The history of chatbots: the journey from psychological experiment to educational object

Michal Černý

1 Department of Information and Library Studies - Faculty of Arts – Masaryk University, Arne Nováka 1, Brno 602 00, Czech, mcerny@phil.muni.cz

Abstract: Chatbots represent a strong and distinctive theme in the current literature on technology in education. What is lacking, however, is an analysis of them in terms of historical development or deeper historical-discursive classification. This paper focuses on the history of chatbots and places it in the context of a critical reflection on studies focusing on chatbots as educational objects between 2006-2021. It offers an analysis of each study and places them in the context of the development of the field as a whole. The study identifies three vital discourses that can be identified in the development of chatbots from a historical perspective - Turing-oriented, Searle-oriented and educational interaction-oriented.

Keywords: chatbot; chatbots; conversational agent; dialogic methods; artificial intelligence; Turing; learning objects; history of education technology

1. Introduction

Dialogue-based learning (Reznitskaya, 2012; Lyle, 2008; Shi et al., 2021) is known from ancient China and, in the Occidental cultural circle, mainly from the region of ancient Greece. Socrates' "grandmotherly obstetric" method, as captured by Plato in his dialogues, consisted of systematically asking questions (Fisher, 2013; Hajhosseiny, 2012). Socrates raises the topic, and the whole process of his teaching consists of asking educated questions to reveal what exactly he means by the subject. What is essential about this method is listening, caring, and, above all, active interest in the one with whom the dialogue is being conducted. Much of the philosophical tradition can be said to be associated with this method of systematically inquiring into what has been uncovered (brought out of hiding - ἀλήθεια) by a previous answer (Grondin, 1982; Campbell, 2017).

The dialogic teaching approach has been in pedagogy for a significant part of history. We can see Socrates, Plato and Aristotle building their teaching methods on it. Jesus chooses to work with parables as the primary didactic method, the goal of which is that actual individualisation
change. Augustine of Hippo (1853), as a teacher of rhetoric, understands conversation as a fundamental means of knowing and thinking, in his book *Confessions* are conceived as an introspective conversation with oneself. He then builds on this method in the Meditations of Descartes (2013). Descartesasticism, we can see dialogue or disputation as a fundamental method of university education at the latest with the expansion of Peter Lombarsky's Sentences (2007) at the turn of the thirteenth century Thomas Aquinas's *Theological Summa* (1981) (like his other textbooks) based on the idea of dialogue between student and teacher or between student and educational object, i.e., the textbook. In this respect, scholastic textbooks can be academic ideological precursors of chatbots. Among scientific works having the character of an educational dialogue, one should mention at least the Dialogue of Galileo Galilei (1990) from 1632, which is the most famous, though not the only, example of a systematic explanation of a scientific problem in the form of a dialogue to persuade and educate the reader. Dialogues follow Galilei or Alfred Rényi (1980), which follows the dialogical method in the first part, or Three Dialogues on Knowledge by Paul Karl Feyerabend (1999).

This enumeration only illustrates that dialogue is irreplaceable and integral in the learning process throughout the occidental concept of education from antiquity to the present. This is also evident in the considerable number of studies that have appeared in recent years on (Meijers & Hermans, 2018; Ruhalahti, 2019; Kloubert & Dickerhoff, 2020; Carvalho, 2022).

This study aims to outline at least the primary line of construction of chatbots as educational objects based on the theoretical concept of dialogue as an educational tool. The study will offer insight into their conceptualisation and development over the last 70+ years. We will show how a device designed primarily for psychological or computational experimentation has been shaped into an educational tool that can relate to the cultural context we have outlined above. As such, the definition of a chatbot is ambiguous (Sánchez-Díaz et al., 2018; Xu et al., 2017) - in the context of this study, we will understand it as a dialogic system that seeks to interact with the user through natural language in written form. It has no material body or similar "material structure", so it can be considered an autonomous software agent.

1.1. *From Turing to the first chatbot*

The beginning of chatbots is usually associated with Alan Turing's 1950 article (Turing & Haugeland, 1950). Turing formulates, among other things, the hypothesis that a machine indistinguishable from a human in its communication through text input and output could be considered intelligent. This idea is based on the premises of Ludwig Wittgenstein, who argued
that the limits of thought are the limits of language and vice versa (Hintikka & Hintikka, 1983; Wright, 1998). Wittgenstein thus represents the intellectual apex of a movement that George Lakoff would refer to as objectivism, which argues that thought is a conceptual matter and, in Wittgenstein's case, even a linguistic matter. Turing takes these premises seriously and applies them to engineering practice, setting out the primary field of research in artificial intelligence primarily up to the present day (Coniam, 2008; Neufeld & Finnestad, 2020).

The Chinese peace argument offers the most famous critique (Warwick & Shah, 2014; Shieber, 2004), formulated in 1980 by John Searle (1984, 1990). His basic idea is that simply manipulating dictionaries and databases, or even controlling a conversation, does not necessarily mean that a given person understands the content of that communicative act. In some ways, Searle modifies Saussure's belief (Harris, 1990) that simple translation is not a trivial matter because it is not a matter of translating labels in a muse but of the need to understand concepts in the context of the signifier and the signified. Turing, however, takes a position in which he argues that the communicative practice is crucial, not the proper understanding of the concepts, which is more a matter of philosophy than a shared understanding of intelligence (Jacquet et al., 2021).

Although Turing's assumptions are not trivial and could be significantly challenged with more careful discussion, they have the advantage of establishing a practical, workable concept for developing dialogue systems that can strive for intelligent interaction. The developer and user may need to be more relevant whether or not the object they are interacting with is intelligent. What is critical is the quality of the dialogue and the experience (Jacquet et al., 2021). This impact of Turing's article was realised by Weizenbaum (1966). In his article, he describes in quite some detail how he constructed the first chatbot called ELIZA. It was a simple program that could create a question from a user's announcement sentence, thus mimicking the work of a psychotherapist using the principles of pure language methods. ELIZA (Thorat & Jadhav, 2020; Natale, 2019) had the advantage that it did not introduce any evaluation, self-reported turns of phrase, or anything of the sort into the conversation with the client. At first glance, it was clear that she did not understand the content of the interview, but at the same time that she was, in some ways fulfilling the ideal of a psychotherapeutic interview. At the same time, the application was able to work with specific schemes of complementary questions that completed the whole interview.

In his article, Weizenbaum points out that this is nothing more than a question of computing natural language, in which English plays more of a role as a model or illustrative example, but
a chatbot could similarly work with other languages, from machine code to algebra. The choice of psychotherapeutic conversation still contains a specific ethos:

"The above remarks intend to further rob ELIZA of the aura of magic to which its application to psychological subject matter has to some extent contributed. Seen in the coldest possible light, ELIZA is a translating processor in Gorn's sense; however, it has been specially constructed to work well with natural language." (Weizenbaum, 1966, p. 43).

It is historically remarkable that the development of chatbots is going in precisely the opposite direction of Weizenbaum's thinking - we are trying to design chatbots that, on the contrary, work with an aura of magic because they enable authentic and functional conversations with real users. The second dimension of this magic is that ELIZA has become part of the cultural milieu, and chatbots based on it can be found in films in literature - from the late 1960s (the film THX-1138) to the present day.

The problem with ELIZA was that the psychologist or psychotherapist is educated with extraordinary verbal skills and the ability to understand the broad context of what the client is saying. It was almost impossible to simulate it using a simple algorithm, so the believability of such an application was low. In addition to efforts to improve the dialogue environment, there has also been a gradual search for scenarios that would significantly reduce such requirements on the chatbot - this is the path that led to the creation of PARRY (Shum et al., 2018; Thorat & Jadhav, 2020).

PARRY is conceived as a patient with schizophrenia characterised by a limited ability to communicate, loss of sensitivity to context and commonly understood insight into context. It was developed in 1971 by Colby et al. In their article, the authors introduce the issue as follows:

"Within the paradigm of computer science, distinctions are sometimes drawn between the activities of computer simulation and artificial intelligence. Yet in constructing models of psychological processes, the distinction can become blurred in places where overlaps emerge, as evident from our account of a model of artificial paranoia." (Colby et al., 1971, p. 1).

Interestingly, even in Weizenbaum's article, the whole construction is different - the authors first carefully examine what the communication of a person with schizophrenia looks like and then carefully model the entire scenario. Thus, the focus is not on the question of computer processing of natural language and the search for universal scenarios but on working on a
specific, very sharply defined topic. The paper also includes an initial test that shows that it was not easy for the psychologists (five psychologists) to make the distinction between human and machine in this form of imitation game (Colby refers directly to Turing) and shows that it can be relatively easy to succeed in this specific area. The general ambitions of such a project could not have been grand. Still, they showed that a well-thought-out script embedded in particular situations could create compelling and functional dialogue.

Jabberwacky was created in 1988 and later saw several other versions. Its concept is groundbreaking because it was the first to use structures that can be understood as elementary artificial intelligence. It had dialogues with users and tried to learn based on dialogical patterns (McNeal & Newyear, 2013). Its content domain was entertainment.

In terms of education, a significant milestone was reached in 1991 when the chatbot appeared in the game TINYMUD. It was a multi-user text-based game, so communication through text was seen as usual and familiar. In this virtual environment, it was possible to create a relatively simple chatbot that is highly believable because users do not expect to interact with a machine and not a human (Masche & Le, 2017; Adamopoulou & Moussiades, 2020). Only an utterly fundamental error can give it away.

The approaches PARRY and the chatbot in TINYMUD show a fundamental design approach. The chatbot did not need to be an all-encompassing, one-size-fits-all tool. Still, working with the appropriate scenario and context in which the chatbot is situated is essential to its design. The evaluation of individual chatbots should always be coupled with thinking through the ecosystem of other information and social interactions in which they are involved.

Clippy (Dale, 2016; Baez et al., 2020; Illescas-Manzano et al., 2021) was an attempt at dialogic and contextually designed user education when working with an office suite, which Microsoft first integrated into its office suite in Office 97. The basic idea was that an office paperclip (gradually supplemented with other appearance variations) tries to find out what the user is doing and offers help and assistance. The fundamental problem was the overly deterministic conception of the whole assistant concept, which could not perform individual functions but only tried to guide (teach) the person how to deal with a particular problem. The overly mechanical concept led to significant inaccuracy and little understanding of what the user needed, which is why Clippy became the subject of various parodies and satirical comments. Crucially, chatbot development worked with pre-prepared dialogues, and the user could only select a response option, which was trivial but valuable because the system was supposed to
react to events in the application and start a dialogue accordingly. The second important
dimension was that it was a chatbot deployed in a ubiquitous and essential application in a form
that could not be overlooked or ignored.

With the development of instant messengers, there is also the emergence of chatbots, which can
easily apply in their environment. In 2001, applications such as StudyBuddy and SmarterChild
(Molnár & Szüts, 2018; Madeira et al., 2022) were developed to support informal learning.
Students could write with the chatbots, and the chatbots used data from publicly available
search engines for their work. The focus was on conveying information and working in an
environment that expects quick and easy textual interactions.

ALICE was a system created by Richard Wallace, which was launched in 1995 and has
undergone many changes and improvements over time (Sharma et al., 2017; Shawar & Atwell,
contributors who were involved in the mathematical models built by the program. Program B
(1998-2000) began using an XML structure to format its knowledge data and has worked with
over 300 contributors. Program C in 2000 entailed porting the code to C, to be replaced by
Program D (based on Program B) in 2000. Program D used JAVA as the programming
language. These improvements brought technological change and were always linked to
developing the chatbot's capabilities in question.

In 2000, 2001 and 2004, he won the Loebner Prize, which focuses not on the appreciation of
artificial intelligence (on which ALICE is based), but on the ability to design dialogue to appear
as natural as possible. The Loebner Prize is a concept based on what Searle criticises in his
argument with the Chinese Room. ALICE works with artificial intelligence, an open-source
chatbot that other users can enhance, customise, or change. The whole concept makes no secret
that the underlying motivation was to use artificial intelligence to improve the ELIZA concept
(Wallace, 2009).

1.2. Personal assistants

Since 2010, there has been a rapid development of systems functioning as personal assistants
built primarily on verbal communication. These are dialogue systems that can be integrated into
special devices (Google Nest, Amazon Echo) or used in the operating system of a computer
(Google Assistant, Microsoft's Cortana) or a mobile phone (Siri, Google Assistant) or as a
specific program (IBM Watson). However, the range of individual solutions is much more
comprehensive. The essential features of these systems are the absence of a visual form
absence of humanisation by face), the attempt to use artificial intelligence, integration with other services (smart home, online services), voice control (limitation of typing) and emphasis on natural language dialogue. The term personal assistant is apt as it clearly shows the aim and purpose of developing these dialogue systems. In the following review, we focus on four products - Siri, Alexa, Assistant, and Cortana (Nobles et al., 2020; López et al., 2017) - that are among the most widely used.

Apple Siri was created in 2010 as a comprehensive voice assistant designed to make controlling Apple devices more convenient. It is currently available in most devices that Apple develops (Aron, 2011; Kepuska & Bohouta, 2018). Siri primarily has a female voice and attempts to communicate in natural language. Nevertheless, some users only use it for simple tasks such as finding their phone. She can control the phone, look up facts, manage devices, or work with navigation. Unlike Google, it works with the concept of a person who speaks with a name in a conversation, which is reflected in pop culture terms when Rajesh Ramayan Koothrappali interacts with her as a woman in The Big Bang Theory. The system uses several advanced algorithms for learning and processing user data, even where it heavily invades their privacy. The fact that the chatbot invades privacy is the most common criticism of Siri (Schönherr et al., 2020; Sharif & Tenbergen, 2020).

Alexa was launched in 2014 as a voice assistant for Amazon Echo systems, i.e. smart speakers (Lopatovska, I., & Williams, 2018; Lopatovska et al., 2019; Zwakman et al., 2021). At the same time, it is a multi-platform system that works on iOS and Android. Like Siri or Cortana, Alexa works with the concept of a dialogue with a person having a name (albeit partially suppressed and gender diversity in the case of Amazon). Alexa builds on Amazon's ambition to offer solutions for smart homes (and cars). It emphasises multimedia control, home automation and working with basic information (sporting events, traffic, weather, etc.). Alexa is also available through AWS to other developers who can integrate it into their products without paying fees. Thanks to Amazon Lex (as of 2017/2018), developers can define their responses by incorporating the artificial intelligence and learning system used by Alexa (Abualsamid & Hughes, 2018). This technology makes the voice assistant an accessible interface for development by more users.

Google Assistant was introduced in 2016, but its predecessor Google Now (Sullivan, 2009; Canbek & Mutlu, 2016), predates it by four years. Google Now aimed to change the paradigm in search and information access. Its designers believed that users had much information they could not access or needed to learn about, which is what this service was intended to provide.
The result was a broad ecosystem of tools capable of searching for information about traffic, movies, calendar reminders, or finding places to park. Google Assistant has brought two significant changes. It was tied to the Android operating system and the Google Home (now Nest) smart speaker, which gave it a whole voice interface. From the beginning, the app's goal was to access information through voice input (possibly through typing on a keyboard) and to perform individual actions with smart homes or devices. Google believes the dialogue system represents a unique way of mediating and retrieving information (Coates, 2019). There are currently many studies on using this technology in education (Lancioni et al., 2020; Chen et al., 2020; Sing et al., 2019; Tai & Chen, 2020).

Google considers its LaMDA tool, which combines various AI tools and pathways with a tremendous amount of training data, to be the best chat tool today (O'Leary, 2022; Griffiths, 2022; Grossman, 2022). While "regular" chatbots are based on the idea that specific dialogue scenarios can be developed and refined in development, Lambda aims for a universal dialogue system capable of discussing anything and has no such preferred set of initiation scenarios.

Microsoft's Cortana was a personal assistant (2014-2020) that was integrated into Microsoft's operating systems (Kim et al., 2016; Elwany & Shakeri, 2014). Cortana was (like Siri) female but needed to gain more popularity. This personal assistant supported simple activities (texting, emailing, calendar management), but the typical dialogue could have been more problematic. Its main weakness was the considerable time lag in development behind competitors from Apple, Amazon and Google.

Like Amazon, Microsoft is also trying to offer solutions developers could use in natural language computing. Since Cortana is no longer developmentally active, these are tools from the Microsoft Azure platform that can be implemented in third-party applications. Google also offers similar solutions to TensorFlow and DialogFlow. Besides, many environments are emerging where chatbots (whether using AI or not) can be exported, such as Tidio, Chatfuel, ManyCahts, Snachbot.me, Wit.ai, Aivo and many others. Chatbots have become an essential marketing and communication tool in the last few years and can also have many uses in other areas.

2. Chatbots in education

So far, we have focused on the historical development of chatbots as such. In this section, we would like to follow their application directly in the educational environment. Analysis will be incomplete and systematic but aims to show the broader possibilities of implementing this
technology in education in a broader pedagogical context. In this research, we will work with the SCOPUS database, from which we will select one study each year (the first available result is from 2006) according to the parameters of citation and relevance. On their selection, we will illustrate the gradual evolution of chatbots in education.

At the same time, it should be stressed that this is a partial analysis. Shawar and Atwell (2004), for example, discuss the possibilities of accessing information systems through chat tools. Heller et al. (2005) work with a chatbot, Freudbot, which models dialogue with Freud as an educational tool. The beginnings of the systematic use of chatbots for educational purposes can be linked to around 2005.

The query was used for the search:

```
TITLE-ABS-KEY (chatbot OR chatbots AND education) AND (LIMIT-TO (LANGUAGE, "English"))
```

Figure 1 shows the gradual increase in studies over time, where it is clear that we can see massive growth from 2017 onwards. The drop-off in 2022 is because not all analyses for this year are published (this study was due in October 2022). The overview study ends in 2021 in selecting texts. Figure 2 shows that although the topic of chatbots in education is fixed on computer science and computer science, it is fundamentally interdisciplinary. Figure 3 then offers a glimpse of some of the cultural limitations of this research, as many languages, regions and cultures are not represented in the research. The English-speaking environment (USA, India, UK, Australia) is dominant. A closer examination (but not the focus of this study) would show that region also influences the thematic focus of each study.
Fig. 1. Data from the SCOPUS database shows the number of documents indexed when working with a search query over time.

Fig. 2. Data from the SCOPUS database show the thematic affiliation of individual studies.
The search produced 438 documents. For each year, we selected one most cited paper, listed below. Within each year, we selected the paper with the highest citation rate and, at the same time, relevant to the study being conducted. Results start from 2006, the first year in which results are available in the Scopus database. Technically, we have always restricted the results to one specific year and ranked the results by citation. We did this for all years - from 2006 to 2021.

Lu et al. (2006) focus on developing a chatbot for teaching English as a second language. Their chatbot is based on the paradigm anchored in instant messaging and combines working with computer processing of natural language and pre-prepared responses. It is on these that most of the dialogue is built. The authors work with the TutorBot chatbot, which serves, on the one hand, to support students in synchronous communication but also provides the teacher with information about individual student progress.

Kerly et al. (2007) highlight the importance of chatbots for adaptive learning that adapts to students' needs. They define the principles that such a chatbot should meet in its development - the ability to keep the learner on topic, manage data on learner behaviour, respond to and learn from previous records, web integration, and a body of knowledge of semantic reasoning. The chatbot is treated in this paper primarily as a technical problem, but one that may - if mastered - have significant practical implications.
Bigham et al. (2008) describe chatbots as integrating and activating blind students. As part of the project, blind students programmed the chatbot themselves, increasing their code-generation competence and affinity for the information. On the other hand, they created a chatbot designed so that other blind users could interact with it. Text input in an austere environment can be a prerequisite for universal design. At the same time shows that a chatbot can be an available assistive technology and an educational object that blind users can work with.

Jia (2009) presents a chatbot for learning English that works with natural language-based dialogue. The author tries to create an environment that will simulate a chat with a natural person, leading to an educational interaction. It strives for dialogue's semantic and syntactic analysis, making it more persuasive and context-working than ELIZA. The study concludes with a statement that we consider significant for the development of chatbots before the massive emergence of artificial intelligence in the likes of neural networks and publicly available frameworks for NPL:

"Solely in NLP, many problems are still hard to be solved, such as the textual ambiguity and entailment. How to overcome these problems with current available technologies is still a great challenge to us. The contradiction between the high response speed and complex, deep, syntactic and semantic analysis should also be paid great attention, especially in the web environment." (Jia, 2009, p. 255)

Lokman and Zain (2010) work with a chatbot to educate people with diabetes to simulate a doctor. They point out a typical problem related to simple chatbots working with individual responses in isolation, leading to less relevant responses and less convincing and natural dialogue. Therefore, their chatbot works in the context of previous questions and answers in the dialogue. This relatively simple procedure does not require sophisticated technological techniques, leading to a rapid improvement in the quality of the outputs.

Crutzen et al. (2011) focused on the information sought by adolescents on the topics of sex, drugs and alcohol. The study worked with a chatbot that dialogically answered students' questions. The research showed positive student interaction with the content and student perceptions. The conversations were relatively long and showed greater satisfaction with information needs than other sources of information, including hotlines. The study's authors show the great potential of chatbots in topics associated with shame, fear, or other social-psychological barriers.
In their study, Niranjan et al. (2012) work with a chatbot as part of an e-learning environment in which questions, usually of an organisational nature, need to be answered. The authors draw attention to the fact that even though online courses are being developed, they require the constant presence of a live teacher to answer students’ questions. In the case of this study, the chatbot uses a naive Bayesian classifier to understand the question and evaluate it and a database of answers to questions authored by the teacher. At the same time, the teacher receives information from the chatbot about which questions are the most frequent, which allows him to modify the educational process.

Lundqvist et al. (2013) investigated ontology models that can be used for developing chatbots or modelling their dialogues. The study’s authors point out that their model using AIML and the OwlLang scripting language achieves high satisfaction with users’ information needs.

Gulenko (2014) describes a design approach to chatbot development. The first step identified essential cybersecurity topics relevant to the target audience. In the second step, a chatbot was created to work with the three most important ones. From this data and pedagogical analysis, sub-goals were constructed, with each goal being accompanied by a problem that prevents its saturation and a suggestion of what needs to change on the user's side to bring about a behaviour change. The study is interesting because it emphasises the development of the chatbot as an educational object rather than a technical solution.

In their theoretical historical study, Maggy et al. (2015) argue that providing an emotional interaction between the chatbot and the user is crucial to the feeling of learning. Combining this inspirational design with a sense of authenticity and personalisation is a fundamental prerequisite for a functional concept of the chatbot as a learning object. The study shows that education is a complex psycho-social process and that a good chatbot for marketing can be something other than an excellent educational object. Focusing on the feelings that chatbots evoke in users is a crucial pedagogical input that needs to be systematically worked with at the level of theory. Maggy et al. (2015) highlight that chatbots are important learning objects, but they need a broader theoretical background for their development.

Pereira (2016) creates a chatbot @dawebot that focuses on practising knowledge for a test in the form of an interactive dialogic test. A relatively simple implementation working with predefined questions on the test and feedback is tested on computer science students. It shows that students would welcome more such chatbots in other courses and feel more engaged and better prepared for the test. The study shows how a relatively simple tool can be further scaled...
up (implemented in other courses) to help students. The author points out that chatbots can have many advantages, but one of the key ones is the interaction environment, which is fundamentally more pleasant and better than an LMS.

Frey et al. (2017) analyse the results of an experiment in foreign language teaching. One group had a live tutor, the other a chatbot. The authors analysed the course's decline in interest and motivation for both target groups. Although the sample size was limited (n=122), and the data is strongly dependent on the chatbot and the tutor's specific configuration, human motivation is essential for a course that lasts for a more extended time. A chatbot is a good tool for occasional activation or as a supplement to the broader learning ecosystem. Still, it is less effective a motivator for students than a live human teacher.

Io and Lee (2018) offer an extensive bibliometric analysis studying the current (2018) state of chatbot development. The authors point out that since 2015, we can identify a sharp increase in studies focusing on chatbots, which they argue is related to a significant shift in the development of artificial intelligence. Still, most of the studies analysed use AIML, which is consistent with the studies analysed above. At the same time, it shows that breakthroughs can be expected in the process of chatbot development and its methodology. According to this analysis, the main area in which chatbots will be applied is in the field of education and implementation in mobile devices.

Shorey et al. (2019) are only the seventh most cited study in the SCOPUS database for 2019 search queries but the first entirely relevant to education. The authors used a chatbot to simulate a virtual patient. The chatbot was complemented with a voice and virtual interface in the UNITI environment, aiming to increase the educational experience's authenticity. Here, the chatbot plays the role of a patient with whom nursing students try to converse and solve specific problems according to selected scenarios. The study notes that despite the relatively small sample of students and the initial stage of development, it can help with feelings of confidence and develop communication skills.

Kramer et al. (2020) is the sixth study in order of sensitivity, and the relevance to education is Perez et al. (2020), the fifth but only a general overview study. The authors focus on the specific phenomenon of embodied chatbots aimed at developing healthy lifestyles. These are animated conversational systems that change user behaviour through conversation. It shows that a chatbot in education may not only be focused on cognitive development and knowledge-based learning, but the study suggests great potential in coaching, personal development, lifestyle change, etc.
The study continues the inclination, evident in the last few years, towards an emotion-oriented emphasis on the design of chatbots as learning objects.

Chocarro et al. (2021) is the fourth most cited paper in this overview study but the first relevant to the content of the study. They focus on teachers’ perceptions of chatbots. The study yields several essential findings - willingness to use chatbots is not related to teachers’ digital competencies, so it does not represent a relevant barrier in this area. Secondly, the actual quality of chatbots is also not a determining factor for the willingness to implement chatbots in the educational process. The authors state, “Our results confirm a positive and substantial impact of the perceived usefulness for using the chatbot on teachers' technology usage intention. Improving the performance and usefulness of chatbots is a critical determinant for teachers when considering adopting this technology for their jobs.” (Chocarro et al., 2021, p. 11) These results should be used when looking for suitable implementation schemes for integrating chatbots into education - it is in actual implementation that chatbots play a crucial role.

3. Discussion

The study aimed to analyse some aspects of the development of chatbots as learning objects. Currently, there are a large number of overview studies (Hwang, G. J., & Chang, 2021; Pérez et al., 2020; Cunningham-Nelson et al., 2019; Kuhail et al., 2022) that focus on the topic of chatbots, but a more systematic or comprehensive grasp of the historical development is lacking, except for a general study by Adamopoulou and Moussiades (2020). Methodologically, this study works with two pillars - the first part analyses the evolution of chatbots as such in a brief historical outline, and the second offers a specifically constructed historical overview study focused on scientific studies. At this point, we would like to identify some developmental landmarks or trends that can be analysed in understanding chatbots as educational objects.

The first trend that we can identify is a gradual transformation in the understanding of what a chatbot is - the Turing tradition seeks to develop a system capable of discussion in a way that is unrecognisable from a human (Powell, 2019; Qaffas, 2019). It works with the notion that indistinguishability is key to communication. This will create a dimension of trust that can be applied to the learning environment. Increasing the quality of dialogue (Abdul-Kader & Woods, 2015) represents a significant turning point in how the possibilities of working with chatbots are changing and evolving. At the same time, however, we must stress that the significance of its impact on the educational process is not only linked to its indistinguishability. Two inherently present concepts in dialogue systems development are juxtaposed here.
The first is related to the question of whether the requirements for a dialogue system used for education should be higher - Black Lemoin has worked for Google on testing LaMDA, the company's most advanced language generative neural network, and has come out in the media with a statement that he believes this neural network has self-consciousness (Tiku, 2022). This path of development and reasoning is towards developing artificial actors that are meant to be as human-like as possible.

On the other hand, it is possible to encounter applications that try to emphasize that the goal of development is not the "other person" but a dialogue partner with whom it is possible to interact through natural language, but at the same time, knowing that it is not a human. Such interactions allow for the design of significantly different educational activities and lead to varying experiences of interacting with the chatbot. A specific (extreme) branch of this research is the development of systems without artificial intelligence (Tamayo et al., 2020).

Studies in recent years by Kramer et al. (2020) and Shorey et al. (2019) show that the direction of development is not only related to actual dialogue based on the written word or voice work but that the goal is to model a virtual actor that has a digital "somatic" structure - face, emotions, expressions, gestures. On the one hand, such a conception leads to a higher authenticity of the chatbot. Still, it also highlights that textual communication constitutes only one part of understanding the information communicated. The findings of Frey et al. (2017) show that the awareness that the communication partner is a chatbot and not a human is not negligible in the education process.

Even systems with such a high degree of believability, such as the LaMDA chatbot, which would feature a graphical interface in Unity as described by Shorey et al. (2019), will face a psychological barrier in terms of long-term interaction. However, the short-term effect of activation is indisputable. Equally indisputable is the positive effect of chatbots on learning when they become part of a broader ecosystem. Their goal is not and cannot be to replace the teacher as a living actor but to complement the educational structure with an educational object that brings specific benefits.

The clear trend is to change the inherent mechanisms on which chatbots are based. We can observe a shift from simple conversational frameworks and models working with keywords (Tamayo et al., 2020; Satyanarayana, 2019) to complex tools using artificial intelligence and neural networks. An unmistakable trend in educational applications is using frameworks that
allow leveraging technical solutions developed by a third party and focusing on the actual educational implementation.

Much attention has been paid progressively to the design of educational interaction (Topal et al., 2021; Stuij et al., 2020; Vázquez-Cano, 2021) - while for the first studies we analysed, the dialogue needed to be made believable, current solutions focus much more their attention on how to use chatbots in a specific educational domain, how to set up the chatbot behaviour and the whole educational process so that the chatbot forms only one particular part of a broader educational interaction.

The fact that the chatbot has gradually become a learning object (FAO, 2021) clearly illustrates the above shift - the goal is no longer to work with the chatbot per se but to find areas in which it can be used, fulfil explicit educational purposes, and function as part of a broader ecosystem of learning objects. In this area, a clear trend can be identified of a gradual focus on "how to build a chatbot" to find areas where its implementation is effective and meaningful.

Although the development of chatbots in education has been systematic since 2006 at the latest (as this study shows) and has been overgrowing since 2017, there still needs to be a more systematic, empirically oriented set of clear recommendations or methodologies for working with chatbots. Individual studies and methodologies tend to be eclectic, and we expect this is where basic pedagogical research will be applied (Chang et al., 2021; Mendoza et al., 2020). A gradual transition needs to be made from an overview study to comprehensively analyse specific aspects of chatbot design.

Typical applications for current applications are in school subjects; history (Noh & Hong, 2021), health education (Hwang, G. J., & Chang, 2021; Palanica et al., 2019), languages (Hwang & Chang, 2021; Sung, 2020; Haristiani, 2019) and many other areas. Alternatively, in specific areas of learning support, such as intelligent feedback (Lundqvist et al., 2013; Lee & Fu, 2019), immediate assistance to the student (Berger et al., 2019; Okonkwo & Ade-Ibijola, 2021) and alternatives to learning management systems (Tamayo, 2020; Villegas-Ch et al., 2020).

4. Conclusion

Chatbots as technology have made significant developments in recent years and have become part of the educational environment. They can now be used for motivation, feedback, as a substitute for information systems, and as a tool to practice knowledge or develop
competencies. They constitute a tool - almost unlimited in terms of the list of possible applications. Nevertheless, the historical development that we have tried to outline in this study shows that the future of technology lies not in replacing all other educational objects but in developing the ability to integrate them effectively into educational environments and developing methodologies for their appropriate use. The chatbot is (perhaps more than other forms) very sensitive to working with language, and students may experience frustration or arrogance when interacting with it, which may be due to minor stylistic problems or the chatbot's inability to accept the correct answer.

In this study, we uncovered three critical historical discourses that have been applied to chatbot development and that frame the developer's approach to chatbots:

a) Turing - the goal is to create a conversational agent indistinguishable from a human in its expressions.

b) Searlean - the goal is to create a conversational robot with consciousness or at least a specific understanding of itself.

c) Technological - the goal is to create an entity that will be different from humans and allow performing interactions inaccessible to humans with a transparent layer of communication in which dialogue will be used. Still, at the same time, it will be evident that it is not an imitation of humans.

The first is the most represented discourse. The second is that many theoretical studies drawing attention to the problem of Turing discourse begin to apply technically. However, from an educational point of view, the third discourse is the key. It allows not to replace the teacher but to extend the range of educational possibilities and complement it. For example, the future of chatbots in online courses may involve a close collaboration between tutor and chatbot on joint comprehensive learner support.

References


Crutzen, R., Peters, G. J. Y., Portugal, S. D., Fisser, E. M., & Grolleman, J. J. (2011). An artificially intelligent chat agent that answers adolescents' questions related to sex, drugs, and


International Conference on Human-Computer Interaction (pp. 93-107). Springer. https://doi.org/10.1007/978-3-030-50506-6_8


About Author

Michal ČERNÝ is an Assistant Professor in the Department of Information Studies and Library Science, where his research focuses on technology in education and its philosophical and sociological reflection. As a senior researcher, he led a project on information literacy and now works as a researcher in projects on personal learning environments and immersive virtual reality learning. He received his PhD in Social Pedagogy when his work focused on the changing philosophy of education in the context of technological change.