, N., Wu, M.-K.D., Hsieh, C.-C., 2018. *Isa: Intuit Smart Agent, A Neural-Based Agent-Assist Chatbot.* 2018 IEEE International Conference on Data Mining Workshops (ICDMW), 17.11.2018 - 20.11.2018 2018 Singapore, Singapore IEEE, 1423–1428.

Appendix 1:

Title	Year	Citation
Interacting with a Conversational Agent System for Educational Purposes in Online Courses (Song et al. 2017)	2017	36
Promoting academically productive talk with conversational agent interventions in collaborative learning settings (Tegos et al. 2015)	2015	22
QuizBot: A Dialogue-based Adaptive Learning System for Factual Knowledge (Ruan et al. 2019)	2019	21
Developing a Conversational Virtual Standardized Patient to Enable Students to Practice History-Taking Skills (Maicher et al. 2017)	2017	20
A Tool for Introducing Computer Science with Automatic Formative Assessment (Benotti et al. 2018)	2018	18
Smart Answering Chatbot based on OCR and Overgenerating Transformations and Ranking (Pichponreay et al. 2016)	2016	18
Sara, the Lecturer: Improving Learning in Online Education with a Scaffolding-Based Conversational Agent (Winkler et al. 2020)	2020	14
Proposal of an Architecture for the Integration of a Chatbot with Artificial Intelligence in a Smart Campus for the Improvement of Learning (Villegas-Ch et al. 2020)	2020	14
Enhancing LMS Experience through AIML Base and Retrieval Base Chatbot using R Language (Vinod Kumar Shukla and Amit Verma 2019)	2019	14
A New Perspective of Negotiation-Based Dialog to Enhance Metacognitive Skills in the Context of Open Learner Models (Suleman et al. 2016)	2016	14
A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation (Oh et al. 2017)	2017	62
MANDY: Towards a smart primary care chatbot application (Ni et al. 2017)	2017	30
The chatbot feels you - A counseling service using emotional response generation (Lee et al. 2017)	2017	25
Quro: Facilitating User Symptom Check Using a Personalised Chatbot-Oriented Dialogue System (Ghosha et al. 2018)	2018	23
A Chatbot-supported Smart Wireless Interactive Healthcare System for Weight Control and Health Promotion (Huang et al. 2018)	2018	18

AI Chatbot Design during an Epidemic like the Novel Coronavirus (Battineni et al. 2020)	2020	17
kBot: Knowledge-enabled Personalized Chatbot for Asthma Self-Management (Kadariya et al. 2019)	2019	16
Social Media–based Conversational Agents for Health Management and Interventions (Wang et al. 2018)	2018	11
Medbot: Conversational Artificial Intelligence Powered Chatbot for Delivering Tele-Health after COVID-19 (Bharti et al. 2020)	2020	10
Question Answering for Suicide Risk Assessment Using Reddit (Alambo et al. 2019)	2019	9
A new chatbot for customer service on social media (Xu et al. 2017)	2017	181
SuperAgent: A Customer Service Chatbot for E-commerce Websites (Cui et al. 2017)	2017	78
An E-business Chatbot using AIML and LSA(Satu et al. 2015)	2015	44
Automated Thai-FAQ Chatbot using RNN-LSTM (Muangkammuen et al. 2018)	2018	8
Ticketing Chatbot Service using Serverless NLP Technology (Handoyo et al. 2018)	2018	7
"Forgot your password again?": acceptance and user experience of a chatbot for in-company IT support (Fiore et al. 2019)	2019	6
Design of E-commerce chat robot for automatically answering customer question (Bhawiyuga et al. 2017)	2017	5
Designing Intelligent Personalized Chatbot for Hotel Services (Putri et al. 2019)	2019	4
Isa: Intuit smart agent, a neural-based agent-assist chatbot (Xue et al. 2018)	2018	4
Motivations, Classification and Model Trial of Conversational Agents for Insurance Companies (Koetter et al. 2019b)	2019	4
Conversational Recommender System (Sun and Zhang 2018)	2018	80
Ava: From Data to Insights Through Conversation (John et al. 2017)	2017	18
Chatbot using TensorFlow for small Businesses(Singh et al. 2018)	2018	12
Intelligent Chatbot for Easy Web-Analytics Insights (Ravi 2018)	2018	8

Applications of Sequence to Sequence Models for Technical Support Automation (Aalipour et al. 2018)	2018	5
Intelligent Conversational Agent for Online Sales (Jusoh 2018)	2018	3
Lstm And Simple Rnn Comparison In The Problem Of Sequence To Sequence On Conversation Data Using Bahasa Indonesia (Prabowo et al. 2018)	2018	3
Conversational Agents for Insurance Companies: From Theory to Practice (Koetter et al. 2019a)	2019	3
Human-in-the-loop conversation agent for customer service (Paikens et al. 2020)	2020	2
Generative model chatbot for Human Resource using Deep Learning (Sheikh et al. 2019)	2019	2
A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods (Ly et al. 2017)	2017	67
Confiding in and Listening to Virtual Agents: The Effect of Personality (Li et al. 2017)	2017	22
Cognitive-Inspired Conversational-Strategy Reasoner for Socially-Aware Agents (Romero et al. 2017)	2017	12
A new friend in our smartphone?: observing interactions with chatbots in the search of emotional engagement (Portela and Granell-Canut 2017)	2017	12
The Power of a "Maverick" in Collaborative Problem Solving: An Experimental Investigation of Individual Perspective-Taking Within a Group (Hayashi 2018)	2018	6
First insights on a passive major depressive disorder prediction system with incorporated conversational chatbot (Delahunty et al. 2018)	2018	5
A Framework for Intelligent Learning Assistant Platform based on Cognitive Computing for Children with Autism Spectrum Disorder (Vijayan et al. 2018)	2018	2
A Virtual Conversational Agent for Teens with Autism Spectrum Disorder: Experimental Results and Design Lessons (Ali et al. 2020)	2020	2
An Empathic Agent that Alleviates Stress by Providing Support via Social Media (Medeiros and Bosse 2017)	2017	1
Expression Tracking with OpenCV Deep Learning for a Development of Emotionally Aware Chatbots (Carranza et al. 2019)	2019	1