EDUBOTS

Chatbots in Higher Education White Paper

Prepared by CYENS Center of Excellence
Date: December 2021

Co-funded by the Erasmus+ Programme of the European Union

THIS PROJECT HAS BEEN FUNDED WITH SUPPORT FROM THE EUROPEAN COMMISSION UNDER THE ERASMUS+ PROGRAMME. THIS PUBLICATION [COMMUNICATION] REFLECTS THE VIEWS ONLY OF THE AUTHOR, AND THE COMMISSION CANNOT BE HELD RESPONSIBLE FOR ANY USE WHICH MAY BE MADE OF THE INFORMATION CONTAINED THEREIN

Table of contents

Table of figures

Consortium

The table presented the consortium of the EDUBOTS project.

DIFFER	P1 – EDTECH FOUNDRY AS
HUBERT	P2 – ANNA & HUBERT LABS AB
UNIVERSITY OF LEEDS	P3 – UNIVERSITY OF LEEDS
foi	P4 – UNIVERSITY OF ZAGREB, FACULTY OF ORGANIZATION AND INFORMATICS
CYENS	P5 – CYENS CENTRE OF EXCELLENCE
UNIVERSIDAD DE GRANADA	P6 – UNIVERSIDAD DE GRANADA

Expert Members

Smart interviews with members from the EDUBOTS expert panel were conducted in an effort to collect important pieces of information with regards to chatbots (in education). The names of the expert names are listed below (see table 1), as well as the interview protocol, are given below (see Annex I).

Table 1. Members of the EDUBOTS expert panel

Name and Surname	**Affiliation**
Asbjørn Følstad	Sintef Digital
Timothy J Wilson	Bellwether College Consortium
Loizos Michael	CYENS Research Centre of Excellence
Rose Luckin	Knowledge Lab, UCL Institute of Education, University College London

Overview

The EDUBOTS project has the ambition of improving best practices of pedagogical chatbot usage in Higher Education. This white paper reflects insights from the literature, desk and field research and inquiry of commercial chatbot solutions that are used in the Higher Education sector. Chapter 1 provides an overview to chatbots, their history and technology used, as documented in the relevant literature and according to the insights of experts in the domain of Artificial Intelligence (AI) and educational chatbots, with whom smart interviews have been conducted, recorded upon the interviewees' consent, transcribed verbatim and analyzed with the adoption of a thematic analysis approach. Among our contributing experts, two come from academia (one from each organization: Bellwether College Consortium, University College London - Knowledge Lab) and two come from the industry (one from each organization: Sintef Digital, CYENS Centre of Excellence). Chapter 2 focuses on the application of chatbots in Higher Education, problems solved, and current practices of using chatbots in Higher Education, followed by a summary of pros, cons, and pricing. Key findings are presented as derived from an iterative needs analysis that was conducted in the EDUBOTS project, with Higher Education users (educators and students) with the aim to identify users' needs and expectations on the potential use of chatbots in Higher Education. Also, exemplary chatbot use cases, as applied in Higher Education settings, as part of the EDUBOTS piloting are presented. Finally, chapter 3 documents what research suggests in relation to chatbots added value in Higher Education students' learning.

Chapter 1: Chatbots and their history

1.1. Introduction

The way people interact with businesses, and technology, is changing rapidly. In the era of instant gratification and digitization, people expect immediate answers to their inquiries, including Higher Education users (e.g., students). To meet the demands of the modern consumer, applications are increasingly integrating with a popular conversational interface called the 'chatbot'.

1.2. What are the chatbots?

Dialogue systems and conversational agents, including chatbots, are becoming ubiquitous in modern society. Chatbots can also be identified in the literature as "chatbot virtual assistants", "conversational agents", "chat bots", "pedagogical agents", "intelligent tutor systems", "dialogue systems", "smart personal assistants" and "smart assistants". They comprise software tools that simulate textual and/or auditory conversations (see Figure 1) and with which users interact on a certain topic or in a specific domain through digital services in a natural, conversational way using text and voice input (Serban, et al., 2017; Smutny & Schreiberova, 2020; Winkler, & Söllner, 2018; ZEMČÍK, 2019). Questions in natural human language are obtained in the chatbot environment, then associated with a knowledge base, and then answers are provided to the user.

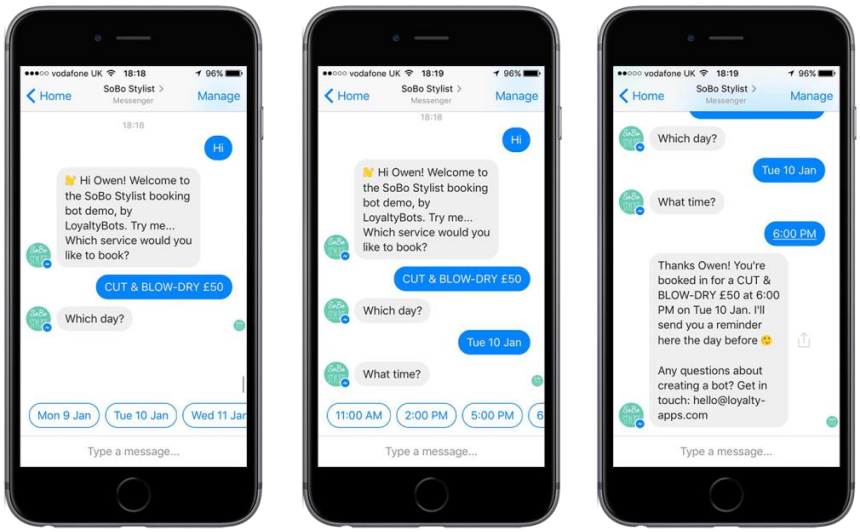


Fig. 1. Chatbots. Image via Loyalty Apps (https://www.loyalty-apps.com/chatbots/)

Chatbots however can accommodate multiple modalities, such as voice chats, text chats, images, and videos. In the animation given in figure 2 you can see an example of a text-based conversation agent. The interaction with a chatbot is much the same way as we do with a friend or a colleague and it responds back in a human-like mode, that demonstrates the personality that the chatbot developer has designed.

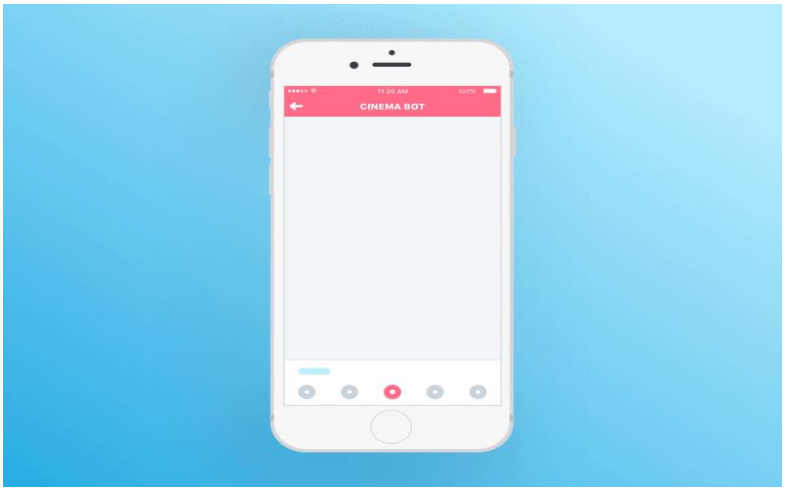


Fig. 2. A chat-based conversational agent. Image source: https://bit.ly/2OKTArW

Chatbot virtual assistants (also called AI or digital assistants, which are application programs that understand natural language voice commands and complete tasks for the user) are becoming a trend in many fields such as medicine, product, and service industry, and lately in education. Powered by AI, chatbots have the potential to identify user intent and provide relevant answers to user requests. They can be programmed to hold simple conversations and may even have an identity (e.g., the teacher). Commonly, chatbots appear in customer service issues, as virtual assistants, and as searchable frequently asked questions (FAQ), as personal assistants on mobile devices, technical support help over telephone lines, in health interventions, as well as online bots in business webpages for sailing products, to offering legal advice (Serban, et al., 2017). Therefore, chatbots have a growing presence in modern society. An international list of chatbots (virtual assistants, chat bot, conversational agents, virtual agents) can be found in the following site: https://www.chatbots.org/.

1.3. The history of chatbots

The possibility of training a computer machine to launch conversations with users, dates to the 1950s, with Alan Turing (Bernardini, Sônego, & Pozzebon, 2018), who proposed a test (the so-called, Turing test) consisting of a program that developed a text message conversation with a prober for five minutes (Turing, 2009). In this test, the user should predict if s/he was talking to a human being or a computer program. The Turing test can be considered as a key milestone in the pre-history of chatbots (Loizos Michael, expert member). In fact, chatbots' conceptualization emerged from the need of human to interact with computes in a natural human language. "*There has been ever since we started with the computers that it would be a good thing to be able to talk to the computers in natural language and that represents the milestone as a starting point of chatbots*" (Asbjørn Følstad, expert member).

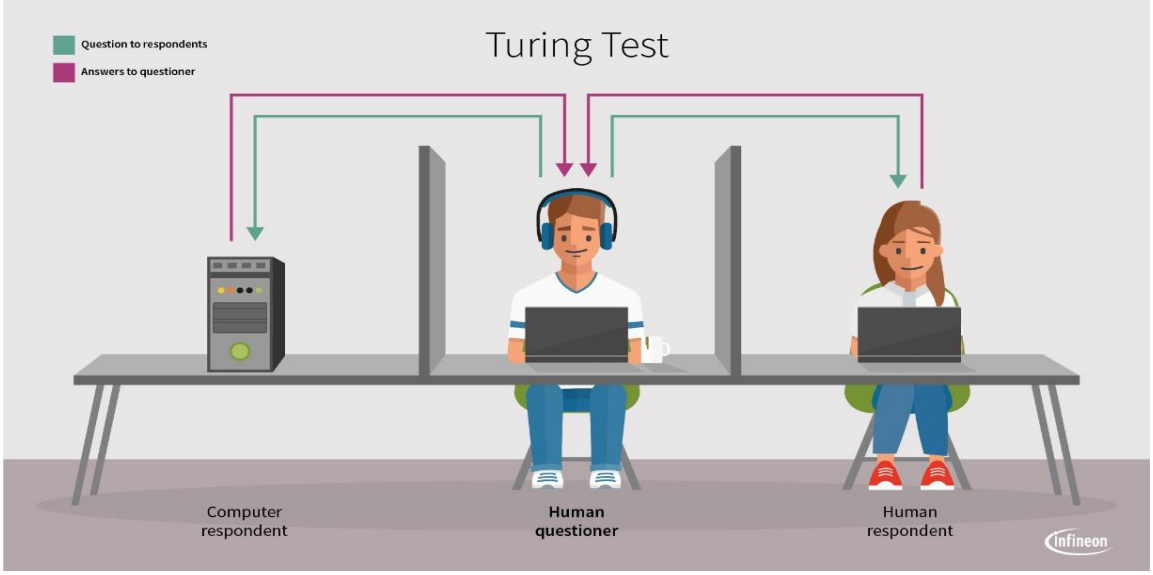


Fig. 3. The Turing Test. Image source: https://bit.ly/3sdcsNT

Following this, the first chatbot in the history of Computer Science was created in 1964-1966 by Joseph Weizenbaum at Massachusetts Institute of Technology (MIT), the so-called ELIZA (Kane, 2016; Khan, & Das, 2017; ZEMČÍK, 2019). ELIZA used simple pattern matching and a template-based response mechanism to imitate the conversational style of a nondirectional psychotherapist, in the early scenario called DOCTOR (ZEMČÍK, 2019).

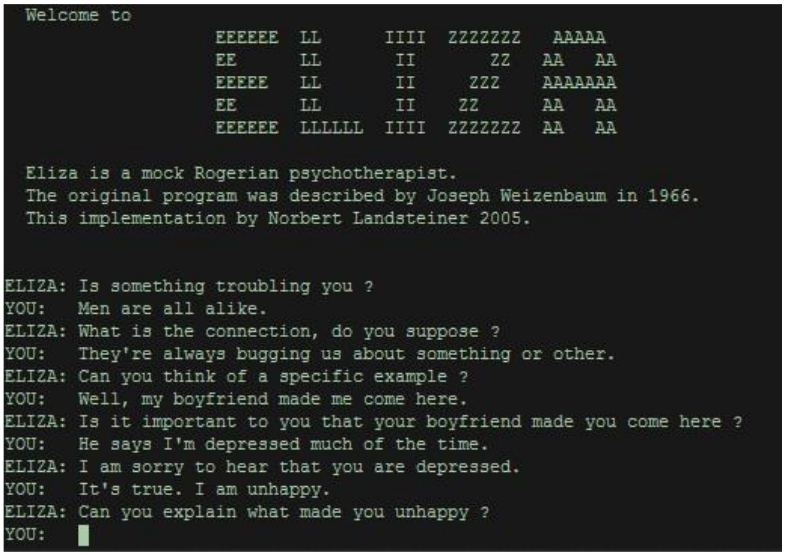


Fig. 4. The chatbot ELIZA (https://yakbots.com/chatbot-history-the-eliza-chatbot/).

ELIZA was made to recognize human interaction by way of simple pattern recognition. “Eliza was the first chatbot developed around 60 years ago and there have been many attempts to create more advanced versions of chatbot” (Loizos Michael, expert member) ELIZA (see figure 4) could analyze input sentences and recognize keywords or phrases from the input and create its response using those keywords from pre-programmed responses, based on reassembly rules associated with a

breakdown of the input. This chatbot functioned in a way that created an illusion of understanding and having an interaction with the user. Weizenbaum was shocked to observe that people thought that ELIZA was a real doctor and they instilled so much confidence in the chatbot that they even discussed personal problems with it. As he notes in his book Computer power and human reason (Weizenbaum, 1976), he was both shocked and felt rewarded at the way humans behaved. Users felt emotional when chatting with ELIZA which is incredible. ELIZA was also the first chatbot to win the Turing Test, which was developed by Alan Turing in 1950.The fact that the users seemed to be misled into thinking that ELIZA was a person rather than a chatbot, inspired a whole community of interest in building chatbots that might one day pass the Turing Test (Turing, 2009).

Another well-known chatbot is PARRY, introduced by Kenneth Mark Colby, a psychiatrist and computer scientist, at Stanford's Psychiatry Department in 1972 (ZEMČÍK, 2019). PARRY followed an opposite strategy from Eliza, behaving like a paranoid schizophrenic patient, aiming to provoke controversies and thus trigger the user in elaborating to one's answers. Racter (short for raconteur - a storyteller) was another interesting chatbot, developed by William Chamberlain and Thomas Etter under the Inrac Corporation in 1983 (ZEMČÍK, 2019). The next milestone in the development of chatbots came in 1991, with a chatbot which used a technologically ground-breaking novelty - the Sound Blaster sound card created by Creative Labs. The program was so-called Dr. Sbaitso (acronym from 'Sound Blaster Artificial Intelligent Text to Speech Operator'). (For more information: https://classicreload.com/dr-sbaitso.html). It was capable of synthesizing speech in a simplified manner.

Fig. 5. Dr. Sbaitso. Image source: https://classicreload.com/dr-sbaitso.html

The term "Chatterbot" was coined a few years later in 1994 by Michael Mauldin, the creator of the Verbot (Verbal Robot) Julia (Khan & Das, 2017; Molnár, & Szüts, 2018). The chatbot A.L.I.C.E (or Alicebot) followed three decades later and was developed in 1995 by Richard Wallace (Wallace, 2003 as cited in Kane, 2016; Khan, & Das, 2017) using the Artificial Intelligence Markup Language (AIML). A.L.I.C.E was able to use natural language processing, which allowed a more sophisticated conversation to take place among the chatbot and the user. AIML allowed pattern matching, but

patterns were kept short, and responses could be combined from multiple categories. In this chatbot, category patterns were matched to find the most appropriate response to user input (Krane, 2016).

Fig. 6. the Verbot (Verbal Robot) Julia. Image source: https://www.web3.lu/verbots/ l

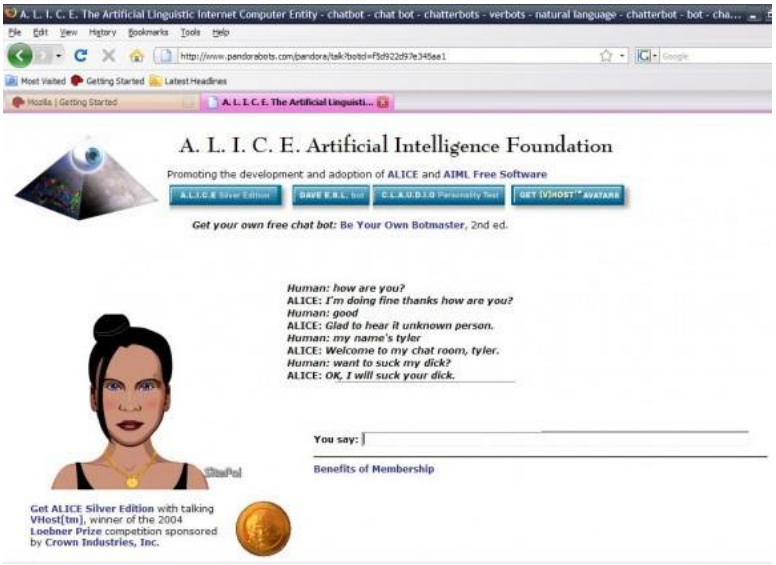


Fig. 7. The chatbot A.L.I.C.E (or Alicebot). Image source: https://bit.ly/3wNnJIi

"So, in the late nineties, we saw the first assistants, customer service chatbots that appeared as part of applications and then appeared in webpages or websites, sort of commercial website; a kind of virtual assistants. I think that is an important milestone because it reflects the first point when one tries to make things into something more than just some kind of chat application that actually serves some kind of purpose" (Asbjørn Følstad, expert member).

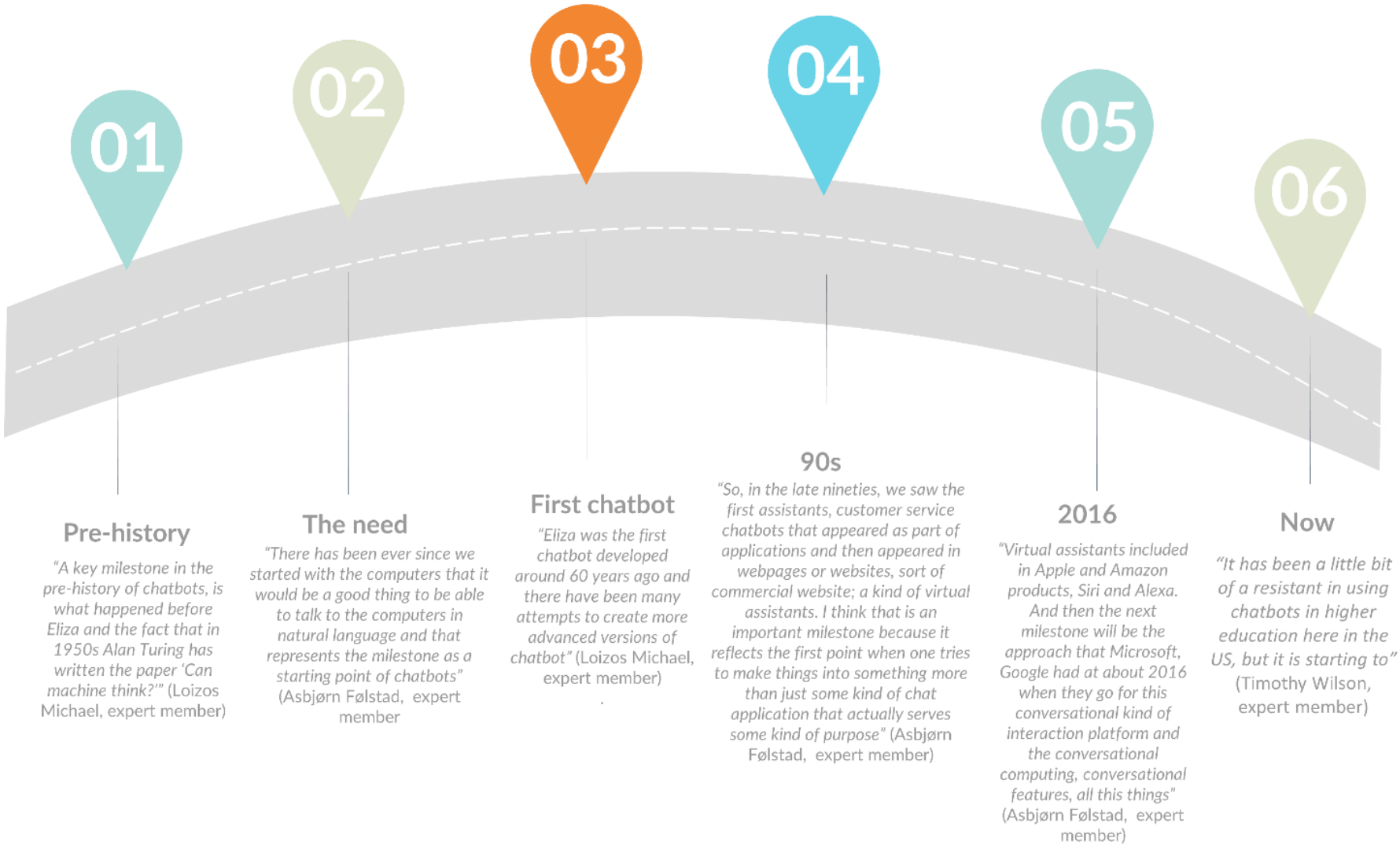


Fig. 8. Input from the EDUBOTS expert panel

The use of chatbots increased dramatically with the massive expansion of the Internet and especially social networking sites (ZEMČÍK, 2019). They have been advanced in the last decade due to the development in natural language processing and in machine learning algorithms, such as deep learning and neural networks which perform AI tasks like Image Recognition, Natural Language Generation, Speech Recognition and Text to Speech Synthesis (Khan & Das, 2017). Since the interaction with technology, using either natural language text or speech, is becoming increasingly feasible nowadays, chatbots can capture a wide range of use cases and can steer the user in the desired direction.

In 2006 IBM designed Watson, a computer system designed to answer questions. Actually, Watson won the popular quiz show Jeopardy. Watson also used natural language processing and is currently available for anyone to use in their apps. In 2008, 'Cleverbot' was launched, and, unlike other chatbots, its responses were not pre-programmed, but instead, the chatbot could learn from the human input (Gehl, 2014).

Fig. 9. Siri, SIRI, a voice-based conversational agent

Yet, chatbots also remained in the domain of tech labs and large organizations, as part of research projects and were not consumer focused. Probably the first ever real consumer-focused product was Apple Siri, a voice -based conversational agent that employs artificial intelligence to provide intelligent responses. Introduced in 2010, Siri allows users to literally talk to their smartphone. Apple's Siri, appeared for the first time on the iPhone in 2011, again in several languages, and became available in 2016 in a desktop version via macOS Sierra. Siri allows the users to engage in random conversations, whilst providing information regarding the weather, stocks etc. (Khan & Das, 2017).

Examples other voice-driven digital assistants that followed include Microsoft's Cortana, Amazon's Alexa and Google's new Assistant (Dale, 2016). Those personal digital assistants became at the forefront of technology of voice recognition and AI (Smutny, & Schreiberova, 2020). Microsoft's Cortana, released in 2014 for Windows Phone, became available in multiple languages, on the Windows 10 desktop operating system in early 2015; whereas in the middle of 2016, Cortana was due to appear in the Xbox One interface.

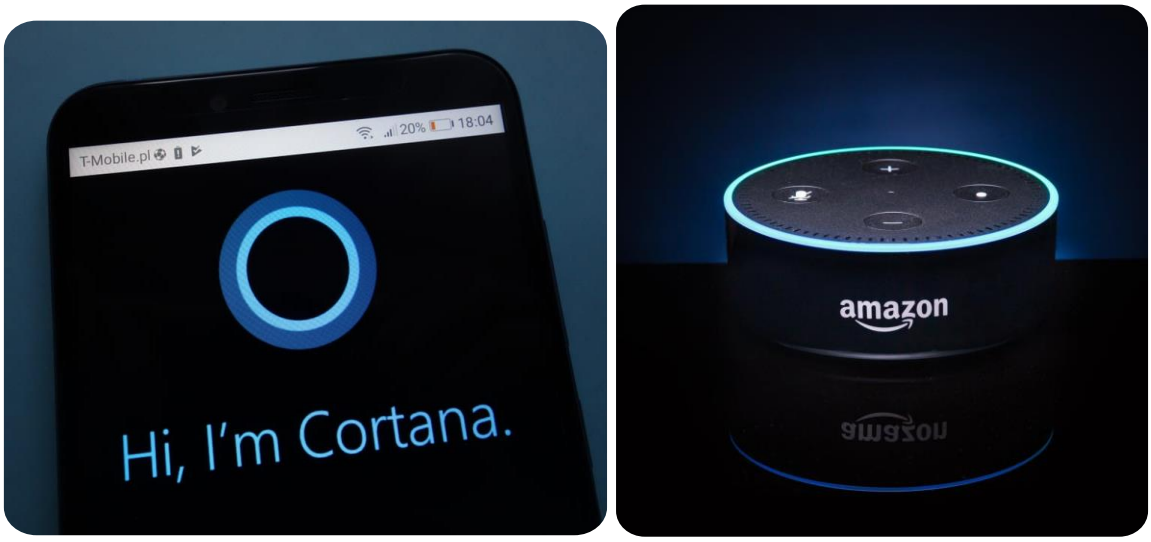


Fig. 10. Microsoft's Cortana (left) and Amazon's Alexa (right)

Amazon's Alexa, embodied in the Amazon Echo smart speaker, became widely available in the USA in 2015. Finally, Google Assistant, announced in May 2016, was an extension of Google Now that could keep track of a conversation. These applications encompass virtual personal assistants, which serve a user's everyday needs and activities, such as, scheduling meetings, checking the user's calendar, making appointments etc.

The above-mentioned devices are voice-based personal assistants. Since the core technology that drives all these products is AI, and more specifically, Natural Language Processing (NLP), there has been an ascent in the chat-based conversational agents as well. A conversational interface also scores high in BOT usability and simplicity. For instance, the news app Quartz uses a conversational interface, where you do not really type anything, but you do use buttons which are pre-set for you. This way, you control what you want to read about. Such pre-set buttons are standard in many chatbots.

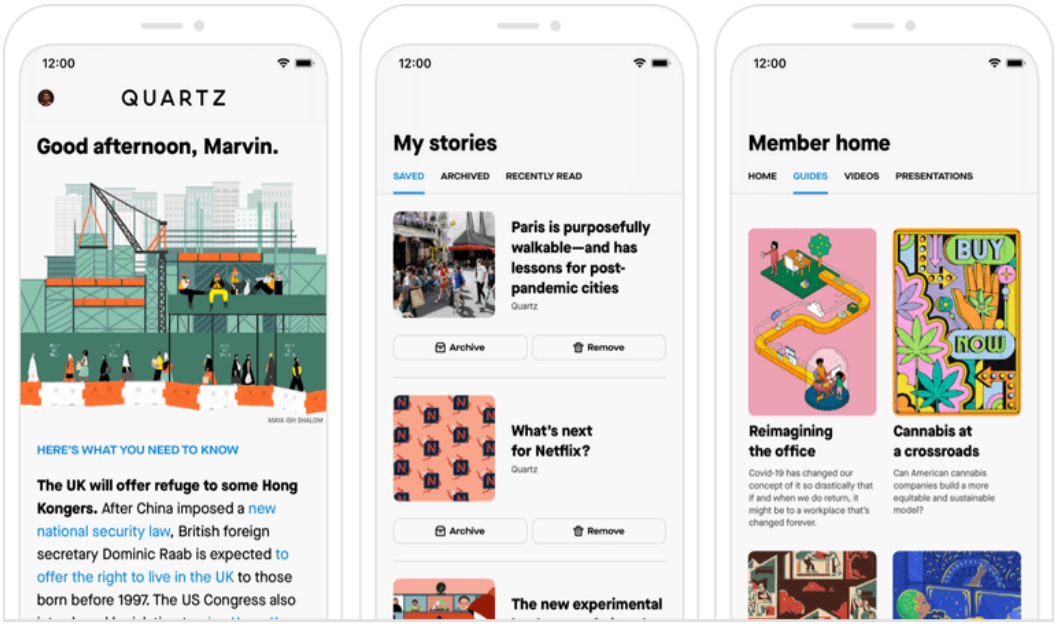


Fig. 11. Text-based conversational agents, here the news app Quartz (https://qz.com/app/)

The real push towards chatbots came in early 2016. The use of chatbots increased dramatically with the massive expansion of the Internet and especially social networking sites. They have been advanced in the last decade due to the development in natural language processing and in machine learning algorithms. Since the interaction with technology, using either natural language text or speech, is becoming increasingly feasible nowadays, chatbots can capture a wide range of use cases and can steer the user in the desired direction. In 2016, Facebook announced its messenger platform, which allows developers to hock in and create chatbots. While other messenger platforms had other bot channels as well, Facebook's announcement has been seen as a game changer, owing over 100 million people who have been using the messenger platform. That is a huge audience to target, and it is not surprising to see that Facebook's move has caused other companies to invest heavily in the same direction.

Along these lines, Google recently launched, 'Allo', a smart AI powered messaging app that learns from your messages and provides contextual options. Besides that, it offers an AI powered chatbot built in, that can create reminders, information about whether and lots more.

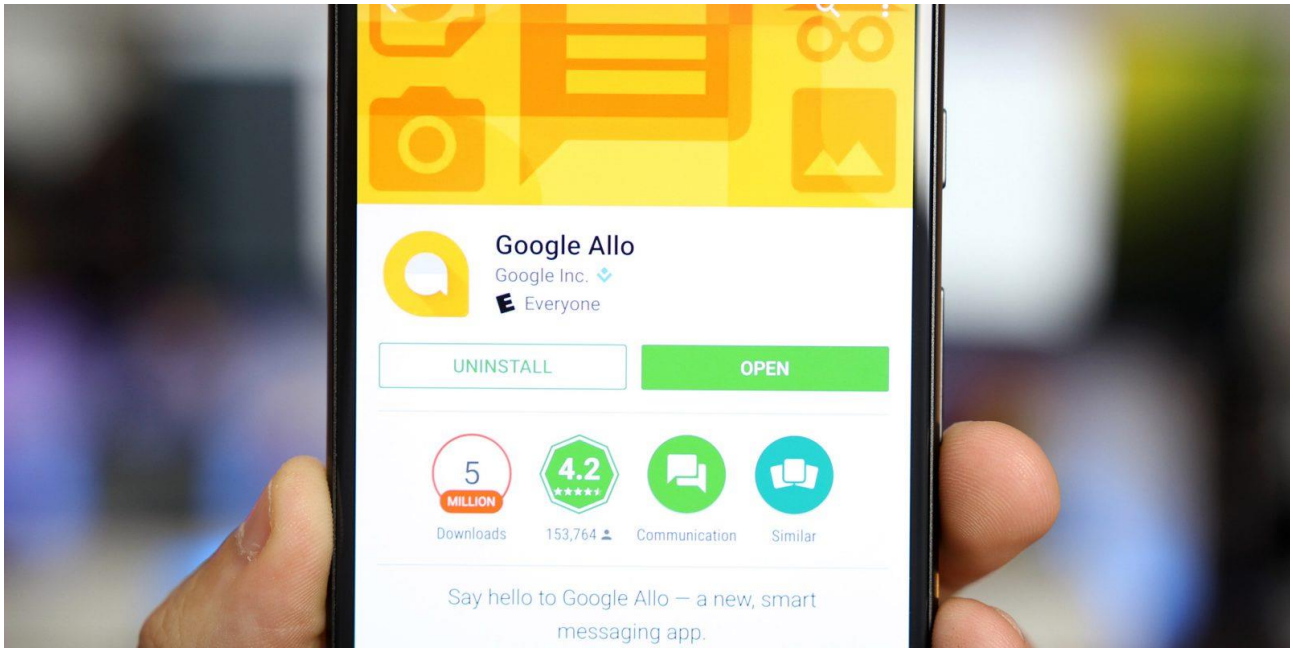


Fig. 12. Google Allo, Image source: https://bit.ly/3uG6I0Q

It must be noted however that not everything goes as planned. In March 2016, Microsoft launched a chatbot called 'Tay', which is an acronym for 'Thinking About You'. Tay was supposed to mimic the personality of a teenage American girl and it was supposed to interact with people on twitter. Tay incorporated AI and machine learning. What that means is that Tay was designed to learn from the users' responses and self-augment its responses accordingly. Yet, humans can be cruel based on the responses that they replace receive. Based on the responses that Tay received in this

case, it started an unfortunate dispute of offensive and racist tweets, that forced Microsoft to take it down in less than 24 hours after its first release.

Fig. 13. The 'Tay' chatbot launched by Microsoft. Image source: https://bit.ly/3a6G5u2

The silver lining here lies in the learning outcomes and a test of capabilities of chatbots. All failures lead to success! Chatbots are nowadays rapidly appearing in many customer services and replacing traditional mobile apps. And while that does not hold true for everything, a lot of online services that require a dedicated app, can now be offered as a chatbot on platforms, such as Facebook messenger.

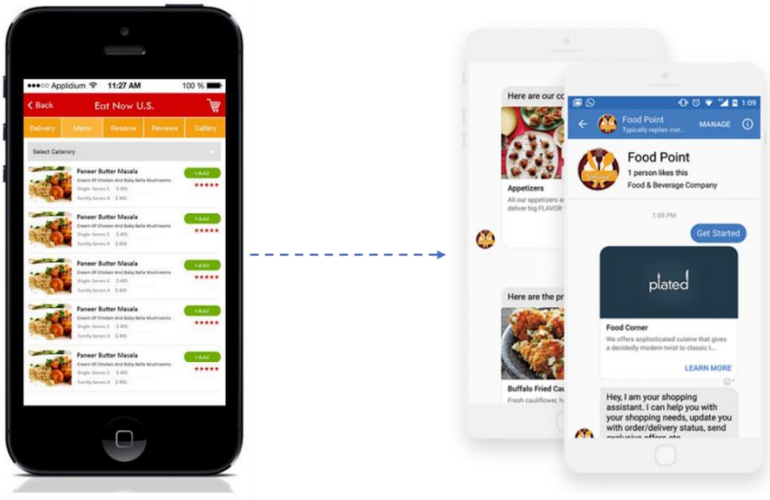


Fig. 14. Online services are now offered by chatbots on messaging apps and platforms.

This also comes since the first in the history of smartphones and apps, people are spending more time in messaging apps than even in social networking apps, which earlier dominated the industry. This has been a catalyst in the push towards chatbots.

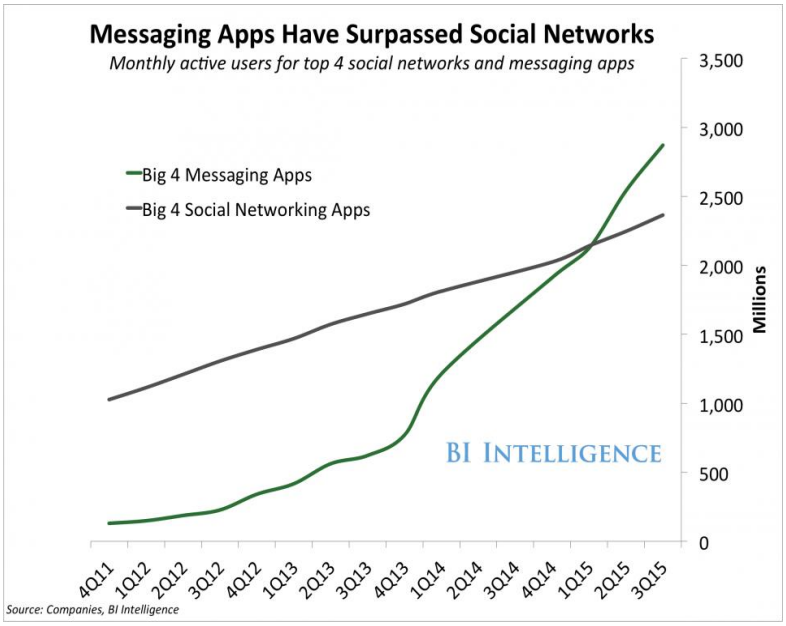


Fig. 15. Messaging apps have surpassed social networks. Image source: https://bit.ly/3sb5WXZ

In 2015, Telegram unlocked its bot platform, allowing developers to build chatbots for serving polls, games, user integration and entertainment (Khan & Das, 2017). Amongst the biggest players in the market, who had a major impact on the buzz of chatbots, is Facebook with the release of the Messenger app in 2016, Skype, Kik and WeChat (Khan & Das, 2017). And while Facebook messenger, WeChat, Telegram, Skype, and so many more platforms are delivery channels, the 'brain' behind the bot is powered by new start-ups like wit.ai, api.ai, motion.ai, Chatfuel, and even big names like Microsoft and IBM Watson.

Nowadays, chatbots potential relies on the exploitation of deep learning technologies, AI, and machine learning techniques. The main challenge for chatbots relates to the AI complete problem. "Completeness is a technical term in theoretical computer science. Essentially says that the problem is as hard as solving the AI problem. In this sense, designing and developing intellectually independent and smart chatbots is as hard as solving the entire problem of building an intelligence machine" (Loizos Michael, expert member). In addition, a meaningful integration of chatbots in education, supported by a theoretical framing is still a challenge. "It has been a little bit of a resistant in using chatbots in higher education here in the US, but it is starting to" (Timothy Wilson, expert member).

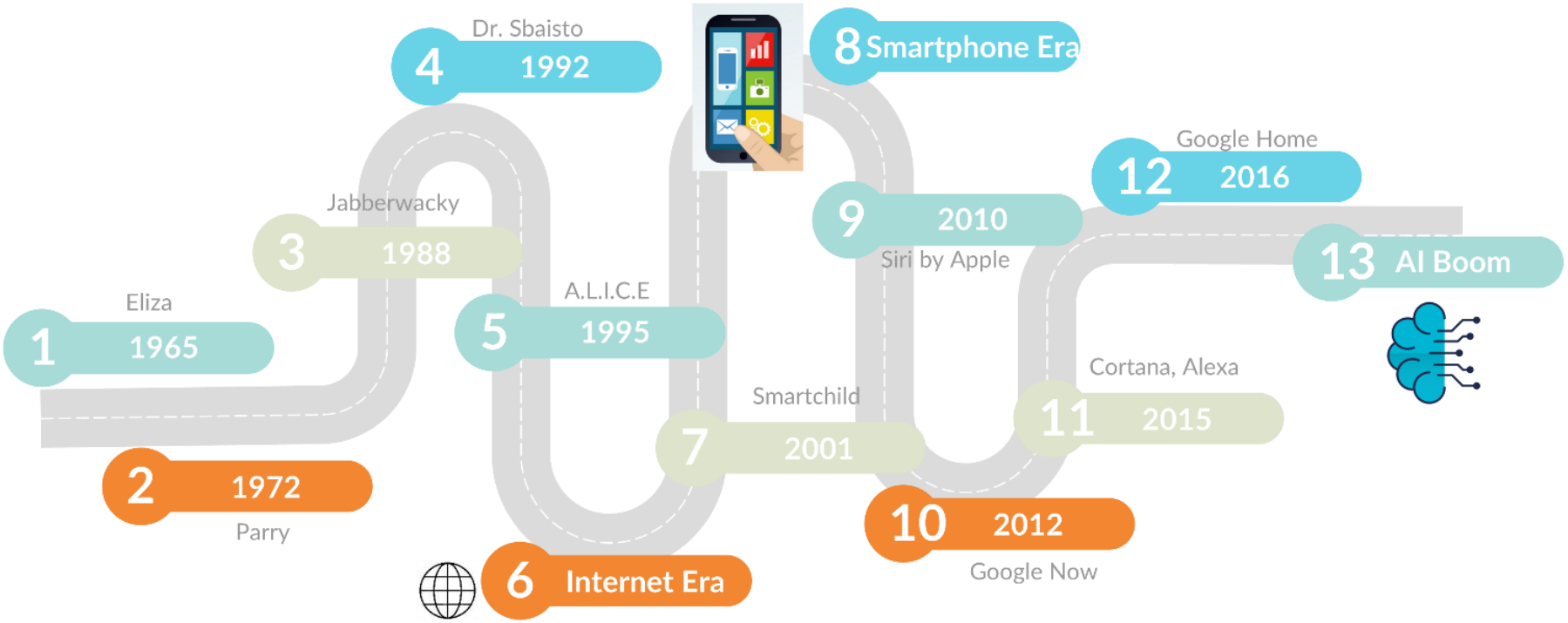


Fig. 16. Timeline of chatbots.

1.4. Areas of chatbot application

While looking at the history of chatbots and their evolution, several chatbot solutions have been presented. Overall, the areas of application cover industries such as:

- Customer services: for product sales and immediate and automated responses to consumers, e.g., Nuance Communications
- Insurance: for finding the correct policy, performing transactions without a human employee, and customer service questions, e.g., ABIE, Kinvey Native Chat , Elafris, Avaamo.ai
- Banking: for checking one's account balance, credit card applications, loan applications, and accomplishing these in more than one language, e.g., Finn AI, Kasisto KAI, IBM Watson Conversation, Personetics
- Financial services: for financial advisory and business trip planning. This type of solution includes budgeting and budget adherence, and hotel and flight recommendations, e.g., Cleo AI, Pana
- Healthcare: for helping with diagnosis, tracking symptoms, and treatment recommendations. Solutions for healthcare providers and individual patients, e.g., Ada Personal Health Guide, Ask Babylon , Symptomate, Woebot.io
- Personal assistants: appearing in mobile devices serving a user's everyday needs and activities, such as, scheduling meetings, checking the user's calendar, making appointments etc. (for instance SIRI, Alexa, Google Now as presented above).
- And lately in Education. We will focus on this area of application in the Chapter 2 (Chatbots in Higher Education).

1.5. Technology used and chatbots' architecture

Chatbots can appear in different types, depending on various parameters, such as: (i) the medium through which they can accessed (i.e., web-based applications that run on a remote serve and can be accessed through a web page, Vs stand-alone applications); (ii) the type of input/ modality (i.e., text vs. speech type of input for the conversation initiation); (ii) the building approaches applied (i.e., retrieval-based models and generative models). The abovementioned classifications were identified in the literature, but also were acknowledged by our experts during the smart interviews. This classification is illustrated in the figure that follows in blue/green shades. In addition, elements of the figure in orange/reddish shades were merely proposed by our experts as potential existing or future classifications of chatbots.

In relation to the medium through which chatbots can be accessed (i.e., web-based, stand-alone applications), the experts agreed on the classification. Expert 2 further proposed the case of specialized devices, by particularly mentioning: *"Alexa for instance is not a web-based application, not either a standalone application. Specialized devices are built for this purpose"* (Loizos Michael, expert member).

With regards to the type of input, chatbots may accommodate different modalities, such as, text, speech type of input, and a combination of voice and text input, but also an avatar representation of

the chatbots, images, photographs, and videos, as proposed by the experts. An illustrative quote from expert 1 is given below: "*In terms of modalities, we can find chatbots with voice, text input. I think it would be also good to include the option for both voice and text so there is a mix modality and then maybe chatbots are with or without an avatar at some point in the future"* (Asbjørn Følstad, expert member). Also, expert 2 proposed the use of images, photographs, and videos as an alternative modality mode. An indicative quote is given: "*I think there are more approaches out there, that could accommodate other modalities. For instance, think of google image search where you can drag an image into the search box and google tries to find similar images. I could imagine the same happening with a chatbot […] It's not just text and speech which are the central modes of communication, but you could envision having other modalities as well"* (Loizos Michael, expert member).

With regards to the building approaches, retrieval-based models, also called rule-based, use a pool of predefined responses (in the form of FAQs) and an algorithm to pick an appropriate response based on the input and/or context. As acknowledged by one of our experts *"…rule-based has to do with how chatbots represent their knowledge and it is not a matter of how they interact with the user*". Eliza is an example of a rule-based chatbot. Although rule-based captures the way the knowledge is represented in the machine, it also implies that most of the time that the knowledge was hand-coded by an expert. Effectively we may say that rule-based machines are machines that were designed by a domain expert, and they do not acquire that knowledge by themselves.

Therefore, this type of chatbots work with predefined rules or keywords. Each rule or keyword needs to be defined by the developer separately to deal with different possible scenarios so that the chatbot understands the questions asked. For example, a student might ask, "What is the registration process for classes?" or more specifically, "How do I register for English 101?" If the chatbot does not understand, the developer must make updates to clarify the response. Simple chatbots are easy for developers to implement but every possible query needs to be hard coded, which can be time consuming.

This contrasts with an intellectually independent chatbot, "…this characterization has to do with how the chatbot acquired its knowledge. It acquired the knowledge by itself, by being trained on data and learning from data". (Expert 2).

Intelligent chatbots are trained to do different tasks. These chatbots are built using artificial intelligence (AI) and NLP technologies to answer and understand user queries. They do not rely on predefined rules to understand; instead, they rely on pretrained models. All questions and answers are recorded and are logged in the algorithm to use for future training. The initial implementation of an intelligent chatbot is difficult because it requires a bigger set of questions but these chatbots are trained to identify and understand when questions are similar, making future updates easier.

Thus, intellectually independent chatbots, corresponding to the generative models, generate responses out of the input with the help of machine learning techniques. Apart from machine learning, another technique for generating responses in the latter category was proposed by one of our experts, that is, the collaborative filtering technique. An indicative quote is given below: *"One technique is called collaborative filtering. It is what Netflix does or Amazon, in which based on the history of choosing books or movies and based on what other people have chosen, a chatbot could suggest to the user to also consider buying this book or see this movie. So that is not necessarily machine*

learning, it is a different approach of gathering data and processing data. Other than that, I agree that there are valid categories and valid variability in each category."

For both types (simple and intelligent chatbots) there are two important tasks that the chatbot performs on the backend. First, it analyzes the request or question asked. Then, it frames an answer to the query. This process seems simple but in practice is complex and works the same whether the chatbot is voice- or text-based.

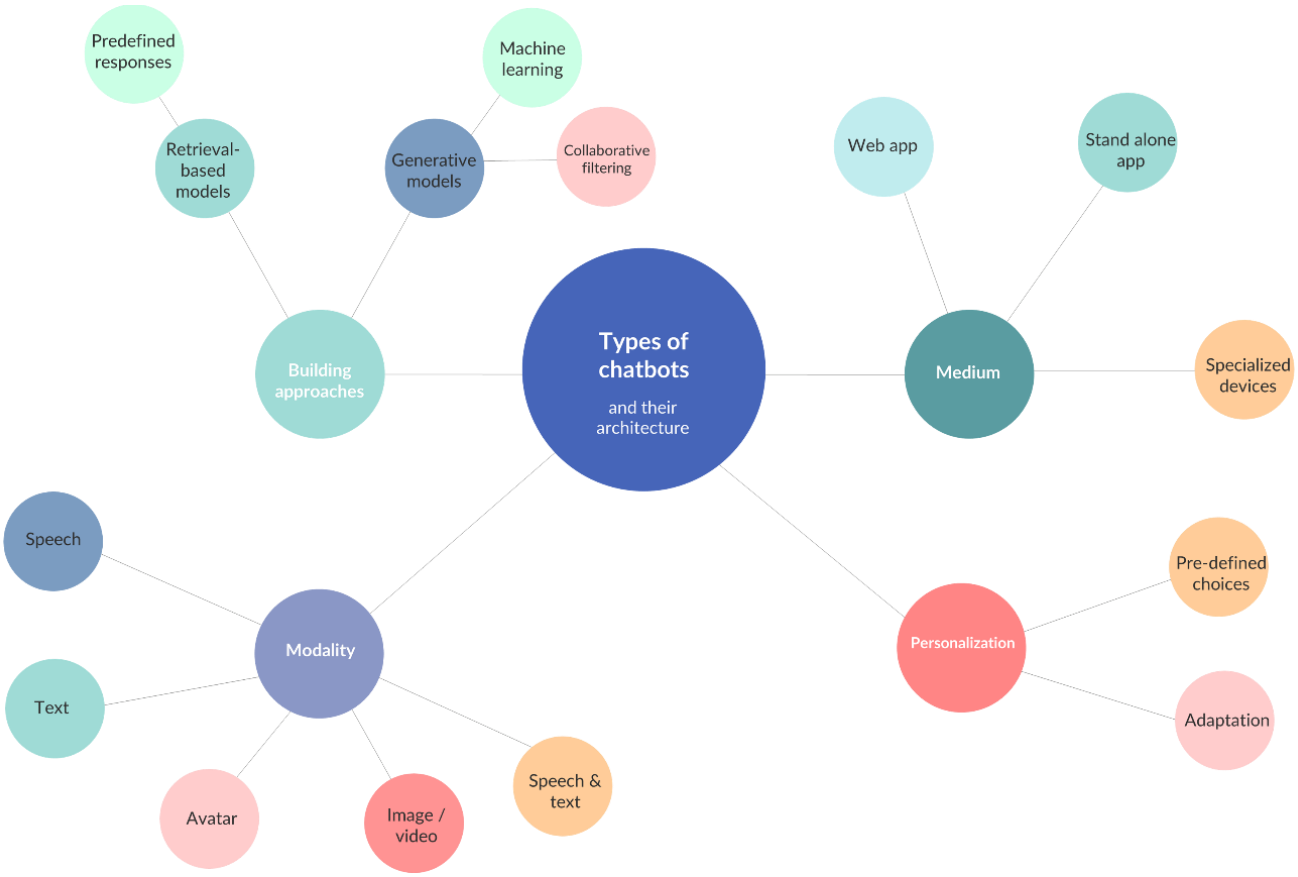


Fig. 17. Types of chatbots and their architecture. Note: Classifications identified in the literature and in experts' interviews appear in blue/green shades; classification identified in the experts' interviews only appear in orange/reddish shades.

A new classification that was proposed by one of our experts deals with the degree of personalization, which we have added as another dimension (see Figure 17). As mentioned by one of the experts, "*For other dimensions that you could use to categorize chatbots, I would say another one is personalization*". Personalization can be then distinguished into the following categories: (a) predetermined choices (e.g., *"Alexa taking in a particular accent is a kind of personalization which is in a sense trivial because there are predetermined choices and you choose one of those"*); (b) adaptation of the machine to the end-user's needs, without the user explicitly asking for this adaptation (e.g., *"another related dimension is the ability to adapt to the user and by that, I mean that, as you interact more and more with the chatbot, it starts adapting its behavior to your needs*").

It must be noted that limitations of chatbots in offering personalization to a satisfied extent at the moment was acknowledged by the experts.

In addition, the ability to identify context (i.e., the setting in which the question or query is asked) and to extract information from the request is the most important part of any chatbot algorithm. If the algorithm fails to understand this, then the chatbot will not be able to respond correctly. Last, for chatbots to be effective, they should provide a consistent and user-friendly customer experience. Both simple and intelligent chatbots should have easy access to data and be able to update that data based on the conversational exchange between the chatbot and the user. This ensures that the chatbot is providing the user with the most relevant and up-to-date information.

Therefore, chatbots range from simple to complex – in the latter case, the aim is to exploit a wide spectrum of artificial intelligence. Nevertheless, modern dialogue systems typically follow a hybrid architecture, combining hand-crafted states and rules with statistical machine learning algorithms (Serban et al., 2017). However, building intellectually independent chatbots remains a major unsolved problem in artificial intelligence research (Serban, et al., 2017).

Perhaps the most critical aspect of implementing a chatbot is selecting the right natural language processing (NLP) engine. In the scenario that the user interacts with the bot through voice, then a speech recognition engine is required. In the paper of Dutta (2017) different chatbot platforms are compared based on their NLP capability and complex feature development ability, namely, Dialogflow.com (Api.ai), Wit.ai, Luis.ai and Pandorabots.com.

Also, the structure levels of chatbot conversations require different programming. For instance, chatbots built for structured conversations are highly scripted, which simplifies programming but restricts the kinds of things that the users can ask, since the questions should be predefined. That kind of chatbots is often used in business-to-business environments for providing immediate responses to customers in the form of frequently asked questions (FAQs). However, once a conversation gets too complex for a chatbot, the call or text window is being transferred to a human service agent. In the case of unstructured conversations, chatbots are powered by AI.

Chapter 2: Chatbots in Higher Education

2.1 Introduction

Education comprises an important field of application for the chatbots as their added value can become significant. The roots of educational chatbots can be found in the early 1970s, with the introduction of pedagogical agents within digital learning environments, the so-called Intelligent Tutoring Systems. The pedagogical agents use AI techniques to enhance and personalize automation in teaching. Several universities and colleges are already using the chatbot technology in specific areas, from admissions to student affairs, career services, financial aid, IT services, general information about the institution or the library, and even test preparations.

Fig 18. Chatbots can provide immediate support to HE students.

In recent years, there has been an attempt to further exploit their capabilities beyond simple querying of information followed by a programmed response and institutions are looking to deploy chatbots with much broader capability. At the same time, *"understandings of these innovations are frequently untheorized and immature*" (Smutny & Schreiberova, 2020, p. 2). According to Winkler and Söllner (2018), chatbots are still in the very beginning of being exploited in education, for supporting students' learning in a meaningful manner. In this chapter, we firstly present key findings from our needs analysis of Higher Education users' needs and expectations for the use of chatbots in Higher Education; secondly, we present chatbot solutions that have been already designed and developed with application in the field of Higher Education, followed by a section focusing on the affordances and weaknesses of using chatbots in Higher Education, along with basic information on pricing. Finally, chapter 2 provides exemplary chatbot use cases, as applied in Higher Education settings, as part of the EDUBOTS piloting.

2.2. Needs Analysis: HE users' needs and expectations.

2.2.1. Insights from the literature

Traditionally, software technology development follows standard processes, including a requirements' gathering phase and a needs assessment. The latter requires a clear understanding of the end users of the software technology to be developed/ refined, their activities and their surrounding environment (Hix & Hartson, 1993). The needs analysis process is a formal, systematic process of identifying and evaluating specific needs of an individual or group of users. Needs are often referred to as 'gaps', or the difference between what is currently done and what should be performed.

In general, the users of a system are not only perceived by those who are going to handle it directly by interacting with its input-output devices. The system designer should record and analyze all direct and indirect users whose activities may be affected by the introduction of the system, and which may affect its acceptance and successful operation. Therefore, users include, among others, those who interact directly with the system, those who control or manage the direct users, those who receive the system output (output), those who take financial decisions regarding the system and those who use competing systems. Based on that, Macaulay (1995) distinguishes the users in the following three categories: primary users, secondary users, trivial users. Other views on the initial analysis phase suggest that all stakeholder groups should be recorded and considered. Stakeholders are all those who have a direct or indirect interest in the introduction and use of the system (MacAulay, 1995).

While the presence of chatbots, on web platforms and/or standalone applications, is already substantial in customer services, business webpages, products sales, and in health interventions, their use for educational purposes is still in its infancy. Higher education comprises an important field for the application of chatbots, especially for large-scale use. A meaningful integration of chatbots in higher education presupposes a good understanding of users' needs and expectations, as well as, an examination of their perceptions towards educational technology (Keller & Cernerud, 2002; Popovici & Mironov, 2015), the adoption of an appropriate pedagogy (Gonda, Luo, Wong, & Lei, 2018), and last but not least, a confrontation of technological challenges and potential limitations, that come relate to the NLP research field (Yan, Castro, Cheng, & Ishakian, 2016).

Previous research focusing on the examination of higher education users' needs and expectations on the use of chatbots in education has already been done to some extent, projecting possible uses of chatbots in education (e.g., Gupta, et al., 2019; Lee, et al., 2019; Thies et al., 2017). Findings of those studies outline the potential use of chatbots as career advisors (see Lee, et al., 2019), or intelligent tutors answering student questions (Gupta, et al., 2019). Thies et al. (2017) in their work analyzed data from exploratory interviews and focus groups with students for the development of chatbot personalities. Their findings suggest that students need a chatbot helping them become knowledgeable and successful in their career aspirations; an entertaining chatbot, with whom they could share fun experiences; and a chatbot to listen to them, help them improve their soft skills and help them become desirable in their social circles (Thies et al., 2017). Moreover, users' expectations

for the design characteristics of chatbots have been already explored (i.e., recognition; visibility of system status; anthropo-morphism in communication; knowledge expertise, linguistic consistency; realis-tic interaction) which may enhance the feeling of trust and support students in a personalized and interactive way (Stathakarou, et al., 2020).

2.2.2. Key findings from the EDUBOTS pilots

In EDUBOTS, meaningful user groups and stakeholders include higher-education students participating in the project's pilots, instructors, mentors, and the consortium partners. In the EDUBOTS project, we focused on fulfilling the needs and expectations of higher-education students and educators, with the introduction of chatbots in higher education settings. In achieving that, we obtained insights from consortium partners' views and needs in relation to chatbots' integration in higher education as well as from end users (i.e., higher-education students and educators).

Table 2. Summary of pilots and data collection

Pilot number	**PILOT 1**	**PILOT 2**	**PILOT 3**	**PILOT 4**
Academic Semester	Spring 2020	Fall 2020	Spring 2021	Fall 2021
Data collection	Semi-structured interviews (n=21)	Semi-structured interviews (n=9) Online survey (n=159)	Online survey (n=158)	Online survey (n=114)
Data analysis	Thematic analysis (interviews)	Thematic analysis (interviews) Descriptive statistics, Comparison tests (survey data)	Descriptive statistics, Comparison tests, e.g., t-test, ANOVA (survey data)	Descriptive statistics, Comparison test, Comparison tests, e.g., t-test, ANOVA (survey data)

In pilot 1, field work was implemented with the conduction of semi-structured interviews, aiming to identify users' needs in relation to the use of chatbots in education. In pilot 1, semi-structured interviews were conducted with 21 Higher Education users, specifically with ten higher education instructors (3 females and 7 males), eight undergraduate students (2 females and 6 males) from different local universities in Cyprus and with different backgrounds (i.e., Multimedia and Graphics Arts, Computer Science, Sports & Exercise Science) and 3 members from the consortium organizations (FOI, University of Leeds, University of Granada). The interview protocol is given in Annex II. All the participants consent to anonymously use the data for research purposes. All the interviews were recorded and then transcribed. The original recordings will be then deleted. A thematic analysis was conducted, of the major ideas in the interviews and across all the interviews. The results of the thematic analysis are reported (see section below) but no quotes or other specific data are associated with names of the interviewees.

As part of pilots 2-4, additional data were collected and analyzed for the needs' assessment task. In pilot 2, data were collected via an online survey which has been constructed accounting on

the themes that emerged from the analysis of the data from the first pilot. The survey was prepared in two versions: one targeting higher-education educators and another one targeting higher-education students (see Annexes IV and V). The survey items were also translated in participants' native language, when needed. The survey has been shared with all academic partners of EDUBOTS, so that survey data were collected from users from four different countries. Survey data were then analyzed quantitatively, through the calculation of descriptive statistics and comparison tests. In addition, as part of the second pilot, CYENS prepared and shared with all academic partners a revised version of the pilot-1 interviews protocol (see Annex II). All the interviews were recorded and then transcribed. The original recordings were then deleted. We present the major ideas in the interviews and across all the interviews, along with useful excerpts. No quotes or other specific data are associated with names of the interviewees.

In pilots 3 and 4, data were collected via survey only, and interviews were not conducted, as it appeared that from pilot 1 to pilot 2 the same needs were identified repeatedly in interviewees' responses. For that reason, we administered an online survey, using the same items as in pilot 2 (see Annexes IV and V), but with the addition of two more questions for the educators' survey, in relation to uses of chat tools in their classes, and a general question on what they would like an automated assistant chatbot to do, if available within a chat tool. We added those questions, because we sought to also examine whether and how educators use chat tools in their classes, and how for which scenario they envision the integration of a chatbot within a chat tool for addressing their needs. Using data from all pilots we proceeded with comparison tests (independent sample t-test, ANOVA test) to explore any potential differences that might exist in the needs and expectations of chatbots uses in different countries and among educators and students. Further to the above, we conducted a thorough user experience (UX) analysis upon the pilot of Differ chat and BO chatbot in pilots 2 and 3. The aim of the UX analysis was to first assess the usability of BO and Differ, and to examine whether users' needs are being addressed with the chatbot solutions that were tested in pilot 2. The interview protocols, the survey items and the UX report are provided in a detailed report dedicated to the needs analysis work and can be retrieved through the EDUBOTS website. Below, we present the key findings from this work. A more detailed report with all findings and the UX report can be retrieved through the EDUBOTS website.

2.2.2.1. Key findings from the interviews

Interviewees' expectations for the use of chatbots were clustered into different themes (see Fig. 19). Chatbots were envisioned as facilitators in the teaching and learning process in supporting: (1) formative assessment, (2) remote tutoring, and (3) administrative work. Also, chatbots were envisaged as (4) digital assistants in research-related processes (e.g., grant preparation), and in (5) other scenarios not related to education (i.e., health care, customer service, in society addressing citizens' needs, and in the fitness industry). Interviewees were explicitly questioned about their potential needs and expectations for the use of chatbots for social bonding; this type of use was tolerated by some of the interviewees, while others did not express concerns. We elaborate on the emerging themes related to educational applications, further below.

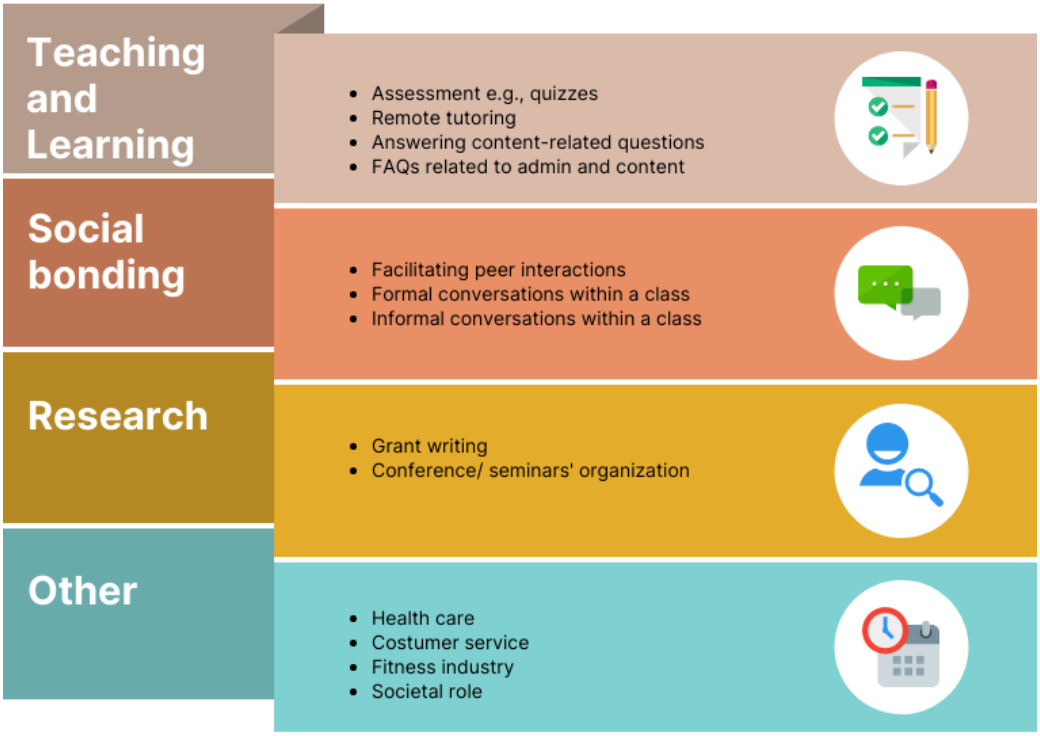


Fig.19. Interviewees' expectations on the use of chatbots

Theme 1: Teaching and Learning

According to the interviewees, chatbots can help to facilitate the teaching and learning process; specifically, chatbots were envisioned as virtual/ digital assistants, smoothing the implementation of formative assessment, by providing grades and/or qualitative and personalized feedback to the students, but also for conducting short quizzes, during an academic semester, for assessing students' conceptual understanding. The interviewees mentioned that this need is imperative especially in large-scale classes, in which the provision of personalized feedback, in a formative and qualitative manner, during the semester might be an unrealistic scenario. Indicative quotes are given below.

"After each lecture you could have some questions, to test students' conceptual understanding" (instructor 2, male, Natural Sciences)

"A chatbot to offer personalized feedback to the students. If there was such a bot, that would be a blessing. Especially for large audiences. But it should be a smart chatbot, so that it does not give feedback just to give, but a chatbot that is really capable of judging the quality of the work" (instructor 3, male, Multimedia and Graphics Art)

"[…] or if the chatbot could be used for giving feedback to our assignments" (student 2, male, Multimedia and Graphics Art)

Furthermore, chatbots could add value by offering content support to the students, in course related topics, especially for addressing students' content-related questions. In this respect, chatbots

could act as remote tutors, offering tutorials to the students, sharing, and filtering relevant to the course resources and having a mediating role to the instructor to intervene when needed.

"If we could have a chatbot offering support step-by-step to the students, that would be useful. That is, to have a chatbot, as a remote tutor, guiding the students on how to solve a problem" (instructor 7, male, Civil Engineering and Geoinformatics)

"We are interested in having something to assist us on bibliography search or to help us find a specific reference […] Also, when you are learning a new language would be nice to have some advice […] chatbots could offer useful tutorials in coding" (instructor 8, male, Computer Architecture and Technology)

"Anything related to lecture content, so content that the students need to learn about; I think the chatbot can replace the teacher in some basic content concepts" (instructor 9, male, Department of Foreign Languages and General Educational Discipline)

The 24/7 availability of chatbots could then allow students to seek support, at any time of the day. Some of the students referred to the value of having a chatbot acting as a remote tutor, especially for those who study late hours and when it is not feasible to directly communicate with the educator. Also, it appears that the students need a facilitator in understanding content-related concepts; as proposed, the user could enter some keywords in the conversation and then, the chatbot could filter and retrieve relevant resources.

"When I need help and support at any time of the day, instead of waiting for an answer from the teacher via email or until I go back to class or through some other social media platforms, I think it would be a good solution to have a chatbot answering my questions" (student 2, male, Multimedia and Graphics Art)

"A chatbot could do what the google scholar does, for example. Where you enter some keywords and then the engine provides you with relevant resources […] For example, you study for the Chemistry class and you come across terms and concepts that you do not know, e.g., molecular coupling. Then you could ask the chatbot for help for those concepts." (Student 6, male, Multimedia and Graphics Art)

"I usually read very late at night and when I have questions it is not possible to directly reach my teacher at that time. If I could have a chatbot answering my questions, well, this way my questions will be answered faster" (student 8, female, Engineering and Geoinformatics Engineering)

Moreover, few of the educators stated that there is a need of having a chatbot that responses to commonly asked questions, posed by students; yet the chatbot should be able to recognize when the educator should intervene in the conversation, and at that point, it should initiate the conversation/ communication between the student and the educator.

"The bot could answer students' content-related questions that are trivial; ideally, the chatbot should be able to understand when I – as an instructor- need to intervene in the chatbot-student

conversation and thus the chatbot should initiate such a communication with my students" (instructor 3, female, Multimedia and Graphics Art)

Apart from that, most of the interviewees acknowledged the potential added value of chatbots in responding to frequently asked questions (FAQs), while most of those often relate to administration (e.g., deadlines). Educators referred to the struggle of handling a large amount of administrative work, on the top of their teaching and research duties, such as, collecting assignments, sending to the students' reminders for their assignments' deadlines, responding to FAQs related to the course, and even responding to questions on how to handle a submission of an assignment. It was proposed by the interviewees that chatbots could provide a lot of support in this direction, thus, reducing the often-observed workload due to such matters. As one of the instructors explicates…

"[…] in this way, the integration of chatbots in education could reduce the teacher's time by saving valuable hours." (Instructor 2, male, Natural Sciences)

The particular educator emphasized that it is more important having a chatbot responding to content-related issues, than administrative matters, since the latter can be also addressed now through the mass amount of information that students already receive in social networks and e-learning systems.

Theme 2: Research

Interestingly, two of the educators expressed the need for having a virtual assistant for fulfilling a grant application (mostly on the technical parts of an application), and for offering support during the organization of conferences and seminars (e.g., a chatbot contacting the participants, responding to queries, sending the agenda). An indicative quote is given below.

"My work is not so standard, and I often fight with colleagues to do the basics, to organize a meeting. Another thing is that I would like to have a tool guiding me, as a researcher, to fill in a grant application. Something like a form filler, something effortless to the user." (Instructor 7, male, Civil Engineering and Geoinformatics)

Theme 3: Social bonding

IVn the social bonding scenario, one of the educators acknowledged the need for using a chatbot for social bonding among students and/or with other colleagues. However, the rest of the educators appeared to be reluctant for the use of chatbots for social bonding; they explicated their view with reference mainly to cultural factors that might hamper such a use. Also, based on their own personal experiences, social bonding with peers was not considered an issue, thus, making the application of chatbots in this scenario inappropriate to them. However, three of the student interviewees referred to the need of having a chatbot facilitating inter-social bonding with peers and with mentors. Such a solution would establish connections among students or between students and mentors, which could

then comprise channels for communication for resolving any course or university-related issues that the students confront during the academic semester.

"A chatbot could actually accommodate social bonding with peers. Ok, with friends and fellow students there are other means of communication, such as messenger, Viber, or anything else that we use now. But there are some researchers, so-called mentors, who are at the university, and we do not see them often, but they help in some way to conduct the course. Other times we have tutoring with mentors. With the use of a chatbot, we will be able to communicate with them, just like having a remote tutor at any time, without requiring getting to know the mentor in person." (Student 8, female, Engineering and Geoinformatics Engineering)

Theme 4: Other needs

Other areas of application in which chatbots could address interviewees 'needs, encompass health care, the customer service, and the fitness industry areas. Finally, one of the instructors proposes a societal role that chatbots could have; for filling several needs of citizens and facilitating the realization of governmental services accessible to the public.

2.2.2.2. Key findings from the survey

Survey results from educators

A total of 113 educators participated in the survey across all pilots, out of which, 14 were retrieved from Cyprus (12.4%), 41 from Croatia (36.3%), 11 from Spain (9.7%) and 47 from the UK (41.6%). In relation to their gender, 52 educators were males (46%), 56 females (49.6%) and 4 preferred not to specify (4.4%). Their preferred uses of chatbots in HE (i.e., percentages of their responses to the 5-point Likert-scale items) are presented in table 3 and figure 20. A percentage of 47.8% of the responders consider as very important the use of chatbots for responding the FAQs related to admin topics of a course, a 31.0% considered also as very important the use of chatbots for addressing students' FAQs related to content, and a 27.5% considered also as very important the use of chatbots for performing quizzes in the class.

Table 3. Educators' preferred uses of chatbots in HE – Data from all partners across all pilots

Item code	Item / importance	1 (not important at all)	2	3	4	5 (very important)
		%	%	%	%	%
E1	Quizzes	11.5	15.9	21.2	23.0	27.4
E2	Tutorial	23.9	13.3	21.2	16.8	24.8
E3	Resources sharing	4.4	8.0	9.7	11.5	13.3
E4	Course evaluation	14.2	11.5	24.8	26.5	23.0
E5	FAQs Admin	12.4	6.2	11.5	22.1	47.8
E6	FAQs content	7.1	2.7	15.9	24.8	31.0
E7	Content-related questions	22.1	11.5	15.9	28.3	22.1
E8	Icebreaking	21.2	15.0	23.9	24.8	15.0
E9	Mentors	19.5	15.0	20.4	20.4	24.8

E10	Study groups	20.4	19.5	19.5	22.1	18.6
E11	Peer-to-peer interactions	23.9	19.5	17.7	21.2	17.7
E12	Informal conversations	21.2	24.8	21.2	15.9	15.9
E13	Formal conversations	0.9	8.8	4.4	8.0	7.1
E14	Find like-minded people	24.8	17.7	20.4	21.2	15.9

1: not important at all, 5: very important.

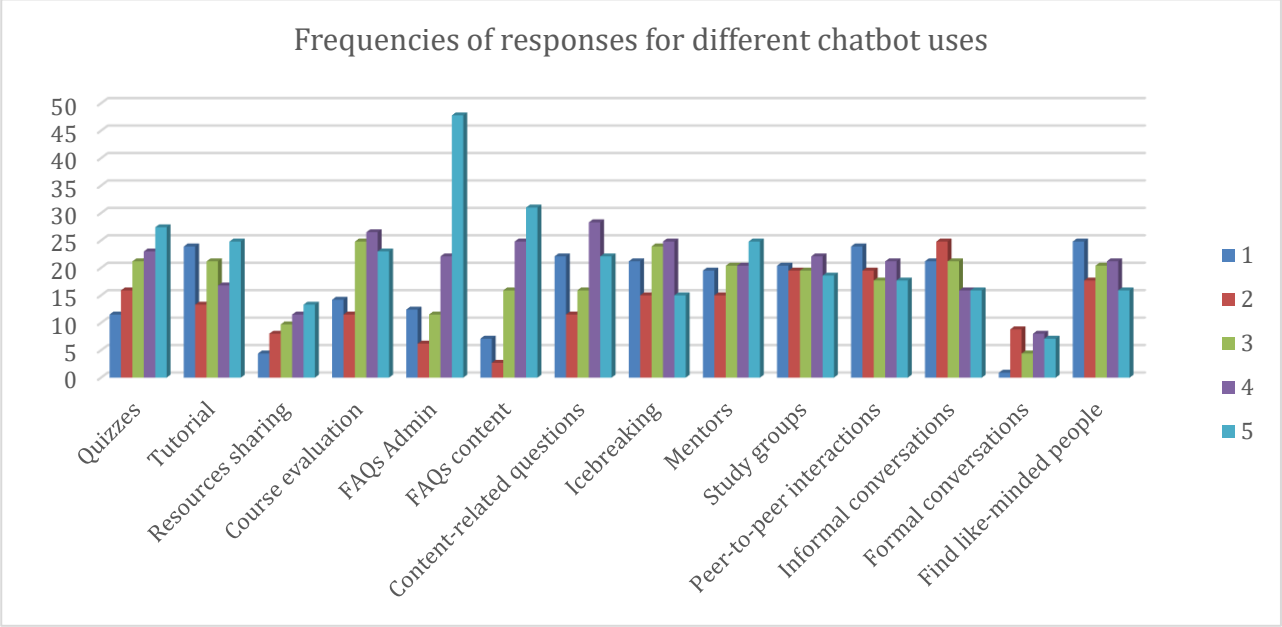


Fig. 20. Frequencies of educators' responses to the 5-point Likert scale items on different chatbot uses. Note: 1: not important at all, 2: not important, 3: neutral, 4: important, 5: very important

Survey results from students

A total of 318 students participated in the survey across all pilots, out of which, 44 were retrieved from Cyprus (13.8%), 171 from Croatia (53.8%), 36 from Spain (11.3%) and 67 from the UK (21.1%). In relation to their gender, 142 students were males (44.7%), 170 females (53.5%) and 6 preferred not to specify (1.9%). Their preferred uses of chatbots in HE (i.e., percentages of their responses to the 5-point Likert-scale items) are presented in table 4 and figure 21. A percentage of 36.2% of the responders consider as very important the use of chatbots for sharing resources relevant to a course, a 34.3% also consider the use of chatbots important for tutoring and for facilitating the communication with the mentors in a class. Also, the 32.4% of the students considers as very important the use of chatbots for addressing their content-related questions, while a 28.3% and 28.6% considers as very important the use of chatbots for FAQs related to admin matters and FAQs related to the course content, respectively.

Table 4. Educators' preferred uses of chatbots in HE – Data from all partners across all pilots

Item code	Item / importance	1 (not important at all)	2	3	4	5 (very important)
		%	%	%	%	%

E1	Quizzes	8.2	9.7	25.5	19.5	16.0
E2	Tutorial	4.7	7.2	21.1	32.7	34.3
E3	Resources sharing	6.3	10.4	17.9	29.2	36.2
E4	Course evaluation	7.5	15.1	24.8	29.6	23.0
E5	FAQs Admin	7.9	7.2	14.8	20.4	28.3
E6	FAQs content	5.7	7.2	19.5	30.5	28.6
E7	Content-related questions	8.2	8.2	19.2	31.4	32.4
E8	Icebreaking	10.4	17.0	31.1	23.0	18.6
E9	Mentors	6.0	10.4	18.9	30.5	34.3
E10	Study groups	11.6	17.0	26.4	25.8	19.2
E11	Peer-to-peer interactions	9.7	16.0	28.6	26.7	18.6
E12	Informal conversations	13.5	18.2	26.4	25.5	16.4
E13	Formal conversations	7.5	11.6	22.6	24.5	15.1
E14	Find like-minded people	13.8	16.0	28.9	25.5	15.4

1: not important at all, 5: very important.

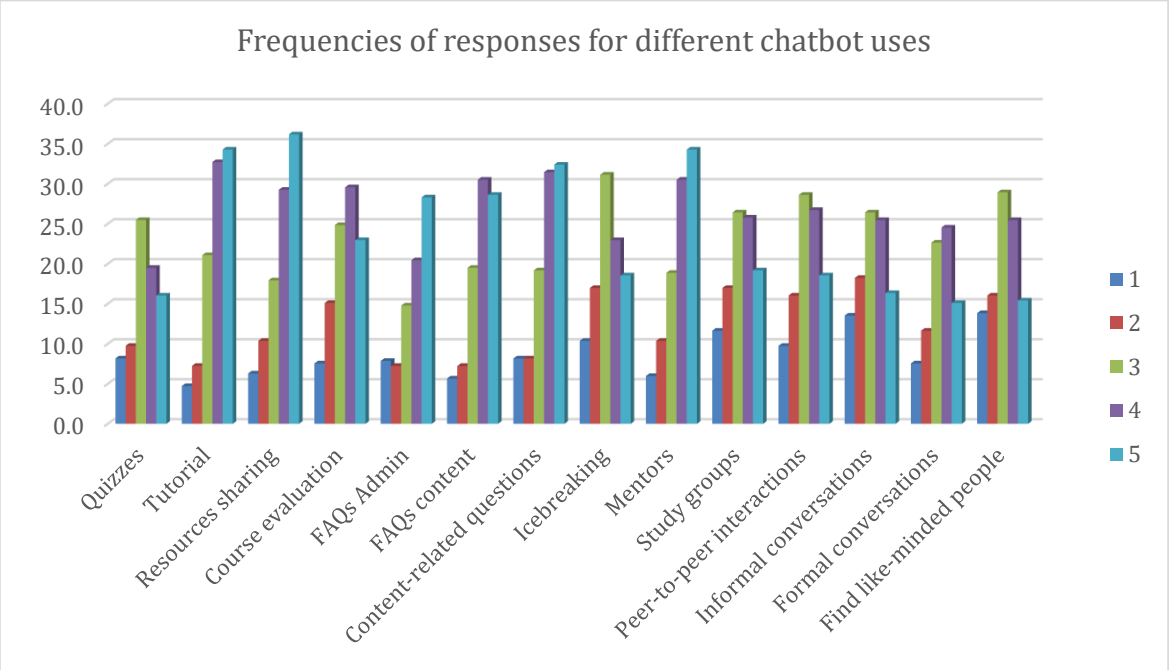


Fig. 21. Frequencies of educators' responses to the 5-point Likert scale items on different chatbot uses. Note: 1: not important at all, 2: not important, 3: neutral, 4: important, 5: very important

Comparison tests

First, an independent sample t-test has revealed that there are differences in educators' and students' perceptions of their needs for the use of chatbots in education. Specifically, it has been revealed that the students assessed as more important the use of chatbots for tutorials (M = 3.84, SD = 1.11) compared to educators' responses (M = 3.05, SD = 1.50) $t_{(157.991)} = -5.121$, p = .000. Likewise, the students assessed as more important the use of chatbots for responding to content-related questions (M = 3.72, SD = 1.23) compared to educators' responses (M = 3.16, SD = 1.46) $t_{(171.458)} = -3.579$, p = .000, as well as for facilitating the communication between mentors and students ($t_{(168.967)}$

= -3.988, p = .000), peer-to-peer interactions in the class ($t_{(172.526)}$ = -2.568, p = .011), and informal conversations within a class ($t_{(428)}$ = -2.278, p = .023). The statistically significant results are presented in the table that follows.

Table 5. Independent sample t-test results – Comparison of students' and educators' responses with survey data from pilots 2-4

Use case	Educators (n=113)		Students (n=318)		t
	mean	SD	mean	SD	
Tutorials	3.05	1.50	3.84	1.11	-5.121***
Content-related questions	3.16	1.46	3.72	1.23	-3.579***
Mentors	3.15	1.45	3.76	1.19	-3.988***
Peer-to-peer interactions	2.89	1.44	3.28	1.22	-2.568*
Informal conversations	2.80	1.37	3.12	1.27	-2.278*

* p <0.05, ** p < 0.01, *** p < 0.001, ns = non-significant

Differences across countries (Cyprus, Croatia, UK, Spain) have been also detected. First, a one-way ANOVA revealed that there were statistically significant differences in educators' perceptions of their needs for the use of chatbots in education between at least two countries, and specifically for the use of chatbots for tutorials ($F_{(3, 109)}$ = 7.863, p < 0.001), resources sharing ($F_{(2, 50)}$ = 4.155, p = 0.021), responding to FAQs related to admin matters ($F_{(3, 109)}$ = 7.347, p < 0.001), responding to FAQs related to the course content ($F_{(2, 89)}$ = 8.532, p < 0.001), and helping people to find like-minded people ($F_{(3, 109)}$ = 3.158, p = 0.028). Also, an independent sample t-test (missing data from two countries for this variable) has shown that educators from Cyprus assessed as more important the use of chatbots for formal conversations (M = 5.00, SD = 0.00) compared to educators' responses from Croatia (M = 3.29, SD = 1.22) $t_{(30.00)}$ = 7.826, p = .000. The results are presented in table 6. Bonferroni Tests for multiple comparisons found that the mean value of educators' ratings were significantly different for the use cases of chatbots concerning: (i) the tutorials use case between CYENS and the University of Leeds (p = 0.003); educators from Cyprus consider more important (M=3.93, SD=1.49) this use case as compared to educators from the UK (M=2.40, SD=1.45); (ii) the FAQs admin use case between the University of Granada and CYENS (p < 0.001) and FOI (p = 0.004); educators from Spain consider as less important (M=2.36, SD=1.57) this use case as compared to educators from Cyprus (M=4.79, SD=0.58) and Croatia (M=3.90, SD=1.41); (iii) the finding like-minded people use case between the University of Leeds and CYENS (p = 0.027); educators from Cyprus consider as more important (M=3.86, SD=1.10) this use case as compared to educators from the UK (M=2.64, SD=1.44). In relation to codes E3 (resources sharing) and E6 (FAQs content) data from the UK and Spain respectively were missing; thus the ANOVA test was performed with data from the other three participating countries. A Dunnett T3 Test (equal variances not assumed) for multiple comparisons found that the mean value of educators' ratings were significantly different for the use cases of chatbots concerning resources sharing use case between CYENS and FOI (p < 0.001) and CYENS and the University of Granada (p = 0.001); educators from Cyprus consider as more important (M=5.00, SD=0.00) this use case as compared to educators from Croatia (M=3.60, SD=1.19) and Spain (M=2.64, SD=1.50). A Bonferroni Test for multiple comparisons found that the mean value of educators' ratings was significantly different for the use cases of chatbots concerning the FAQs

content between the University of Leeds and CYENS (p = 0.017) as well as between the University of Leeds and FOI (p = 0.001), with educators from the UK considering less important this use case, as compared to educators from the other two countries.

Table 6. ANOVA results analysis – Comparison of educators' responses across partner countries with survey data from pilots 2-4

Use case	Cyprus (n=14)		Croatia (n=41)		Spain (n=11)		UK (n=47)		F
	M	SD	M	SD	M	SD	M	SD	
Quizzes	3.64	1.34	3.40	1.32	2.45	1.37	3.53	1.33	2.159^{ns}
Tutorial	3.93	1.49	3.61	1.18	2.64	1.63	2.40	1.45	7.863^{***}
Resources sharing	5.00	0.00	3.60	1.19	2.64	1.50	-	-	4.155^{*}
Course evaluation	4.07	1.21	3.22	1.39	2.64	1.50	3.36	1.21	2.653^{ns}
FAQs Admin	4.79	0.58	3.90	1.41	2.36	1.57	3.91	1.27	7.347^{***}
FAQs content	4.36	0.84	4.35	0.71	-	-	3.38	1.39	8.532^{***}
Content-related questions	3.64	1.60	3.51	1.31	2.45	1.57	2.89	1.46	2.774^{ns}
Icebreaking	3.50	1.65	3.20	1.12	2.45	1.37	2.74	1.42	2.081^{ns}
Mentors	3.86	1.51	3.37	1.36	2.64	1.63	2.89	1.42	2.435^{ns}
Study groups	3.07	1.49	3.29	1.36	2.64	1.50	2.79	1.40	1.250^{ns}
Peer-to-peer interactions	3.36	1.50	3.02	1.42	2.64	1.50	2.70	1.43	0.988^{ns}
Informal conversations	3.57	1.55	2.88	1.38	2.36	1.57	2.62	1.21	2.229^{ns}
Formal conversations	5.00	0.00	3.29	1.22	-	-	-	-	$t=7.826^{***}$
Find like-minded people	3.86	1.10	2.88	1.35	2.45	1.57	2.64	1.44	3.158^{*}

* p <0.05, ** p < 0.01, *** p < 0.001, ns = non-significant. Note: missing data for E3 (UK), E5 (Spain), E14 (UK and Spain.

Second, a one-way ANOVA revealed that there were statistically significant differences in students' perceptions of their needs for the use of chatbots in education between at least two countries for the use of chatbots in quizzes ($F_{(2, 48)} = 3.696$, p = 0.026), content-related questions ($F_{(3, 313)} = 3.395$, p = 0.018), ice-breaking activities ($F_{(3, 314)} = 12.79$, p = 0.000), facilitating the communication among mentors and students ($F_{(3, 314)} = 3.039$, p = 0.029), and assisting in finding like-minded people ($F_{(3, 313)} = 2.836$, p = 0.038). The results are presented in table 7. Bonferroni Tests for multiple comparisons found that the mean value of students' ratings were significantly different for the use cases of chatbots concerning: (i) the content-related questions use case between the University of Granada and the University of Leeds (p = 0.017); students from UK consider more important (M=4.09, SD=0.97) this use case as compared to students from Spain (M=3.33, SD=1.39); (ii) the ice-breaking activities between the University of Leeds and CYENS (p < 0.001), FOI (p < 0.001) and the University of Granada (p = 0.021). Students from UK consider more important (M=3.99, SD=1.04) this use case as compared to students from the other participating countries; (iii) the find like-minded people use case between CYENS and the FOI (p = 0.036), with the students from Croatia considering more important (M=3.92, SD=1.18) this use case as compared to students from Cyprus (M=3.36, SD=1.22); (iv) the quizzes use case between CYENS and the FOI (p = 0.027), with the students from Cyprus considering more important (M=3.77, SD=1.16) this use case as compared to students from Croatia (M=3.23, SD=1.22).

Table 7. ANOVA results analysis – Comparison of students' responses across partner countries with survey data from pilots 2-4

Use case	Cyprus (n=44)		Croatia (n=171)		Spain (n=36)		UK (n=67)		F
	M	SD	M	SD	M	SD	M	SD	
Quizzes	3.77	1.16	3.23	1.22	3.19	1.24	-	-	3.696^{*}
Tutorial	3.73	1.26	3.98	1.14	3.56	1.16	3.75	0.89	1.952^{ns}
Resources sharing	3.73	1.19	3.94	1.28	3.44	1.27	3.63	0.98	2.250^{ns}
Course evaluation	3.50	1.25	3.50	1.20	3.33	1.37	3.36	1.12	0.317^{ns}
FAQs Admin	3.86	1.15	3.68	1.33	3.53	1.44	-	-	0.666^{ns}
FAQs content	3.70	1.25	3.85	1.15	3.33	1.50	3.61	1.07	1.133^{ns}
Content-related questions	3.70	1.25	3.66	1.26	3.33	1.39	4.09	0.97	3.395^{*}
Icebreaking	3.02	1.19	2.96	1.18	3.28	1.30	3.99	1.04	12.79^{***}
Mentors	3.36	1.22	3.92	1.18	3.83	1.18	3.61	1.19	3.039^{*}
Study groups	3.18	1.23	3.35	1.28	3.44	1.40	2.90	1.14	2.413^{ns}
Peer-to-peer interactions	3.07	1.23	3.29	1.26	3.36	1.29	3.36	1.08	0.591^{ns}
Informal conversations	2.95	1.26	3.08	1.28	3.33	1.39	3.25	1.20	0.876^{ns}
Formal conversations	3.25	0.97	3.46	1.16	3.44	1.33	3.04	1.28	2.004^{ns}
Find like-minded people	3.05	1.28	3.21	1.23	3.47	1.25	2.79	1.27	2.836^{ns}

* p <0.05, ** p < 0.01, *** p < 0.001, ns = non-significant. Note: missing data for E1, E5 from UK.

2.2.2.3. Concluding remarks from the needs analysis

Overall, survey data were collected from 432 Higher Education users across all pilots, 113 educators (26.2%) and 318 students (73.6%), out of which 58 were retrieved from CYENS (13.4%), 212 from FOI (49.1%), 47 from the University of Granada (10.9%) and 114 from the University of Leeds (26.4%). 194 survey respondents were males (44.9%), 226 were females (52.3%) and 11 individuals preferred not to say (2.5%). In addition, interview data have been collected during pilots 1 and 2 for a more in depth understanding and examination of educators and students' needs in relation to the use of chatbots in education. In pilots 3 and 4, data were collected via survey only, and interviews were not conducted, as it appeared that from pilot 1 to pilot 2 the same needs were identified repeatedly in interviewees' responses. In pilot 1 semi-structured interviews were conducted with 21 Higher Education users, specifically with ten higher education instructors (3 females and 7 males), eight undergraduate students (2 females and 6 males) from different local universities in Cyprus and with different backgrounds (i.e., Multimedia and Graphics Arts, Computer Science, Sports & Exercise Science) and 3 members from the consortium organizations (FOI, University of Leeds, University of Granada). partners of the consortium. In pilot 2, a total of nine interviews were conducted with participants; 5 interviews were conducted with educators (CYENS=2, FOI=3) and 4 interviews with students (CYENS=2, FOI=2).

It has been shown that higher-education students and instructors have several needs related to the teaching and learning processes, research and to social bonding matters, that chatbots could potentially address, through an appropriate design. Findings from this study indicate that higher education instructors currently struggle with content delivery, formative assessment implementation and time-consuming administrative work. Likewise, students confirmed the need of having a tutor

for content support provision, a finding consistent with the work of Gupta, et al. (2019). Students also express their need for having a chatbot responding to FAQs at any time of the day (such as: Watson, reported in Goel et al., 2015; EduBot, reported in the study of Verleger & Pembridge, 2018).

In fact, solutions have previously been developed for addressing such issues to some extent. For instance, Goel and colleagues (2015) developed a chatbot, called 'Jill Watson', built on IBM's Watson platform, that functioned as a teaching assistant, to handle forum posts by students enrolled in a computer science course at the University of Georgia (Goel, et al., 2015). In this case, there was a need by the instructor to address numerous questions of students during the semester. The Jill chatbot is recorded as the world's first AI teaching assistant that was available 24/7 for responding to M.S. Computer Science students' queries. It was observed that the use of the chatbot by students enhanced their engagement in the course. Therefore, the development of the teaching assistant chatbot helped in solving this problem to a great extent. In the same direction, Dutta (2017) describes in detail the development process of an intelligent chatbot to assist high-school students for learning general knowledge subjects. The proposed intelligent web-based and free access chatbot tool makes use of NLP techniques to answer the queries by high-school students and be trained on a knowledge base consisting of general knowledge questions and answers. In relation to formative assessment solutions, Yi-Chieh and Wai-Tat (2019) explored the use of a conversational chatbot interface for guiding students to perform peer assessments. Even though their findings are promising in this area, as grading consistencies were exhibited among students' and teachers' evaluations, the use of chatbots for the provision of qualitative feedback continues to comprise an area for further investigation.

The challenge remains on building and sustaining the use of a smart chatbot, capable of delving into content-related topics of a course and offering personalized guidance and feedback to the learners. Even though chatbots can evolve via machine learning techniques and through evaluating conversations with users, the error rate at which a chatbot works is initially high (Molnár & Szüts, 2018). Also, according to the same authors, even though chatbots can simplify the administrative work of educators by disclosing supplementary information to students about their courses, they often fail to solve content issues. This becomes a bit problematic when chatbots are meant to be used for meaningful formative assessment purposes and content-related guidance provision. Therefore, the need for having chatbots that support formative assessment and tutoring in a meaningful manner and content-wise remains.

Overall, all participants in this study valued the use of chatbots in higher education, but at the same time they were aware and knowledgeable of limitations that chatbots have from a development perspective, so as to address all their needs. This remark was evident especially in the interview data. The comparison tests that we have performed on quantitative data which have been collected across all pilots, demonstrate statistically significant differences among educators and students on their perceived needs and expectations for the use of chatbots in education. Likewise, statistically significant differences have been found among educators and students from the four participating countries (Cyprus, Croatia, Spain, UK). Those differences can be attributed to their different previous experiences with chatbots, and therefore the realization of the real affordances and constraints of chatbots; to cultural factors and social norms that exist in each country; to the different educational systems and methodologies that might be used across the different higher education institutions; but

also, to students' and educators' own attributes and characteristics. The outcomes of the needs analysis that was conducted as part of the EDUBOTS project work, have implications for the design, development, and testing of specific pedagogical scenarios for the meaningful integration of chatbots in higher education.

2.3. Chatbot solutions and problems solved

An increasing number of universities and colleges has already started deploying virtual assistants or chatbots to communicate with students on all aspects of college life, creating a virtual "one-stop-shop" for student queries. Chatbot solutions have been already applied in multiple educational contexts, such as, in health and well-being interventions (Alepis & Virvou, 2011; Bickmore, Schulman, & Sidner, 2013), in medical education (Kerfoot et al., 2006), in mathematics education (Knill, Carlsson, Chi, & Lezama, 2004), and in language learning (e.g., Ayedoun, Hayashi, & Seta, 2014; Fryer, Ainley, Thompson, Gibson, & Sherlock, 2017). For instance, in medical education, chatbots have been used for offering support to students during their education, but also as medical agents for supporting patients with their therapy and answering questions related to their health situation (Alepis & Virvou, 2011). In this manner, patients learn to autonomously follow their therapy, without the immediate support of a human being (Bickmore et al., 2010). In addition, chatbots have been used in the well-being domain, for triggering healthy habits to users (e.g., prompting users to be more physically active, to consume greater amounts of vegetables and fruits) (Bickmore, Schulman, & Sidner, 2013). In language education, chatbots have been embedded in a learning-friendly environment that allows students to practice a foreign language, with reduced levels of anxiety while doing the practice (Ayedoun, Hayashi, & Seta, 2015). However, research has shown that students tend to lose interest in engaging with chatbots over time, probably due to the weaker value of chatbots compared to human assistants (Fryer, Ainley, Thompson, Gibson, & Sherlock, 2017).

Apart from their application in different disciplines and domains, educational chatbots can serve different purposes and accommodate different learning scenarios. They can facilitate students' learning, ease instructors' workload, and act as teaching assistants, offer administrative and technical support. Below, we summarize different use case scenarios in which chatbots offer aids in Higher Education.

2.3.1. Safeguarding students' successfully transition to Higher Education

Chatbots may smoothen the transition of secondary-education students into the university environment and increase university enrolment. For instance, the Pounce chatbot, a text-based chatbot, was introduced at the Georgia State University (GSU) in 2015, for safeguarding students' successfully transition to college.

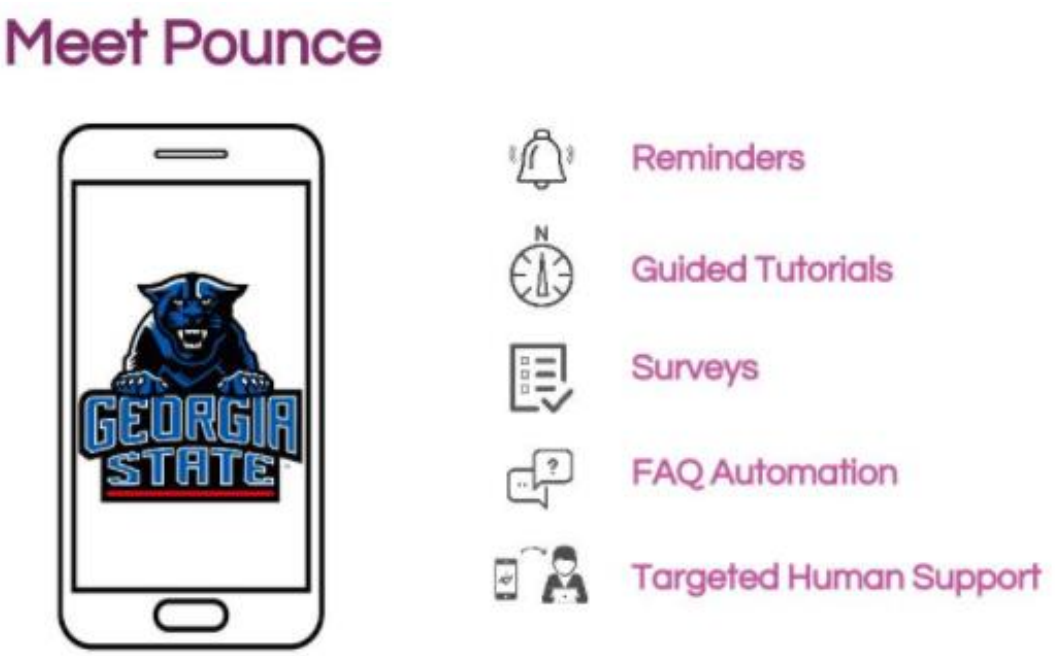


Fig. 22. The Pounce chatbot. More information: https://bit.ly/3rKslMf

Other chatbot solutions may help students choose a university according to their career aspirations and preferences.

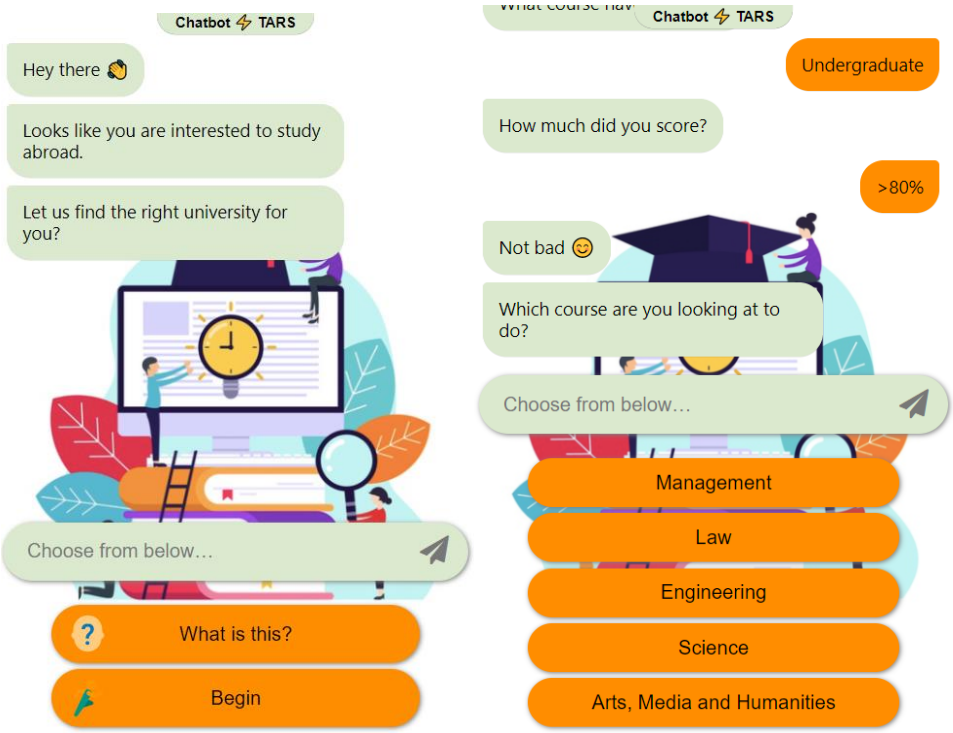


Fig. 23. Chatbot supporting students to choose a university for studies. More information: https://hellotars.com/

2.3.2. Simplify the Enrollment Process

For any college or university, swift communication is crucial when it comes to converting prospective applicants into enrolled students. A prospective student could be browsing the university's website in the middle of the night and suddenly have questions that, which if answered immediately, could determine whether they apply to the institution or not, without risking losing their interest altogether. Chatbots can also make enrollment easier. Universities and colleges deploy chatbots to answer students' FAQs, such as "Which courses are available this semester?" "How do I register?" "Who is my advisor?". These questions extend to financial aid as well. "Am I eligible?" "How do I apply?" "Why haven't I received my financial aid?" In the lead up to the start of a new academic year candidates are buzzing with questions relating to courses, fees, accommodation, and the like. Overall, chatbots can be designed and programmed appropriately and according to each institution's needs, for providing information about the university campus, facilities, scholarships and benefits, and enrollment processes. Also, chatbots can provide basic information to students about a specific course, for convincing them to enroll in the course. For instance, a chatbot solution offered by Tars (https://hellotars.com/) can introduce a new course to a student and then provide some more information about how to enroll, in case of some interest on the behalf of the student.

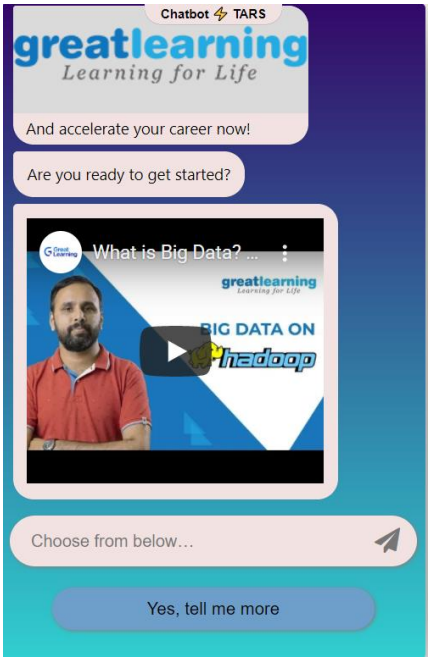


Fig. 24. A course enrollment chatbot. More information: https://hellotars.com/

2.3.3. Accessibility of website information

University websites see traffic all year round and carry a lot of different kinds of information that students may not actively search for but could benefit from. For instance, scholarships and select internship programs often take voluntary applications, which an eligible student might not submit if they are unaware of the existence of these programs. A solution to this can be a chatbot which directs students, and users in general, on what and where to search within the institution's website. In

addition, chatbots may appear in the organization's website, for initiating a conversation with the user and then making the connection to a human assistant, who can further handle the conversation and respond to students' more complex questions. An example of such a chatbot solution is illustrated in the figure below.

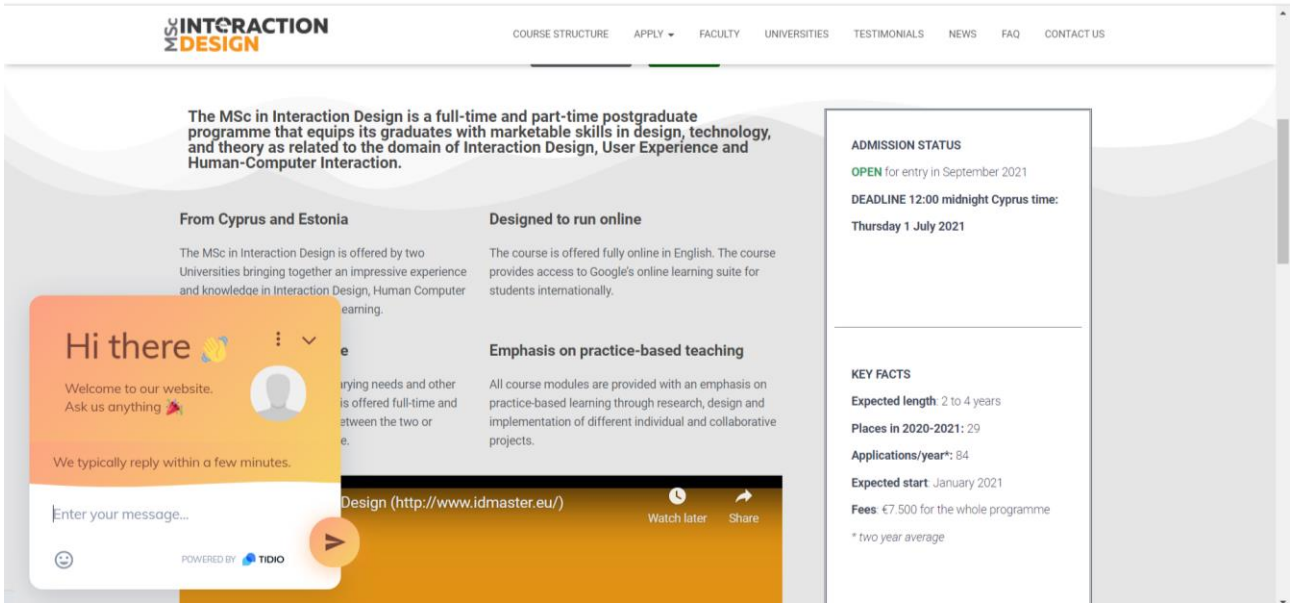


Fig. 25. An example of a chatbot availability in an academic institution's website. Source: https://www.idmaster.eu/

2.3.4. Simplification of administrative formalities

Facilitating information access and retrieval is the first step in a series of tasks that education chatbots are mandated to perform. They also have a major role to play in simplifying the administrative formalities that are part and parcel of enrolling at an educational institute. To begin with, the admissions process itself entails a significant amount of paperwork, which includes filling and submitting forms, taking online tests, furnishing certificates and referrals, and making fee payments. Through each of these processes, potential students have a variety of doubts and questions. They are not always located near the university, nor do they have the time to file their queries and complaints via emails or phone calls. University website chatbots become indispensable in these situations. Chatbots can simplify administrative formalities, including the administrative work of educators, by disclosing supplementary information to students about their courses and the university in general (Ma et al., 2019). An example of such a chatbot is BARICA (https://github.com/AILab-FOI/B.A.R.I.C.A./), which is a chatbot for students. B.A.R.I.C.A.[1] is an acronym behind the name Beautiful ARtificial Intelligence Cognitive Agent. Students can interact with BARICA and pose questions about their schedule, about their professors, about the building, where the rooms etc. BARICA is a cognitive agent, it is visualized as a virtual personality in a way, it interacts with voice recognition and she answers with a voice.

[1] https://www.foi.unizg.hr/en/news/barica-artificial-intelligence-chatbot-developed-foi-students-assistance

2.3.5. Facilitation of task management

A scheduling meeting bot service allows people to delegate the work of scheduling a meeting by cc'ing the bot in an email conversation (Cranshaw et al., 2017). Similarly, there is a myriad of bots available in work-centric chat platforms like Slack and Microsoft Teams. Technologies for team coordination of tasks (Anh, Cruzes, & Conradi, 2012; Sanchez, Jin, Maheswaran, & Szekely, 2008) are well-established, yet research work has been further expanded towards the development of bot prototypes, such as the Taskbot, which act as a mediator of task management for individuals and teams. The bot can help individuals and teams to create, assign, and keep track of tasks, all within their main communication channel (Toxtli, Monroy-Hernández, & Cranshaw, 2018).

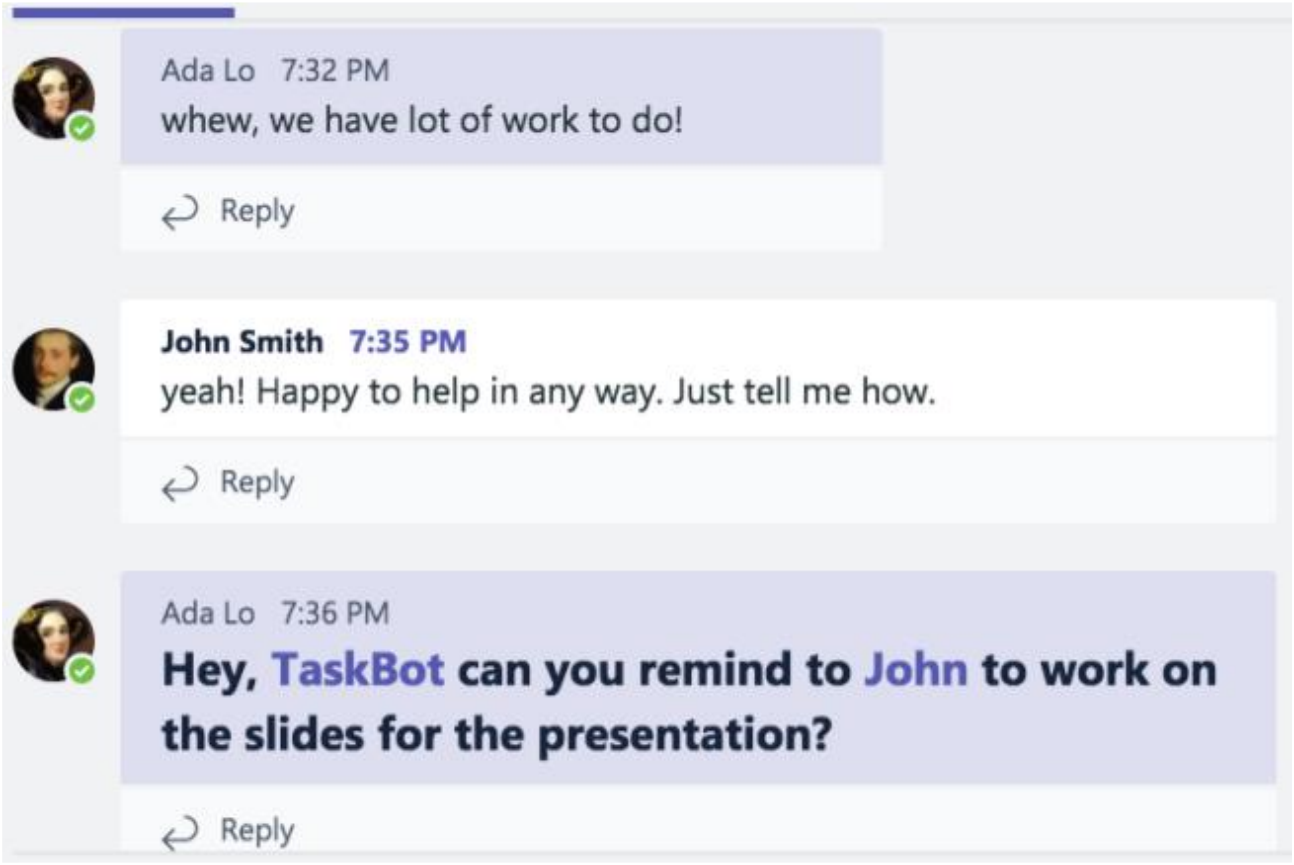


Fig. 26. A group chat conversation on Microsoft Teams showing a user assigning a task to her teammate and asking TaskBot to help keep track of it (Toxtli et al., 2018).

2.3.6. Online learning and teaching using LMS

Many universities use Learning Management Systems (LMS) for their distance learning programs, but also for conventional courses the past year, due to the covid-19 pandemic and the transition to online learning. One of those LMS is Moodle, with a wide range of users worldwide.

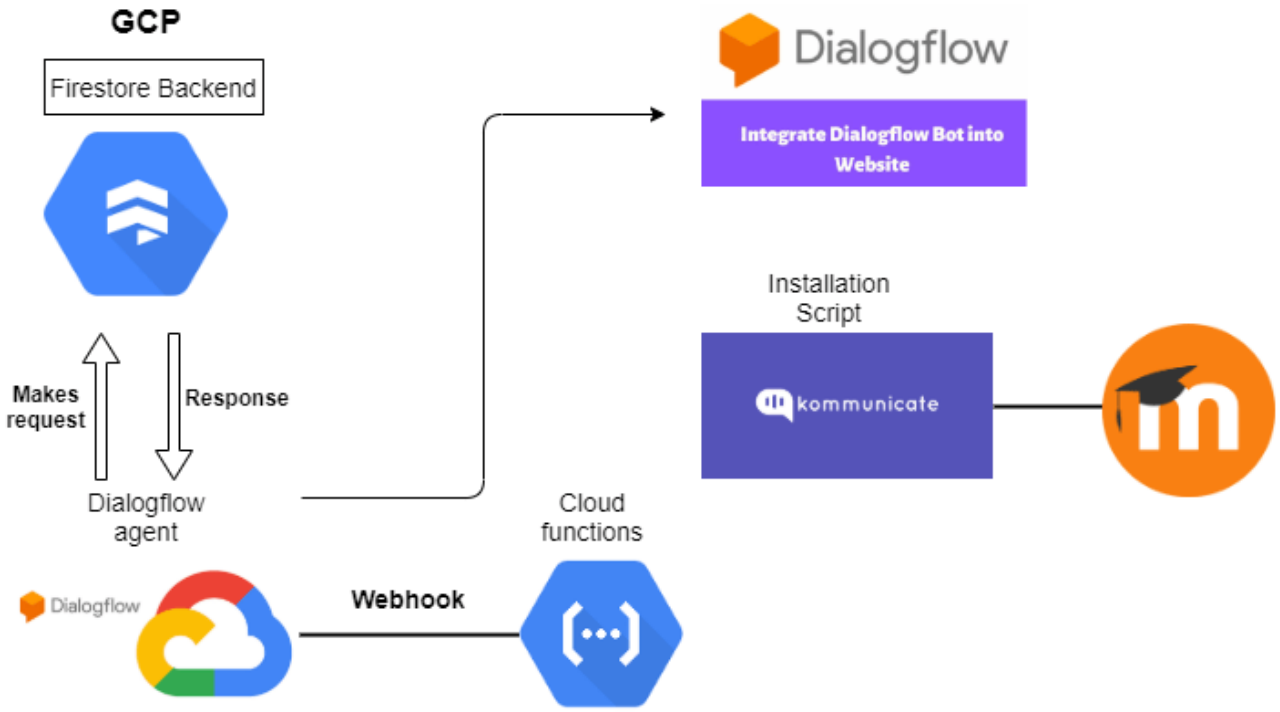


Fig. 27. Chatbot integration into Moodle: the solution architecture. Source: https://bit.ly/3skeU5n

In response to the problem of providing online information to students seeking clarifications around their course curriculum, at any time of the day, Moodle has integrated chatbot solutions to aid/assist in information triage and provide a speedy response (real time) to questions and queries of interest. The area of focus is on student needs and addresses the academic, enrollment, student life questions and clarifications. The chatbot is created using the Dialog Flow framework.

2.3.7. Social connection with peers and mentors

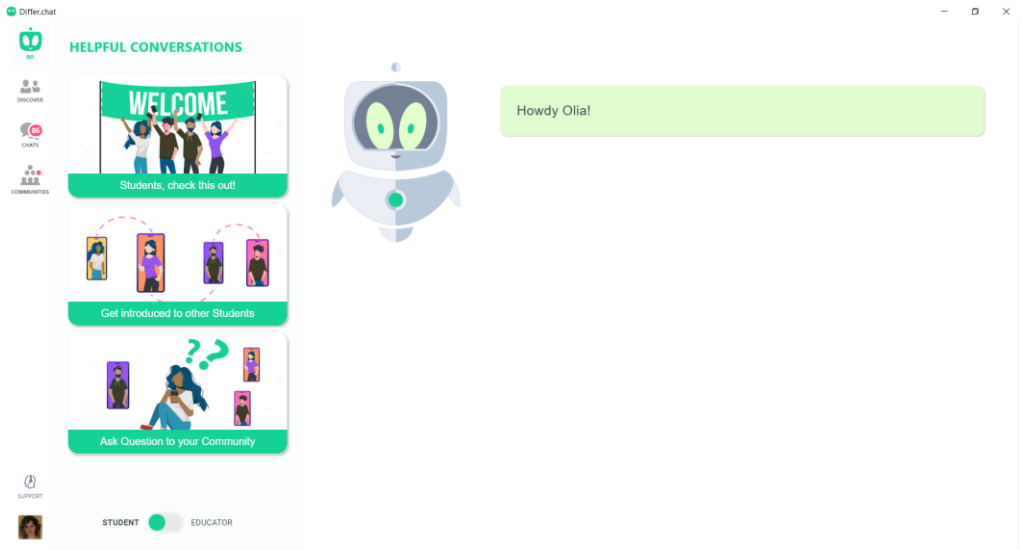


Fig. 28. Differ and BO chatbot can support peer-to-peer connections

Other chatbot solutions can help students, especially newcomers, in connecting with other students at the University, their mentors and instructors, and therefore facilitate social connection and interaction. An example of such a chatbot solution is BO, a social chatbot integrated into a chat platform, Differ (more information: https://www.differ.chat/).

2.3.8. Facilitation group chat discussions

GroupfeedBot has been developed to facilitate and improve group chat discussions, by managing the discussion time, encouraging members to participate evenly, and organizing members' opinions, in response to the difficulties faced when conducting goal-oriented discussions through a group chat. The particular challenges that GroupfeedBot aims to address involve: (i) the increased difficulty in reaching timely consensus in a group chat, in comparison to face-to-face meetings, as procrastination and loss of concentration are common in a group chat (Hill Duin & Archee, 1996); (ii) uneven participation in the group chat conversation due to the absence of in person contact during messenger-mediated interactions and subsequent weakened positive group dynamics (Frey, 2004); (iii) the increased difficulty in the exchange of diverse opinions, due to the unstructured nature of group chats.

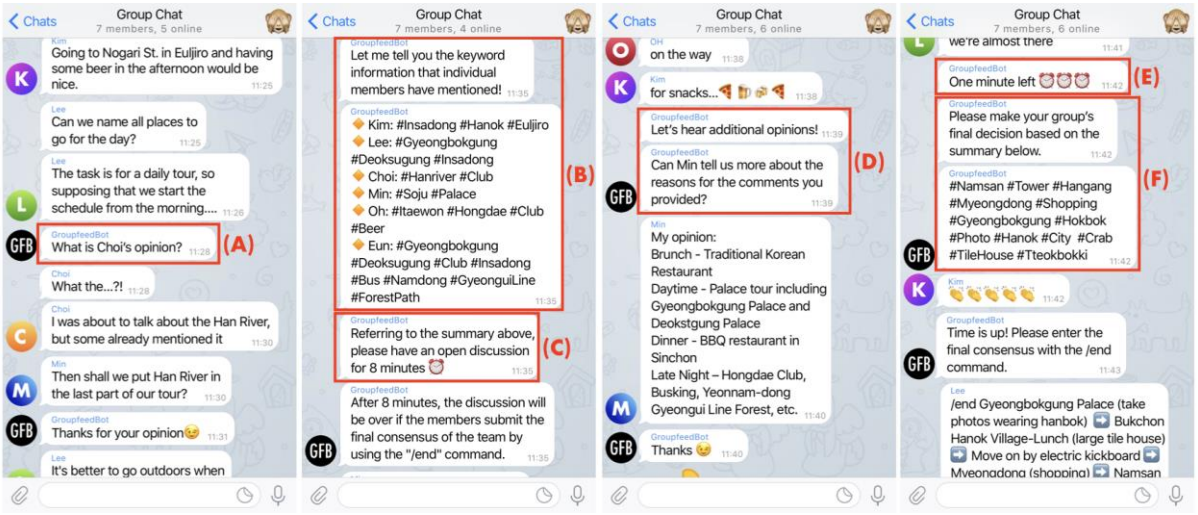


Fig. 29. Conversational design strategies applied in GroupfeedBot to facilitate group chat discussions. The chatbot manages the discussion time (1C, 1E), facilitates even participation by encouraging lurkers to speak up (1A, 1D), and organizes individual members' (1B) and overall groups' (1F) opinions (Kim et al., 2020).

2.3.9. Teaching assistants and tutors

Chatbots may also act as teaching assistants and tutors in specific courses. For instance, Goel and colleagues (2015) developed a chatbot, called 'Jill Watson', built on IBM's Watson platform, that functioned as a teaching assistant, to handle forum posts by students enrolled in a computer science course at the University of Georgia (Goel, et al., 2015). In this case, there was a need by the instructor to address numerous questions of students during the semester. The Jill chatbot is recorded as the world's first AI teaching assistant that was available 24/7 for responding to M.S. Computer

Science students' queries. Therefore, the development of the teaching assistant chatbot helped in solving this problem to a great extent; it was recorded that 10 thousand questions were received per semester (Molnár & Szüts, 2018). It was observed that the use of the chatbot by students enhanced their engagement in the course.

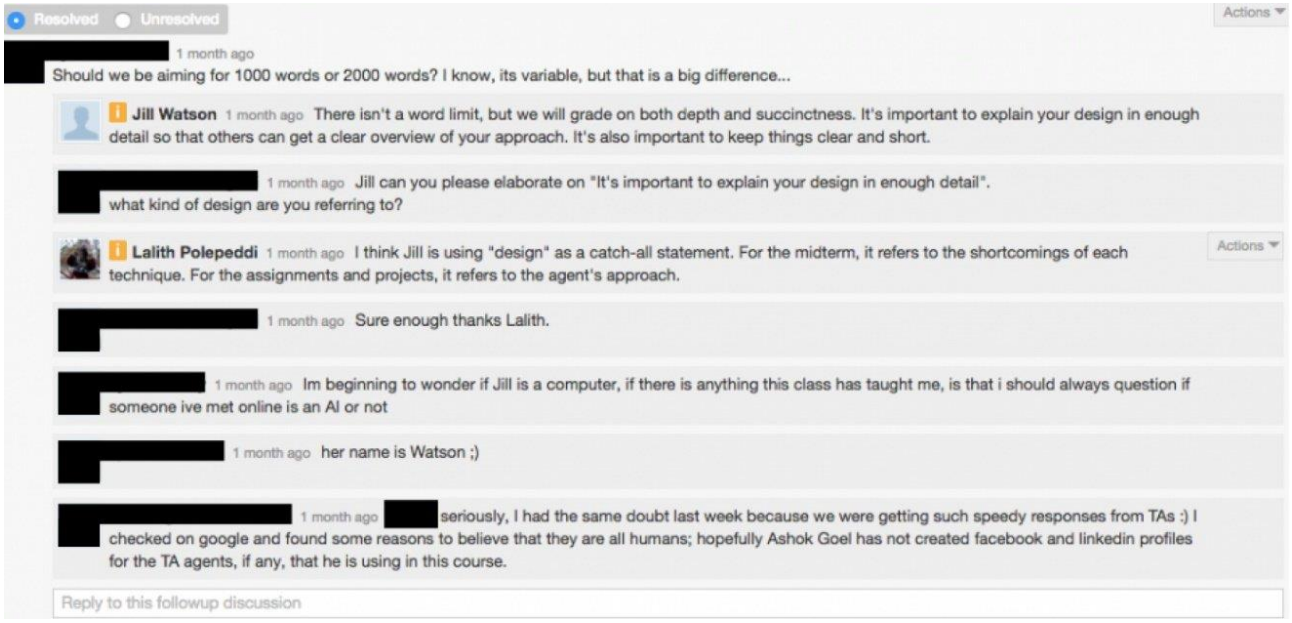


Fig. 30. The Jill Watson chatbot. Photo Credit: Gatech.edu

Dutta (2017) describes in detail the development process of an Intelligent chatbot to assist high school students for learning general knowledge subjects. The proposed intelligent web-based and free access chatbot tool makes use of NLP techniques to answer the queries by high-school students and be trained on a knowledge base consisting of general knowledge questions and answers.

Huang et al. (2017) in their paper present a chatbot for a dialogue-based computer-assisted second language learning system (Genie Tutor). Genie Tutor's operations are structured based on semantic and grammar correctness evaluations (Kwon, Lee, Kim, & Lee, 2015). Given scenarios are incorporated in the system which leads dialogues by posing questions to the users. Genie Tutor promotes dialogue with the users through two types of English learning stages, called Think & Talk and Look & Talk (see figure 31). The schematic diagram of Genie Tutor consists of automatic speech recognition and tutoring modules. Once the user-chatbot conversation is completed, feedback closing session is taking place, assessing "task proficiency", "grammar accuracy", "vocabulary diversity", and "syntactic complexity". An example of this is illustrated in figure 32.

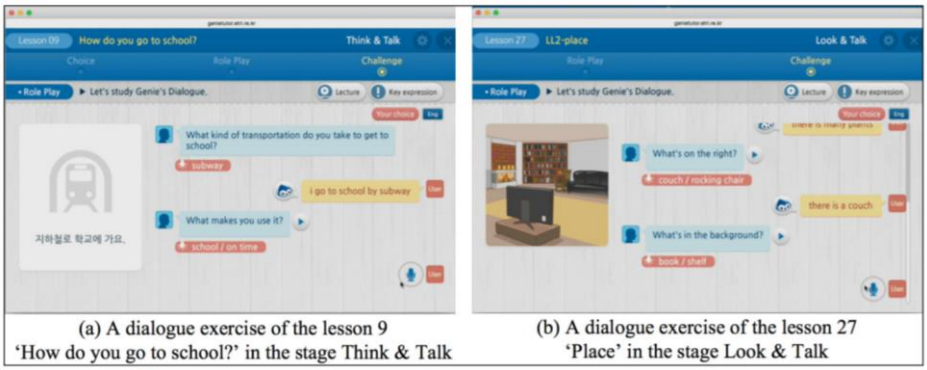


Fig. 31. Examples of dialogue exercises in Think & Talk and Look & Talk (Kwon, et al. 2015).

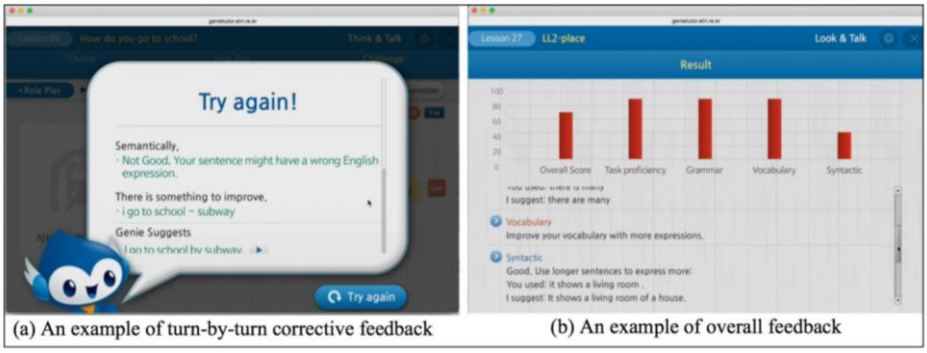


Fig. 32. The educational feedback of Genie Tutor (Kwon, et al. 2015).

Smutny and Schreiberova (2020) in their recent literature review of educational chatbots for the Facebook Messenger evaluated 47 chatbots among 2000 identified Facebook Messenger chatbots in the Botlist.co directory list (https://botlist.co/). Among those, chatbots were serving teaching purposes, by supporting the following functionalities: recommending learning content to users, providing feedback, Q&A, setting goals and monitoring learning progress (e.g., Wordsworth and EnglishWithEdwin chatbots, see figure 33).

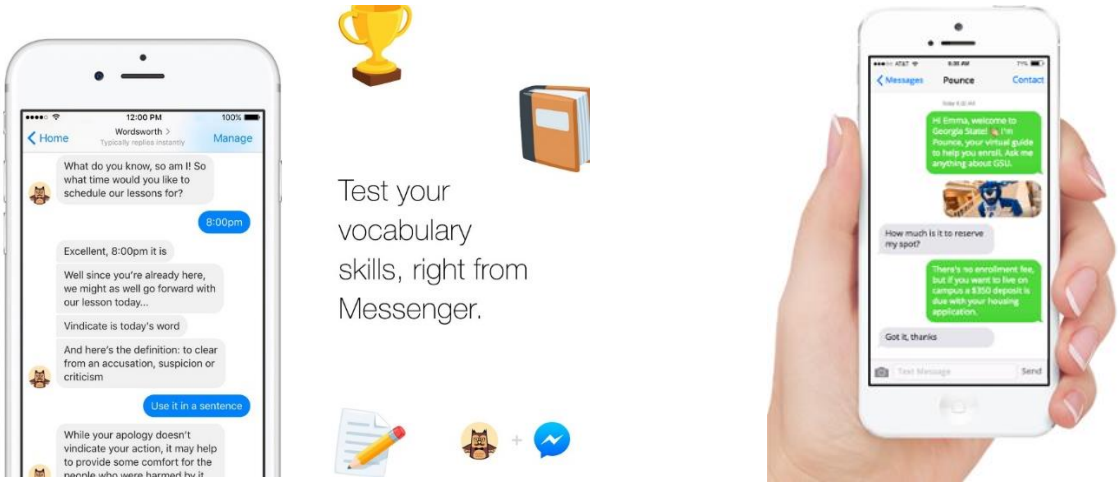


Fig.33. The EnglishWithEdwin chatbot. Photo Credit: Facebook (on the left). The Pounce chatbot. Photo Credit: AdmitHub (on the right).

2.3.10. Course evaluation

Course evaluation typically takes place at the end of the semester in most academic institutions, in the form of standardized surveys with close-ended questions. Yet, chatbots can make this procedure more interactive and informal to students. A chatbot solution in this direction is Hubert (https://hubert.ai/), which allows instructors to customize a course evaluation and choose a template with predefined questions that the chatbot will pose to students. Through short, interactive dialogues with a chatbot, the students have the opportunity to provide informal feedback to their educator on their course. This feedback can be used by the educator for improving the applied teaching practices and approaches. Instructors can use chatbots for receiving feedback from students during the semester, in an informal manner, with direct access to raw data (students' responses and interactions) but also to thematic analysis that the chatbot can do.

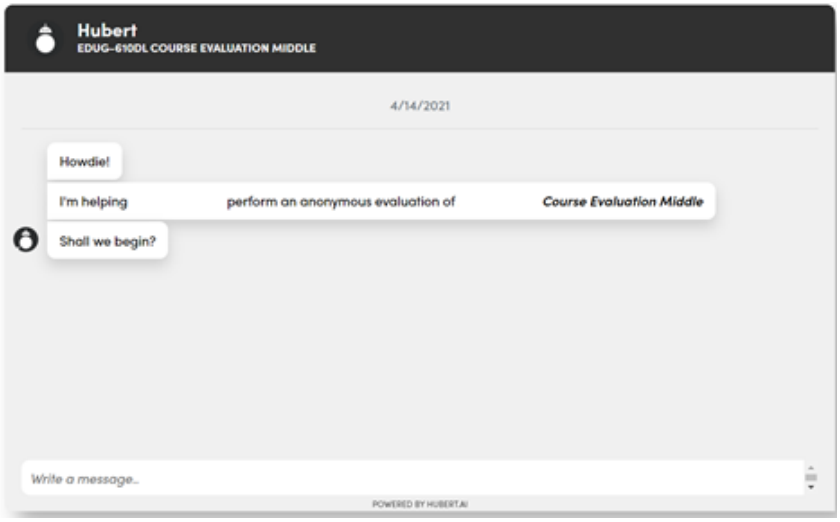
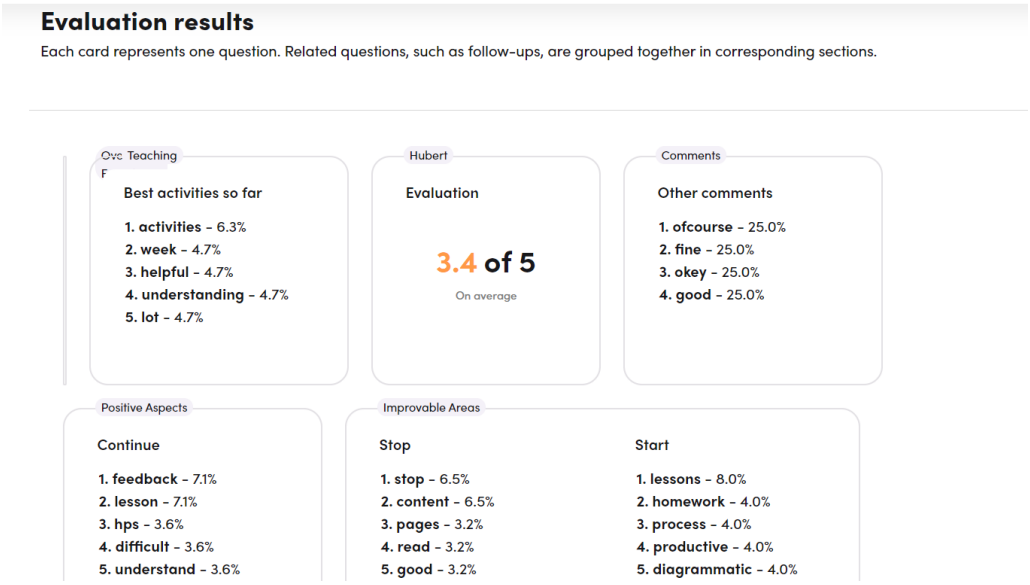


Fig.34. Hubert chatbot interacting with a student for course evaluation.



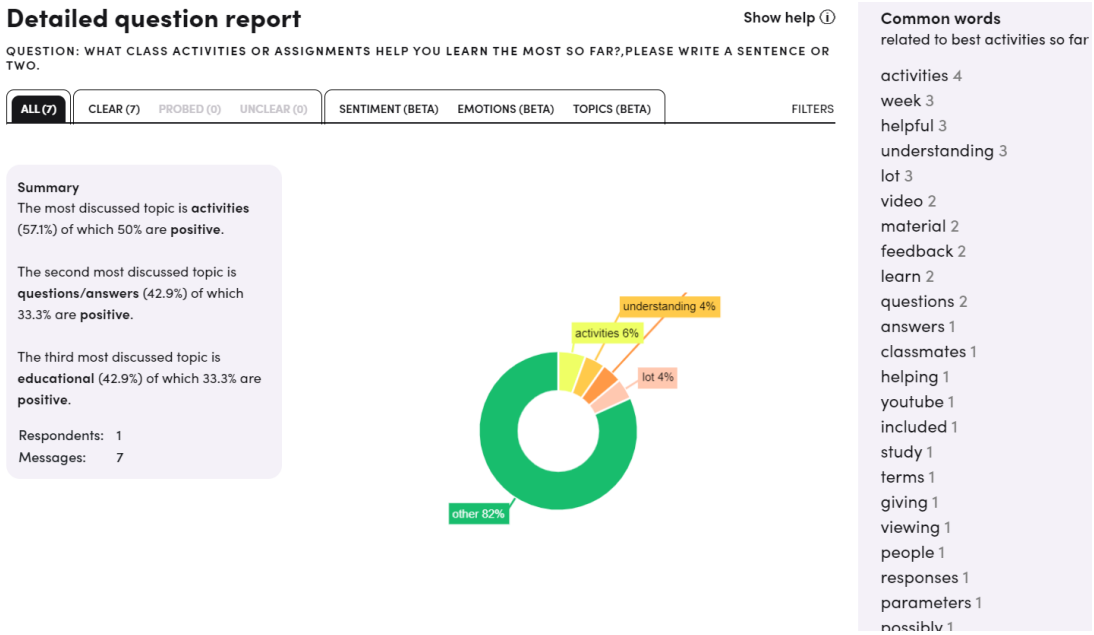


Fig.35. Access to a thematic map of students' responses via Hubert.

2.3.11. Formative assessment

Chatbots have also the potential to provide formative feedback to students on their performance and prompt metacognitive thinking. Considering that the importance of formative approaches in the assessment field has been acknowledged by several scholars, where attention is redirected from the summative evaluation of student performance toward the use of assessment during the learning procedure in order to support the improvement of learner outcomes (Bell & Cowie, 2001; Pellegrino, Chudowsky, & Glaser, 2001), it is unequivocal that chatbots can act as formative feedback tools and contribute to this direction.

2.3.12. Self-reflection and motivation

Finally, chatbots can be considered as tools for triggering metacognitive thinking, but also for increasing students' motivation in engaging in the learning process. With regards to metacognition, Kerly, Ellis, and Bull (2008) in their study, revealed that the major advantage of chatbots in comparison to asynchronous ways of communication is that students reflect their beliefs during discussion. As for motivation, considering that chatbots can pose challenging questions to students during a learning scenario, thus, fostering their curiosity and engagement towards learning (Oudeyer, Gottlieb, & Lopes, 2016).

2.4. Pro and cons

The usage of chatbots bears several advantages. First, in terms of customer service costs, the implementation of chatbots costs less compared to other conventional methods of using human assistants (for more information see section on pricing). Second, chatbots can increase user satisfaction by speeding up response times and being available at any time of the day. Third, chatbots

can interact proactively with the users and display precisely the information that the users are searching for. This proactive role can be achieved because chatbots can initiate a 'dialogue' with the user and adjust the content of the communication appropriately, considering the user's location or clickstreams, making in this way that the customer feel that is personally addressed (Howlett, 2017). Forth, and in the field of education, chatbots can be a solution to the inadequate individual support that students receive in large-scale courses and/or MOOCs, with no further financial and organizational costs for the providers. Therefore, chatbots can provide essential individual student support especially in large scale classrooms, in which the provision of individual feedback and support is demanding for educators. Moreover, the interactions among chatbots and the users can be automatically analyzed, thus providing a fourth advantage of using them. This analysis can be used to understand the users' requirements better and therefore improve the service/ product and in general, the purpose that the chatbot is serving.

The analysis of the smart interview data with experts confirms the aforementioned benefits that chatbots have. In addition to those, the cognitive offload of the users was mentioned. That is, from the users' perspective cognitive offload was explained in the sense that they do not need to memorize pieces of information when a chatbot can offer this type of information easy and quickly via a short interaction with the user. Then, from the provider's perspective, cognitive offload was explained in the sense that the chatbot can reply to commonly asked questions by the users. Indicative codes are given below:

"Typically, people that want to find some information or solve a task, ask other peers to help them solve the problem so in this context the person asking for information benefits by asking that information from a chatbot, this is a form of cognitive offloading. Instead of having to remember things from myself or having to solve a certain task in a prescribed way myself or search offloading my cognitive effort to a machine to a chatbot" (expert).

"On the other side of the conversation we have the person that we have given that information to the requester, the information provided or the task solver, and again we have a benefit in that this person does not have to repetitive request and menial tasks to solve because their job is automated by having a chatbot take on that role" (expert).

In addition, the type of interaction between chatbots and users can be characterized as a human-like and enjoyable interaction and in an educational context, it can be considered appropriate for online teaching and learning. In addition, the personalization aspect has been acknowledged as an advantage, since a sufficiently advanced chatbot should be able to accommodate the particularities of the user or the student interacting with a chatbot.

Despite all these advantages, chatbots are still at the beginning stages of their expansion especially in the field of education in which there is a great potential for their usage. Several further issues need to be confronted when chatbots are being applied in educational settings. Personalization was mentioned by the experts as also being one of the challenges since the technology is not ready yet to fully support this functionality. Related to the previous constraint, another challenge proposed by the experts, includes the limitation of a chatbot to handle complex conversations and the high error rate at the beginning of their use (which can evolve via machine learning techniques). Therefore, even though chatbots can simplify the administrative work of educators by disclosing supplementary

information to students about their courses, they often fail to solve content issues (Molnár & Szüts, 2018). This becomes a bit problematic when chatbots are meant to be used for meaningful formative assessment purposes. Let alone, when chatbots might become inefficient enough or do not comprehend the users' requests and questions, even in administrative matters, then, in this case, they could cause frustration originating from ineffective communication.

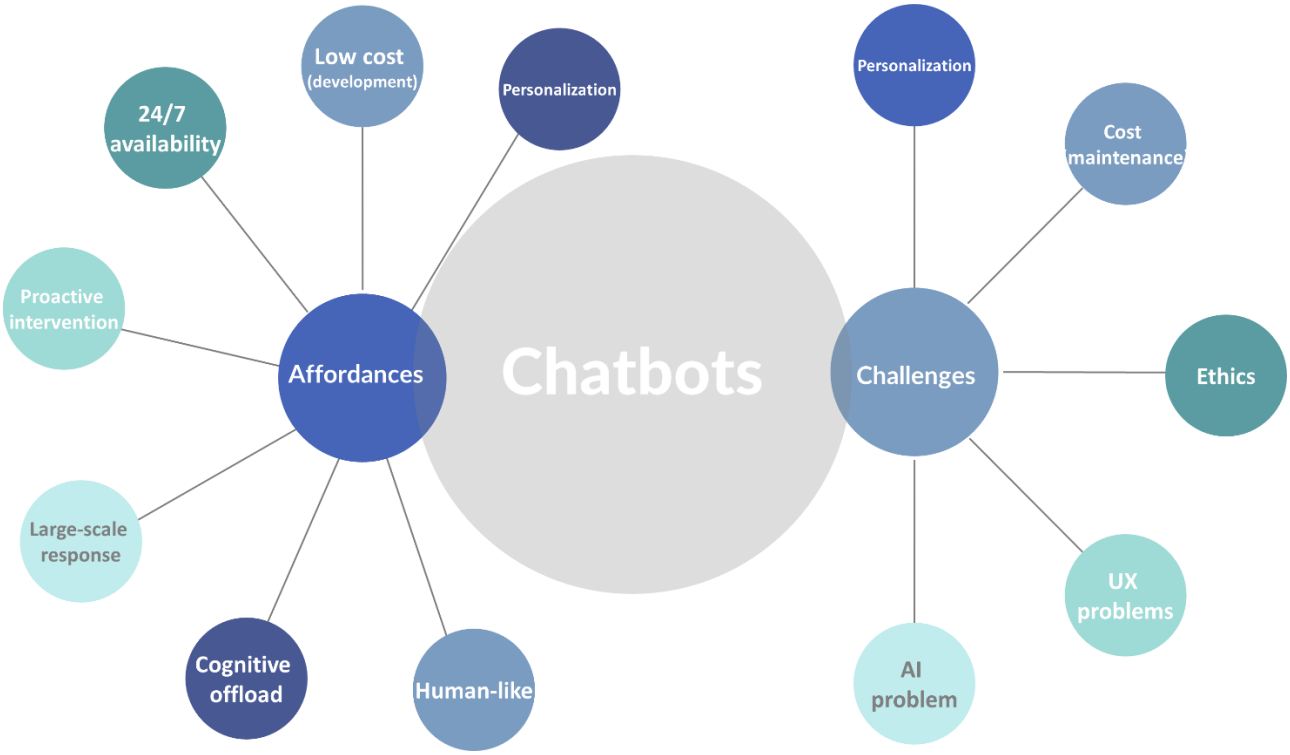


Fig. 36. Challenges and affordances of Chatbots

"I think we are not yet at the place where we have real personalization to the students and an unconstrained in a sense interaction with the chatbot. Chatbots, for the most part to my understanding, come with pre-determined skills for a rather narrow domain of discourse. That is to be expected because having unconstrained chatbots would essentially solve the AI problem and we are far from doing that" (expert).

Chatbots can evolve via machine learning techniques and through evaluating conversations with users. In this way, the more conversations a chatbot has, the more intelligent it becomes over time (Molnár & Szüts, 2018).

Shortcomings in the use of chatbots include the usual disconnect between the vision of what AI powered chatbots, and more broadly, intelligent tutoring systems could be, and what they are (Baker, 2016); this could be attributed to the approaches used in practice, which are mainly simple. Also, research innovations in the field often do not get integrated into the systems deployed at scale, that is, systems being used at scale in education are generally not representative of the full richness that research systems demonstrate. Therefore, even though there is an initial intent from researchers to develop systems that can use reinforcement learning to improve themselves (e.g., Beck, Woolf, & Beal, 2000), few systems incorporate this capacity (Baker, 2016). As Hobert and Meyer von Wolff

(2019) propose, there is a need for comprehensive, in-depth evaluation studies and corresponding process models in this direction.

Another major challenge that was revealed in the interview data, deals with ethical considerations around the application of chatbots in education. This challenge bears several dimensions. First, a chatbot-student interaction can be viewed as an educational intervention which raises ethical questions in case of undesirable effects. Second, an orthogonal ethical consideration involves data privacy issues; that is the data that is gathered from a chatbot interacting with a student and its further use for research or policy making purposes.

Finally, the experts argued that sustainability of use, high maintenance cost and potential user experience (UX) problems might cause feelings of frustration to the users were also outlined. Related to the latter point, it was explained by one of the experts that it is necessary for the users (i.e., learners) to be aware of the capabilities of the chatbots and adjust their expectations accordingly, in order to avoid feelings of frustration.

2.5. Pricing

As already mentioned earlier, one of the advantages of using chatbots in any business but also by educational organizations, is the low cost, as compared to other conventional methods of using human assistants (Winkler & Söllner, 2018). On average, a chatbot interaction with a customer saves them a cost of 0.70 $ per interaction compared to traditional support conversations (source: IBM blog: https://ibm.co/3dVTxSL). There are different schemes that providers follow in terms of pricing. For standalone chatbots, providers may offer a flexible plan, starting from a free access plan with basics functionalities, or a free trial of 14 days, to more advanced and customized chatbot solutions with specific prices for personal or institutional (business, enterprise) use. Likewise, for chatbots integrated in other platforms and/or chat tools. For example, in terms of the cost of BO, the social chatbot integrated within the Differ chat (https://www.differ.chat/), the price is integrated in the total price of the Differ solution. Differ applies a "freemium model" with a basic free solution available for individual communities of up to 200 students. Community managers can pay for a premium annual license to increase the size of the community and/or access premium features. Institutions can get a custom quote for large scale implementation projects, including a multitude of communities and students. However, as explained by the experts during the smart interviews, maintenance costs can be listed in the drawbacks of chatbots, especially when it comes to customized and adapted solutions for serving the needs of Higher Education users from different disciplines and domains and depending on the exact technology used.

Chapter 3: Chatbot-mediated learning

3.1. Introduction

Chatbots can offer solutions to many problems for Higher Education Institutions and have the potential to facilitate students' learning process. The smooth integration of chatbots into the educational context can be facilitated by the fact that nowadays students are more familiar with technology than ever before, and the use of communication tools is an integral part of their daily lives. Moreover, scholars argue that chatbots have the potential to scaffold students' learning and provide new ways in which students search for information. Following a student-centered approach with a clear focus on the promotion of students' self-regulation, reflection, autonomy, and responsibility using formative assessment and timely feedback is a challenge for educators, especially in cases of large student groups. As Winkler and Söllner (2018) suggest, in large-scale learning scenarios at universities or in massive open online courses (MOOCs), chatbots have the potential to provide individual support and feedback to students with no further financial and organizational costs for the providers.

Previous studies have shown that the effectiveness of chatbots in students' learning is complex and depends on a variety of factors (Winkler & Söllner, 2018), while measuring the effects and impact of chatbots' use in students' learning comprises a crucial challenge. Also, the process of using chatbots in education is crucial for a full understanding of how technology impacts learning (Söllner, Bitzer, Janson, & Leimeister, 2018), since learning comprises, a complex process including cognitive, affective and psychomotor dimensions and individual differences among learners (Gupta & Bostrom, 2013). For instance, if chatbots do not meet requirements such as easy use and access, then they add little value to students' learning process and there is a high possibility of not being used by students. In addition, and as stated by one of our experts, *"even with quasi-experimental or true experimental designs that we usually apply in such studies, it is hard to measure the impact of chatbots in students' learning".* Randomization in sampling and large samples are deemed essential. In addition, ethical considerations on the data collection processes and the intervention itself remain challenging issues.

3.2. Chatbot-mediated learning (CML)

Chatbot-mediated learning (CML) involves the use of chatbots for educational purposes and for enhancing and supporting the learning process and ultimately user's learning outcomes by providing an individual learning experience. CML is synchronous, self-paced, and it is supposed to address the needs of the individual user / student. For underlining the potentiality of chatbots in enhancing the learning process and therefore increasing the learning outcomes, we summarize below the main outcomes of previous studies which explored the following parameters: (i) students' personal characteristics and how do they affect the CML processes and outcomes; (ii) approaches endorsed in designing and developing chatbots; (iii) elements for assessing the CML quality and how is the latter related to the chatbot design and CML learning outcomes. The abovementioned parameters are addressed below in the sections to follow.

3.2.1. Students' personal differences affecting CML

Users' characteristics (i.e., attitudes towards technology, learning characteristics, educational background, social and technological skills, self-regulated skills) may affect that CML process, as it takes place, as well as the potential outcomes. First, the attitude and trust of students toward technology play a significant role. Students with a positive attitude towards the value of chatbots feel more content in their CML processes (Söllner et al., 2018). Second, learning characteristics, such as trait emotions (e.g., anger, anxiety, joy) and personality traits (e.g., sociability), influence chatbot-directed emotions (Harley, Carter, Papaionnou, Bouchet, Landis, Azevedo, & Karabachian, 2016). Third, educational background and technical skills influence how the students interact with the chatbots. In particular, the type of language (e.g., formality) used during the interaction with chatbots by students with a computer science background vs. students with a humanities background, differs. Also, according to Mimoun and Poncin (2015), students with better technical skills have higher possibilities for benefiting from chatbots, as compared to students with lower such skills. Lastly, students' self-regulated skills influence CML processes and the corresponding learning outcomes. Students with well-developed self-regulation have better CML learning outcomes (Söllner et al., 2018).

3.2.2 Approaches used to build and design an educational chatbot

As already mentioned in earlier sections, chatbots' architecture may vary, resulting in different types of chatbots, such as rule-based to intellectually independent chatbots powered by AI. In the first case scenario, chatbots are tree-based, using an already predefined by the developer path. In many of the studies identified in the literature, chatbots of this kind have been used in educational settings. The rationale of their function is that the user enters input, then the chatbot tries to match this input with the database and answer, respectively. This kind of chatbot requires a large dataset in order to keep the frustration level of students caused by wrong responses low (Allison, 2012).

However, AI chatbots provide flexibility in the conversations that take place between the user and the chatbot, and in this manner, they better simulate a real human-to-human interaction. AI chatbots are then divided into one-way and two-way chatbots. In the first case, the chatbot is using machine learning techniques for understanding the user's input (Dutta, 2017). On the other hand, two-way chatbots use their AI power to respond in a smart way back to the user. Instead of using a predefined response from an extensive database of potential responses, the two-way chatbots try to construct, on situ, the most accurate response. These chatbots have the capability to 'learn' over time. According to Winkler and Söllner (2018), two-way AI chatbots have not been yet identified in educational settings; one-way chatbots, though, have been identified in previous studies, whose function facilitated the process of understanding the students' intents and therefore increased CML process quality and learning outcomes.

3.2.3. Determinants of CML process quality and how do they relate to chatbot design

Communication and its content are key components of the success of the CML process. Short messages and the lack of richness of vocabulary makes it harder for chatbots to understand users'

intents, during a human-to-chatbot interaction. It has been revealed in previous studies, that when students interact with chatbots during a learning scenario, they tend to communicate for longer durations with shorter messages, the communication lacks much of the richness of vocabulary, and they exhibit more considerable profanity (Corti & Gillespie, 2016; Hill, Ford, & Farreras, 2015; Mou & Xu, 2017). In other words, their behavior is driven by the fact that they interact with a machine. At the same time, it has been observed that when users interact with chatbots, they tend to adapt their language to the language of chatbots (Hill et al., 2015), and this can potentially increase the effectiveness of chatbots. The chatbot developers need to understand the types of communication that take place between users and chatbots, in order to maximize the learning outcomes in the case of educational applications and in order to improve the CML process.

3.2.4. Chatbot design, individual differences of students and CML process quality

Several factors need to be encountered when integrating chatbots in a learning scenario, in an effort to maximize students' learning outcomes and safeguard a successful application of chatbots in the educational context. Such factors comprise the visualization of chatbots, context-awareness, and the integration of existing learning messaging tools (Winkler & Söllner, 2018).

Contradictory findings on the visualization of chatbots and its impact on the learning process have been found in previous studies. In particular, Berry, Butler and Rosis (2005) in their work, claim that chatbots should be embodied, show emotions, and have a personality. On the contrary, Mimoun and Poncin (2015) in their study showed that an embodiment of chatbots is not required since this factor does not affect the levels of attention paid by students during the learning process, as emerged from an eye-tracking methodology employed in their research design. Also, according to Brahnam and Angeli (2012), chatbots' gender does not affect chatbot measures such as usage, spoken topics.

Regarding context-awareness, the chatbot design may include contextual information in order to maximize cognitive and affective learning outcomes. In order to achieve that, chatbots should be smart enough to consider the cognitive and emotional situation of learners in order to assist them in optimizing their learning goals. For instance, the potential ability of chatbots to identify students' misconceptions and/or knowledge gaps can contribute towards the enhancement of learning outcomes (Pérez et al., 2016). Likewise, the potential ability of chatbots in identifying students' emotional situation and the exhibition of relational behavior, such as caring and empathy, can enrich students' affective learning outcomes (Foster & Oberlander, 2010).

In any case, chatbots can only be effective if they are easily accessible to students and if they respond fast to students' input. It is suggested that chatbots should be integrated in existing learning and messaging systems, which students already use for implementing their learning tasks (i.e., learning management systems) and for communication with each other (e.g., Facebook, WhatsApp) (Pereira, 2016).

Concluding remarks

In this report we tried to document the current state-of-the-art of chatbots with application in higher education. The report reflects insights from the literature, complemented with findings from smart interviews which were conducted with experts in the domain of AI and educational chatbots. Insights into the history of chatbots, their evolution and their current state in terms of technology and applications are presented. Focusing on the educational context, significant attempts towards integrating chatbots in education have been identified. In addition, the report reflects key findings of the EDUBOTS Erasmus+ project, in relation to needs analysis and HE users' expectations in terms of chatbots' use in HE. Last, good practices for the use of two chatbots in HE (BO as provided in Differ and Hubert) are presented, as implemented with real users in educational settings.

References

Alepis, E., & Virvou, M. (2011). Automatic generation of emotions in tutoring agents for affective e-learning in medical education. *Expert Systems with Applications*, *38*(8), 9840-9847.

Allison, D. (2012). Chatbots in the library: is it time?. *Library Hi Tech*, *30*(1), 95-107.

Anh, N. D., Cruzes, D. S., & Conradi, R. (2012, September). Dispersion, coordination and performance in global software teams: A systematic review. In *Proceedings of the 2012 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement* (pp. 129-138). IEEE.

Ayedoun, E., Hayashi, Y., & Seta, K. (2015). A conversational agent to encourage willingness to communicate in the context of English as a foreign language. *Procedia Computer Science*, *60*, 1433-1442.

Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, *26*(2), 600-614.

Beck, J., Woolf, B. P., & Beal, C. R. (2000). ADVISOR: A machine learning architecture for intelligent tutor construction. *AAAI/IAAI*, *2000*(552-557), 1-2.

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education, 85,* 536–553.

Bernardini, A. A., Sônego, A. A., & Pozzebon, E. (2018, November). Chatbots: An Analysis of the State of Art of Literature. In *Workshop on Advanced Virtual Environments and Education* (Vol. 1, No. 1, pp. 1-6).

Berry, D. C., Butler, L. T., & De Rosis, F. (2005). Evaluating a realistic agent in an advice-giving task. *International Journal of Human-Computer Studies*, *63*(3), 304-327.

Bickmore, T. W., Mitchell, S. E., Jack, B. W., Paasche-Orlow, M. K., Pfeifer, L. M., & O'Donnell, J. (2010). Response to a relational agent by hospital patients with depressive symptoms. *Interacting with computers*, *22*(4), 289-298.

Bickmore, T. W., Schulman, D., & Sidner, C. (2013). Automated interventions for multiple health behaviors using conversational agents. *Patient education and counseling*, *92*(2), 142-148.

Brahnam, S., & De Angeli, A. (2012). Gender affordances of conversational agents. *Interacting with Computers*, *24*(3), 139-153.

Corti, K., & Gillespie, A. (2016). Co-constructing intersubjectivity with artificial conversational agents: people are more likely to initiate repairs of misunderstandings with agents represented as human. *Computers in Human Behavior*, *58*, 431-442.

Cranshaw, J., Elwany, E., Newman, T., Kocielnik, R., Yu, B., Soni, S., ... & Monroy-Hernández, A. (2017, May). Calendar. help: Designing a workflow-based scheduling agent with humans in the loop. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 2382-2393).

Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, *22*(5), 811-817.

Di Blas, N., Lodi, L., Paolini, P., Pernici, B., Raspa, N., Rooein, D., & Renzi, F. (2019, June). Sustainable Chatbots supporting Learning. In *EdMedia+ Innovate Learning* (pp. 1358-1363). Association for the Advancement of Computing in Education (AACE).

Dutta, D. (2017). Developing an Intelligent Chat-bot Tool to assist high school students for learning general knowledge subjects. Georgia Institute of Technology.

Foster, M. E., & Oberlander, J. (2010). User preferences can drive facial expressions: evaluating an embodied conversational agent in a recommender dialogue system. *User modeling and user-adapted interaction*, *20*(4), 341-381.

Frey, L. R. (2004). The symbolic-interpretive perspective on group dynamics. *Small group research*, *35*(3), 277-306.

Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, *75*, 461-468.

Gehl, R. W. (2014). Teaching to the Turing test with Cleverbot. *Transformations: The Journal of Inclusive Scholarship and Pedagogy*, *24*(1-2), 56-66.

Goel, A., Creeden, B., Kumble, M., Salunke, S., Shetty, A., & Wiltgen, B. (2015, September). Using watson for enhancing human-computer co-creativity. In *2015 AAAI Fall Symposium Series*.

Gonda, D. E., Luo, J., Wong, Y. L., & Lei, C. U. (2018, December). Evaluation of Developing Educational Chatbots Based on the Seven Principles for Good Teaching. In *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)* (pp. 446-453). IEEE.

Gupta, A., Zhang, P., Lalwani, G., & Diab, M. (2019). CASA-NLU: Context-Aware self-attentive natural language understanding for task-oriented chatbots. *arXiv preprint arXiv:1909.08705*.

Gupta, S., & Bostrom, R. (2013). Research note—An investigation of the appropriation of technology-mediated training methods incorporating enactive and collaborative learning. *Information Systems Research*, *24*(2), 454-469.

Harley, J. M., Carter, C. K., Papaionnou, N., Bouchet, F., Landis, R. S., Azevedo, R., & Karabachian, L. (2016). Examining the predictive relationship between personality and emotion traits and students' agent-directed emotions: towards emotionally-adaptive agent-based learning environments. *User Modeling and User-Adapted Interaction*, *26*(2-3), 177-219.

Hill Duin, A., & Archee, R. (1996). Collaboration via e-mail and internet relay chat: Understanding time and technology. *Technical Communication*, *43*(4), 402-412.

Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior*, *49*, 245-250.

Hix, D., & Hartson, H. R. (1993). *Developing user interfaces: ensuring usability through product & process*. New York, NY, USA: John Wiley & Sons, Inc.

Hobert, S., & Meyer von Wolff, R. (2019). Say Hello to Your New Automated Tutor–A Structured Literature Review on Pedagogical Conversational Agents.

Howlett, N. (2017). How machine learning is developing to get more insight from complex voice-of-customer data. *Applied Marketing Analytics*, *3*(3), 250-254.

Huang, J. X., Lee, K. S., Kwon, O. W., & Kim, Y. K. (2017). A chatbot for a dialogue-based second language learning system. *CALL in a climate of change: adapting to turbulent global conditions*, 151.

Kane, D. A. (2016). The role of chatbots in teaching and learning. E-Learning and the Academic Library: Essays on Innovative Initiatives, 131.

Keller, C., & Cernerud, L. (2002). Students' perceptions of e-learning in university education. *Journal of Educational Media*, *27*(1-2), 55-67.

Kerfoot, B. P., Baker, H., Jackson, T. L., Hulbert, W. C., Federman, D. D., Oates, R. D., & DeWolf, W. C. (2006). A multi-institutional randomized controlled trial of adjuvant Web-based teaching to medical students. *Academic Medicine*, *81*(3), 224-230.

Kerly, A., Ellis, R., & Bull, S. (2007, December). CALMsystem: a conversational agent for learner modelling. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence* (pp. 89-102). Springer, London.

Khan, R., & Das, A. (2017). Build better chatbots: a complete guide to getting started with chatbots. Apress.

Khan, R., & Das, A. (2018). Introduction to chatbots. In *Build Better Chatbots* (pp. 1-11). Apress, Berkeley, CA.

Kim, S., Eun, J., Oh, C., Suh, B., & Lee, J. (2020, April). Bot in the bunch: Facilitating group chat discussion by improving efficiency and participation with a chatbot. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

Knill, O., Carlsson, J., Chi, A., & Lezama, M. (2004). An artificial intelligence experiment in college math education. *Preprint available at http://www. math. harvard. edu/knill/preprints/sofia. pdf*.

Kwon, O. W., Lee, K., Kim, Y. K., & Lee, Y. (2015, August). GenieTutor: a computer assisted second-language learning system based on semantic and grammar correctness evaluations. In *Critical CALL–proceedings of the 2015 EUROCALL conference, Padova, Italy* (pp. 330-335). Research-publishing. net.

Ma, W., Cui, Y., Shao, N., He, S., Zhang, W. N., Liu, T., ... & Hu, G. (2019). TripleNet: Triple attention network for multi-turn response selection in retrieval-based chatbots. *arXiv preprint arXiv:1909.10666.*

Macaulay L., (1995). *Human-Computer Interaction for Software Designers*. London: International Thom-son Publishing.

Mimoun, M. S. B., & Poncin, I. (2015). A valued agent: How ECAs affect website customers' satisfaction and behaviors. *Journal of Retailing and Consumer Services*, *26*, 70-82.

Molnár, G., & Szüts, Z. (2018, September). The Role of Chatbots in Formal Education. In *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)* (pp. 000197-000202). IEEE.

Mou, Y., & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, *72*, 432-440.

Olson, C. Just say it: The future of search is voice and personal digital assistants; https://www.campaignlive.co.uk/article/just-say-it-future-search-voice-personal-digitalassistants/ 1392459, Assessed 04 December, 2019.

Oudeyer, P. Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. In *Progress in brain research* (Vol. 229, pp. 257-284). Elsevier.

Patton, M. Q. (2015). Qualitative research and methods: Integrating theory and practice. Thousand Oaks, CA: SAGE Publications.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academic Press.

Pereira, J. (2016, November). Leveraging chatbots to improve self-guided learning through conversational quizzes. In *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 911-918). ACM.

Pérez, J., Cerezo, E., Serón, F. J., & Rodríguez, L. F. (2016). A cognitive-affective architecture for ECAs. *Biologically Inspired Cognitive Architectures*, *18*, 33-40.

Popovici, A., & Mironov, C. (2015). Students' perception on using eLearning technologies. *Procedia-Social and Behavioral Sciences*, *180*, 1514-1519.

Sanchez, R., Jin, J., Maheswaran, R. T., & Szekely, P. (2008, January). Interfaces for team coordination. In *Proceedings of the 13th international conference on Intelligent user interfaces* (pp. 427-428).

Serban, I. V., Sankar, C., Germain, M., Zhang, S., Lin, Z., Subramanian, S., ... & Rajeshwar, S. (2017). A deep reinforcement learning chatbot. *arXiv preprint arXiv:1709.02349*.

Shum, H. Y., He, X. D., & Li, D. (2018). From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, *19*(1), 10-26.

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, 103862.

Söllner, M., Bitzer, P., Janson, A., & Leimeister, J. M. (2018). Process is king: Evaluating the performance of technology-mediated learning in vocational software training. *Journal of Information Technology*, *33*(3), 233-253.

Stathakarou, N., Nifakos, S., Karlgren, K., Konstantinidis, S. T., Bamidis, P. D., Pattichis, C. S., & Davoody, N. (2020). Students' Perceptions on Chatbots' Potential and Design Characteristics in Healthcare Education. In *The Importance of Health Informatics in Public Health during a Pandemic* (pp. 209-212). IOS Press.

Statista. (2017). Chatbot Market size 2015-2024; https://www.statista.com/statistics/656596/worldwidechatbot-market/

Thies, I. M., Menon, N., Magapu, S., Subramony, M., & O'neill, J. (2017, September). How do you want your chatbot? An exploratory Wizard-of-Oz study with young, urban Indians. In *IFIP Conference on Human-Computer Interaction* (pp. 441-459). Springer, Cham.

Toxtli, C., Monroy-Hernández, A., & Cranshaw, J. (2018, April). Understanding chatbot-mediated task management. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1-6).

Turing, A. M. (2009). Computing machinery and intelligence. In *Parsing the Turing Test* (pp. 23-65). Springer, Dordrecht.

Xu, Z., Wijekumar, K., Ramirez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K-12 students' reading comprehension: A meta-analysis. *British Journal of Educational Technology*.

Weizenbaum, J. (1976). Computer power and human reason: From judgment to calculation.

Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis.

Yan, M., Castro, P., Cheng, P., & Ishakian, V. (2016, December). Building a chatbot with serverless computing. In *Proceedings of the 1st International Workshop on Mashups of Things and APIs* (pp. 1-4).

Yi-Chieh, L. & Wai-Tat, F. (2019). Supporting peer assessment in education with conversational agents. In *Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion (IUI '19).* ACM, New York, NY, USA, 7-8. DOI=10.1145/3308557.3308695

ZEMČÍK, M. T. (2019). A Brief History of Chatbots. *DEStech Transactions on Computer Science and Engineering*, (aicae).

Annex I

Interview protocol with expert members

General Questions

1. Which are the key milestones in the history of chatbots?
2. What are the challenges and what are the benefits of using chatbots?
    a) In general
    b) In education contexts

Types of chatbots and technologies used

We understand that chatbots appear in different types depending on their various parameters, for example

- **the medium through which they can accessed** (i.e., web-based applications that run on a remote serve and can be accessed through a web page, Vs stand-alone applications)
- **the type of input** (i.e., text vs. speech type of input for the conversation initiation)
- **the building approaches applied** (i.e., retrieval-based models and generative models). [Retrieval-based models use a pool of predefined responses (in the form of FAQs) and an algorithm to pick an appropriate response based on the input and/or context; whereas, generative models generate responses out of the input with the help of machine learning techniques.]

1. Can you please comment briefly on the validity of the above?
2. Are there other parameters that we should encounter when trying to identify different types of chatbots, that we are missing above?
3. What are the constraints and what are the affordances of those?
4. Relying on their architecture and structure, chatbots can be characterized as rule-based chatbots, intellectually independent chatbots and AI-powered chatbots (one-way and two-way).
5. Considering that I am a person with no background in CS, what is the key distinction among those chatbots?
6. What type of programming is required for setting up a chatbot?

Chatbots in education

In this project we focus on the application of chatbots in Higher Education.

1. What is the added value for chatbots in higher education?
2. What is the state of the art on pedagogical chatbots in higher education now?
    a) In terms of available technologies
    b) In terms of instructional designs integrating chatbots
    c) In terms of areas of application and learning scenarios
3. How could chatbots facilitate the implementation of formative assessment in higher education?

4. How could chatbots facilitate students' learning in higher education?
5. What is the potential of chatbots in offering personalised feedback to learners? What types of features should chatbots have, serving this purpose?
6. How could chatbots facilitate social bonding, peer interaction and communication in higher education? What types of features should chatbots have, serving this purpose?
   In the literature, we came across the term '*Chatbot-mediated learning'*, as per Winkler and Söllner (2018). According to the authors, CML involves the use of chatbots for educational purposes and for enhancing and supporting the learning process and ultimately user's learning outcomes by providing an individual learning experience.
7. How do students' individual characteristics affect the CML process and outcomes?
8. How can we assess CML quality?
9. How does chatbot design affect CML outcomes?
10. Any other comments?

Thank you for your input.

Annex II

Pilot 1 - Interview protocol with students and educators

The interview is taking place as part of the work of the EDUBOTS - Best practices of pedagogical chatbots in higher education - Erasmus Plus Project, funded by the EC. First, you will be asked for your demographics. Next, questions will be posed about your previous experience with chatbots, your expectations of a potential use of a chatbot, your needs for formative assessment and finally, your experience of using Differ and / or Hubert. The aggregated data from the interviews will be used anonymously in any publication concerning research results. It will not be used in any manner which would allow identification of your individual responses. You have the right to withdraw your consent at any time and that your data will be removed. Do you give us your consent to move on?

Section A: Demographic data

1. Please specify your gender. Male, Female, Other
2. In which university do you study? / In which university are you employed?
3. In which department do you study? / In which department are you employed?
4. In which year of your studies are you? /What is your position in your institution?

Section B: Previous experience with chatbots

A chatbot is a piece of software that conducts a conversation via auditory or textual methods with the user. It is often also called, virtual agent, conversational agent, bot.

Have you ever interacted with a chatbot before? If not, we move on to Section C.

1. In which case instances did you have the opportunity to interact with a chatbot? Please choose what applies to you:
    a. In education
    b. In business
    c. In customer service
    d. In healthcare
    e. In social media
    f. Other, please specify
2. A chatbot is a piece of software that conducts a conversation via auditory or textual methods. Does this definition meet your previous experience?
3. How did you use the chatbot exactly?
4. How was your overall experience? Positive, negative? Why?
5. What did you find positive about interacting with the chatbot(s), in this previous experience?
6. What did you find negative about interacting with the chatbot(s), in this previous experience?
7. After this experience did you want to use a chatbot again?

8. With reference to this previous experience, what would you change in the chatbot(s), so that it would better address your needs?

Section C: Your expectations on a potential chatbot use

A chatbot is a piece of software that conducts a conversation via auditory or textual methods with the user. It is often also called, virtual agent, conversational agent, bot.

1. In which aspects of your daily life would a possible interaction with a chatbot make sense to you?
   a. Why?
   b. Which specific features would you like a chatbot to have, in order to serve your needs?
2. In your opinion, could a chatbot facilitate your learning/teaching progress?
   a. How? Please explain your reasoning.
   b. Can you describe a scenario where the chatbot serves your needs in this context?
3. In your opinion, could a chatbot facilitate your social life (e.g., social bonding with peers)?
   a. How? Please explain your reasoning.
   b. Can you describe a scenario where the chatbot serves your needs in this context?
4. Can you describe a scenario where a chatbot supports you in anything else that you might think of?

Section D: Your needs in terms of formative assessment and evaluation

"Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited." (Black, & Wiliam, 2009, p. 9).

1. Do you implement formative assessment in your university courses?
   a. If yes, which methods and means do you use for formative assessment purposes?
   b. If no, what are the reasons?
2. Do you receive personalized feedback from your instructor(s) and/or peers? / Do you provide personalized feedback to your students?
   a. If yes, how? With what means? Does personalized feedback support your learning process? / Does personalized feedback support your students' learning process?
   b. If no, why?
3. What kind of technology tools do you use in your courses for:
   a. Communication purposes?
   b. Assessment purposes?
   c. Peer interaction purposes?

4. If you were provided with a technology tool that facilitates formative assessment and the provision of personalized feedback, what specific functions should this tool support? What specific features should this tool have?

Annex III

Pilot 2 – Interview protocol with students and educators

The interview is taking place as part of the work of the EDUBOTS – Best practices of pedagogical chatbots in higher education – Erasmus Plus Project, funded by the EC. First, you will be asked for your demographics. Next, questions will be posed about your previous experience with chatbots, your expectations of a potential use of a chatbot, your needs for formative assessment and finally, your experience of using Differ and/or Hubert. The aggregated data from the interviews will be used anonymously in any publication concerning research results. It will not be used in any manner which would allow identification of your individual responses. You have the right to withdraw your consent at any time and that your data will be removed. Do you give us your consent to move on?

Section A: Demographic data

5. Please specify your gender. Male, Female, Other
6. In which university do you study? / In which university are you employed?
7. In which department do you study? / In which department are you employed?
8. In which year of your studies are you? /What is your position in your institution?

Section B: Previous experience with chatbots

A chatbot is a piece of software that conducts a conversation via auditory or textual methods with the user. It is often also called, virtual agent, conversational agent, bot.

Have you ever interacted with a chatbot before? If not, we move on to Section C.

9. In which case instances did you have the opportunity to interact with a chatbot? Please choose what applies to you:
    a. In education
    b. In business
    c. In customer service
    d. In healthcare
    e. In social media
    f. Other, please specify
10. A chatbot is a piece of software that conducts a conversation via auditory or textual methods. Does this definition meet your previous experience?
11. How did you use the chatbot exactly?
12. How was your overall experience? Positive, negative? Why?
13. What did you find positive about interacting with the chatbot(s), in this previous experience?
14. What did you find negative about interacting with the chatbot(s), in this previous experience?
15. After this experience did you want to use a chatbot again?

16. With reference to this previous experience, what would you change in the chatbot(s), so that it would better address your needs?

Section C: Your expectations on a potential chatbot use

A chatbot is a piece of software that conducts a conversation via auditory or textual methods with the user. It is often also called, virtual agent, conversational agent, bot.

1. In which aspects of your daily life would a possible interaction with a chatbot make sense to you? Please explain your reasoning and specific features would you like a chatbot to have.
2. In your opinion, how can a chatbot facilitate your learning/teaching progress? Please explain your reasoning.
3. In your opinion, could a chatbot establish and sustain an informal online community among students? Please explain your reasoning.
4. In your opinion, could a chatbot establish and sustain a course community? Please explain your reasoning.
5. Can you describe a scenario where a chatbot supports you in anything else that you might think of?

Annex IV

Educators' survey items

https://forms.gle/TWufftkrAbo3TWZv7

Dear educators,

There is a growing interest in the use of chatbots in educational settings as they can provide efficient and timely services to students and to lecturers. Chatbots are software programs that communicate with users through natural language interaction interfaces. They conduct a conversation via auditory or textual methods with the user. They are often also called virtual agent or assistant, conversational agent, bot (short for robot).

This survey is taking place as part of EDUBOTS - Best practices of pedagogical chatbots in higher education - Erasmus Plus Project, funded by the EC, https://www.edubots.eu/. Your feedback is very important to us for expanding knowledge about the use of chatbots in education.

Consent
The survey is anonymous and the aggregate data will be used anonymously in any publication/report related to the survey results. By clicking "next" you give your consent to participate in the survey.

Section A - Uses of Chat in Classes

If you use chat tools in your classes, how is it used?

- Yes, students can chat with everyone
- Yes, students can chat to talk to the lecturer
- Yes, but private chat between participants is disabled
- No, I disable chat in my classes

If you had an automated assistant chatbot in the chat, what would you like it to do?

- Highlight aspects of the chat to the tutor
- Promote discussion
- Social interaction

Section B - Use of chatbots in Higher Education

Below you will find potential uses of chatbots in Higher Education. How do you personally evaluate the importance of the following chatbot uses? (1: not important at all, 5: very important)

I can see a great value of using a chatbot:		1 Not important	2	3	4	5 Very important
As a digital assistant enabling feedback loops	by offering personalized feedback to my students on their conceptual understanding of a topic.					
	by conducting short quizzes with my students on their conceptual understanding of a topic.					
	allowing my students to offer to me feedback on my teaching methods for course evaluation.					
As a remote tutor for	offering tutorials to my students related to my courses.					
	sharing and filtering with my students relevant to the course resources.					
	responding to frequently asked questions (FAQs) of my students that relate to administration (e.g., deadlines).					
	responding to FAQs of my students that relate to the course content.					
	addressing my students' content-related questions on my course.					
As a social bonding facilitator for	introducing my students to their peers (e.g., with the use of ice breaking activities).					
	facilitating communication between students and mentors (e.g., teaching assistants, senior students).					
	establishing study groups within my courses.					
	facilitating 1-1 peer interactions.					
	facilitating informal peer group conversations.					
	helping my students to find like-minded people in open community topics around hobbies/interests etc.					

Section C- Demographic data

- Gender. Male, Female, Prefer not to say
- University
- School / Faculty / Department

Your student-facing responsibilities include:

- Teaching
- Admin
- Support
- Other

What level of students do you work with?

- Undergraduate
- Postgraduate
- Other

Thank you a lot for your time and effort!
EDUBOTS research team
https://www.edubots.eu/

Annex V

Students' survey items

https://forms.gle/4uB9NvPknF7FSREj8

Dear student,

There is a growing interest in the use of chatbots in educational settings as they can provide efficient and timely services to students and to lecturers. Chatbots are software programs that communicate with users through natural language interaction interfaces. They conduct a conversation via auditory or textual methods with the user. They are often also called virtual agent or assistant, conversational agent, bot (short for robot).

This survey is taking place as part of EDUBOTS – Best practices of pedagogical chatbots in higher education – Erasmus Plus Project, funded by the EC, https://www.edubots.eu/.

Your feedback is very important to us for expanding knowledge about the use of chatbots in education.

Consent

The survey is anonymous and the aggregate data will be used anonymously in any publication/report related to the survey results. By clicking "next" you give your consent to participate in the survey.

Section A - Use of chatbots in Higher Education

Below you will find potential uses of chatbots in Higher Education. How do you personally evaluate the importance of the following chatbot uses? (1: not important at all, 5: very important)

I can see a great value of using a chatbot:		1 Not important	2	3	4	5 Very important
As a digital assistant enabling feedback loops	giving me personalized feedback on a course topic.					
	Giving me short quizzes on a course topic.					
	Allowing me to offer feedback to my instructor on his/her teaching method.					
As a remote tutor for	offering tutorials related to my courses.					
	Sharing and filtering relevant to the course resources.					
	Responding to frequently asked questions (FAQs) that relate to administration (e.g., deadlines).					
	Responding to FAQs that relate to the course content.					

	Addressing my content-related questions on my course.					
As a social bonding facilitator for	introducing me to other peers (e.g., with the use of ice breaking activities).					
	Facilitating communication with my mentors (e.g., teaching assistants, senior students).					
	Establishing study groups within my course.					
	Facilitating 1-1 peer interactions.					
	Facilitating informal peer group conversations.					
	Helping me to find like-minded people in open community topics around hobbies/interests etc.					

Section B - Demographic data

- Gender. Male, Female, Prefer not to say
- University
- Faculty / Department / Department
- Year of studies (multiple choice: 1st, 2nd, 3rd, 4th and above, MA/MSc, PhD)

Thank you a lot for your time and effort!
EDUBOTS research team
https://www.edubots.eu /