Machine Learning with Applications 2 (2020) 100006

Contents lists available at ScienceDirect

Machine Learning with Applications

journal homepage: www.elsevier.com/locate/mlwa

Chatbots: History, technology, and applications

Eleni Adamopoulou *, Lefteris Moussiades

Department of Computer Science, International Hellenic University, Agios Loukas, 65404 Kavala, Greece

A R T I C L E I N F O

Keywords:
Chatbot
Pattern matching
Machine learning
Natural dialog interfaces
Natural language processing
Human–computer interaction

A B S T R A C T

This literature review presents the History, Technology, and Applications of Natural Dialog Systems or simply chatbots. It aims to organize critical information that is a necessary background for further research activity in the field of chatbots. More specifically, while giving the historical evolution, from the generative idea to the present day, we point out possible weaknesses of each stage. After we present a complete categorization system, we analyze the two essential implementation technologies, namely, the pattern matching approach and machine learning. Moreover, we compose a general architectural design that gathers critical details, and we highlight crucial issues to take into account before system design. Furthermore, we present chatbots applications and industrial use cases while we point out the risks of using chatbots and suggest ways to mitigate them. Finally, we conclude by stating our view regarding the direction of technology so that chatbots will become really smart.

1. Introduction

Artificial intelligence (AI) has influenced how we engage in our every day activities by designing and evaluating advanced applications and devices, called intelligent agents, which can perform various functions. A chatbot is an artificial intelligence program and a Human–computer Interaction (HCI) model (Bansal & Khan, 2018). According to the dictionary, a chatbot is ''A computer program designed to simulate conversation with human users, especially over the Internet'' (*Chatbot | Definition of chatbot in English by Lexico Dictionaries*, 2019). It uses Natural Language Processing (NLP) and sentiment analysis to communicate in human language by text or oral speech with humans or other chatbots (Khanna et al., 2015). Artificial conversation entities, interactive agents, smart bots, and digital assistants are also known as chatbots.

Apart from imitating human interaction and amusing people, chatbots are useful in various other fields in education, business and e-commerce, health, and entertainment (Shawar & Atwell, 2007). Productivity is the most important motivation for chatbot users, although different motivations include entertainment, social factors, and novelty interaction. Moreover, in business, chatbots have become so common because they reduce service costs and can handle many customers simultaneously. Chatbots are more friendly and attractive to users than, for example, the static content search in frequently asked questions (FAQs) lists. They offer users comfortable and efficient assistance when communicating with them; they provide them with more engaging answers, directly responding to their problems (Brandtzaeg & Følstad, 2017) (R. Ranoliya, Raghuwanshi, & Singh, 2017).

Most of the time, users feel chatbots as friendly companions and not just as mere assistants (Costa, 2018). Forty percent (40%) of user requests are emotional than informative (Xu, Liu, Guo, Sinha, & Akkiraju, 2017). It is the evolution of Machine Learning and sentiment analysis that equipped chatbots with the ability to respond emotionally to customers (Følstad, Bertinussen Nordheim, & Alexander Bjørkli, 2018).

The degree of trust a chatbot gains from its use depends on factors related to its behavior, appearance, and others related to its manufacturer, privacy issues, and protection (Wallace, 2009). The development of this relationship of trust is also supported by the level to which the chatbot is human-like, which depends on the visual characteristics, how closely its name is related to a person, its personality, and its efficiency to handle human language (Go & Sundar, 2019). Emotion is another essential aspect to humanize a chatbot, and there have been many approaches to building an emotionally aware chatbot (Pamungkas, 2019).

Developments in AI enhance the abilities of chatbots to mimic human agents in conversation. However, human–chatbot communication has noticeable differences in the content and quality in comparison to the human–human discussion. The duration of a human–chatbot conversation is long. People often use concise language with poor vocabulary or even lousy language (Hill, Randolph Ford, & Farreras, 2015). It is worth noting, that the crucial difference among chatbots and humans is the perception of empathy, as chatbots are less capable of conversational understanding than humans are. However, progress is being made, and chatbots are gradually becoming more fully aware of their interlocutor's feelings (Fernandes, 2018).

* Corresponding author.
E-mail addresses: eladamo@cs.ihu.gr (E. Adamopoulou), lmous@cs.ihu.gr (L. Moussiades).

https://doi.org/10.1016/j.mlwa.2020.100006
Received 7 July 2020; Received in revised form 19 October 2020; Accepted 19 October 2020
Available online 9 November 2020

Moreover, human–chatbot communication changes depending on the disclosure or not of the conversational partner. Undisclosed chatbots are four times more productive than novice sales staff, and their ability reaches that of specialized consumer shopping employees (Luo, Tong, Fang, & Qu, 2019). However, subjective human perception makes people consider disclosed chatbots less informed and emotionally intelligent. Thus, when customers discover during a conversation that they talk to a chatbot, they get upset and buy fewer products. Therefore, a method for delayed disclosure of a chatbot was used (Luo et al., 2019). In Mori, MacDorman, and Kageki (2012), the so-called "uncanny valley theory" examines the uncomfortable feelings that a person is experiencing when he/she does not know if the interlocutor is a human or a computer program (Skjuve, Haugstveit, Følstad, & Brandtzaeg, 2019). Personification and contact in people's disclosures on sensitive topics, such as social stressors, have also been examined (Sannon, Stoll, DiFranzo, Jung, & Bazarova, 2018).

Although we live in an age when our interlocutor can be a real person or a chatbot without caring about his true identity (Dale, 2016), a bias against gender is exposed (Costa, 2018). Most chatbots are typically used as personal assistants and secretaries to execute activities that mimic traditionally feminine stereotypes and convey these characteristics through stereotypic behaviors.

In this context, we set the research question of this literature review as follows: What are the evolution and current state in chatbots' technology and their applications?

Our objective is through a descriptive investigation of the research question to organize critical information that is a necessary background for further research activity in the field of chatbots. In this context, we present the evolution history of chatbots and state out drawbacks at each stage; we also present the different approaches for chatbot construction and discuss chatbots applications and industrial use cases.

The main contribution of our work consists of:

- We enhanced the chatbots classification system proposed by (Nimavat & Champaneria, 2017) by adding the 'Permissions' category, which distinguishes chatbots to Open Source or commercial and extending the 'Communication channel' category by adding the image input.
- We propose an architectural design for chatbots that is general, includes sufficient details, and complete compared to other related works.
- We highlight critical issues that every developer should take into account before the chatbot design.
- We point out the drawbacks of today's chatbots and comment out how one can mitigate them.
- We state our view on the direction that technology should take for chatbots to meet the needs of human communication through speech.

The rest of the paper is structured as follows. In Section 2, we take a historical look back from the beginning of chatbot creation to the present day, pointing out the scientific community's interest. Then, in Section 3, we propose a classification of existing chatbots. Pattern-matching and machine learning approaches are described in Section 4. In Section 5, we suggest a general chatbot architecture, while in Section 6, we discuss issues related to chatbot development. In Section 7, the weaknesses and threats of chatbots are addressed, while in Section 8, some chatbots applications are presented. Finally, in Section 9, we report the limitations of this study, discuss its implications, and present our view on the direction the technology should take.

2. History

In 1950, Alan Turing wondered if a computer program could talk to a group of people without realizing that their interlocutor was artificial. This question, named Turing test, is considered by many to be the generative idea of chatbots (Turing, 1950). The first chatbot with ELIZA name was constructed in 1966. ELIZA simulated a psychotherapist's operation, returning the user's sentences in the interrogative form Weizenbaum (1966). Its ability to communicate was limited, but it was a source of inspiration for the subsequent development of other chatbots (Klopfenstein, Delpriori, Malatini, & Bogliolo, 2017). ELIZA uses pattern matching and a response selection scheme based on templates (Brandtzaeg & Følstad, 2017). A drawback of ELIZA is that its knowledge is limited, and therefore, it can discuss only in a particular domain of topics. Also, it cannot keep long conversations and cannot learn or discover context from the discussion.

In 1972, PARRY appeared; It acted as a patient with schizophrenia (Colby, Weber, & Hilf, 1971). PARRY is considered more advanced than ELIZA is as it is supposed to have a "personality" and a better controlling structure. It defines his responses based on a system of assumptions and "emotional responses" activated by the change of weights in the user's utterances (Colby, Hilf, Weber, & Kraemer, 1972). PARRY was used in an experiment in 1979 when five psychiatrist judges interviewed by teletype a patient to decide whether he was a computer program or a real schizophrenic patient. Therefore, psychiatrists gave ten diagnoses. The first psychiatrist gave two correct diagnoses; another gave two incorrect ones. The third considered that both subjects were real patients, and the other two diagnosed that both subjects were chatbots (Heiser, Colby, Faught, & Parkison, 1979). However, the sample of five psychiatrists is small, and the meaning of the findings is not clear as people with schizophrenia have a degree of incoherence in their speech. In general, PARRY is considered a chatbot with low capabilities concerning language understanding and the ability to express emotions. It also has a low speed of responding, and it cannot learn from the conversation.

Artificial Intelligence is firstly used in the domain of the chatbots with the construction of Jabberwacky in 1988 (Jabberwacky, 2019). Jabberwacky was written in CleverScript, a language based on spreadsheets that facilitated the development of chatbots, and it used contextual pattern matching to respond based on previous discussions. Still, Jabberwacky cannot reply to high speed and work with a massive number of users (Jwala, 2019).

The term Chatterbot was first mentioned in 1991. It was a TINYMUD (multiplayer real-time virtual world) artificial player, whose primary function was to chat. Many real human players seemed to prefer talking to Chatterbot than a real player. The Chatterbot succeeded because, in the TINYMUD world, players assumed that everybody was a human and might cause doubts only if it made a significant mistake (Mauldin, 1994). Dr. Sbaitso (Sound Blaster Artificial Intelligent Text to Speech Operator) (Dr. Sbaitso, 2019), a chatbot created in 1992, was designed to display the digitized voices the sound cards were able to produce. It played the role of a psychologist without any sort of complicated interaction (Zemčík, 2019).

Another step forward in the history of chatbots was the creation, in 1995, of ALICE (Artificial Linguistic Internet Computer Entity), the first online chatbot inspired by ELIZA (Wallace, 2009). ALICE was based on pattern-matching, without any actual perception of the whole conversation (Marietto et al., 2013) but with a discussion ability on the web that allowed longitude and included any topic. However, a few years had to pass before it was improved to win the title of the Loebner Prize of the best human-like computer program (Bradeško & Mladenić, 2012). ALICE was developed with a new language created for this purpose, Artificial Intelligence Markup Language (AIML), which is the most critical difference between ALICE and ELIZA. ALICE's Knowledge Base consisted of about 41,000 templates and related patterns, a vast number comparing to ELIZA that had only 200 keywords and rules (Heller, Procter, Mah, Jewell, & Cheung, 2005). However, ALICE did not have intelligent features and could not generate human-like answers expressing emotions or attitudes.

In 2001, there was a real evolution in chatbot technology with the development of SmarterChild (Molnár & Zoltán, 2018), which was available on Messengers like America Online (AOL) and Microsoft

(MSN). It was the first time that a chatbot could help people with practical daily tasks as it could retrieve information from databases about movie times, sports scores, stock prices, news, and weather. This ability marked a significant development in both the machine intelligence and human–computer interaction trajectories as information systems could be accessed through discussion with a chatbot.

The development of Artificial Intelligence chatbots went one step further with the creation of smart personal voice assistants, built into smartphones or dedicated home speakers, who understood voice commands, talked by digital voices, and handled tasks like monitoring home automated devices, calendars, email and other. Apple Siri (Siri), IBM Watson (*Watson Assistant | IBM Cloud*, 2020), Google Assistant (*Google Assistant, your own personal Google*, 2019), Microsoft Cortana (*Personal Digital Assistant—Cortana Home Assistant—Microsoft*, 2019), and Amazon Alexa (*What exactly is Alexa? Where does she come from? And how does she work?*, 2019) are the most popular voice assistants. There are also many other less famous voice assistants owing unique characteristics, but the same core functions. They connect to the Internet and, in contrast to their predecessors, they create quickly meaningful responses (Hoy, 2018).

Siri (Siri), developed by Apple in 2010, pioneered the way for personal assistants. Users make inquiries and conversations with it through Messengers using voice commands, and it includes integration with audio, video, and image files. Siri makes recommendations and responds to user requests using various internet services, while it adapts, with constant use, to the users' language usages, searches, and desires (Siri, 2020). Although Siri is sophisticated, it is not without weaknesses. It requires an internet connection. It is multilingual, but there are many languages it does not support, while navigation instructions are supported only in English. It also has difficulties hearing the interlocutor, who has a heavy accent or in the presence of noise (Soffar, 2019).

In 2011, a chatbot called Watson (*Watson Assistant | IBM Cloud*, 2020) was created by IBM. Watson could understand the natural human language well enough to win two previous champions on the quiz competition "Jeopardy", in which participants received some information in the form of answers and should guess the corresponding questions. Years later, Watson enabled businesses to create better virtual assistants. Moreover, Watson Health was designed to help doctors in healthcare diagnose diseases. However, a drawback of Watson is that it supports only English.

Google Now (Google now, 2020), developed in 2012, was initially used to give information to the user taking into account the time of day, location, and preferences. Google Assistant (*Google Assistant, your own personal Google*, 2019), which was developed in 2016, constitutes the next generation of Google Now. It has a more in-depth artificial intelligence with a friendlier, more conversational interface and delivers information to users predicting their requirements. However, it has no personality and its questions may violate the user's privacy as it is linked directly to their Google Account.

Microsoft designed a personal assistant Cortana developed in 2014 (*Personal Digital Assistant—Cortana Home Assistant—Microsoft*, 2019). It recognizes voice commands and performs tasks such as identification of time and position, support people-based reminders, send emails and texts, create and manage lists, chitchat, play games, and find information the user requests. The major drawback of Cortana that has been reported is that it can run a program that will install malware (*Cortana security flaw means your PC may be compromised*, 2018).

The same year, Amazon (*What exactly is Alexa? Where does she come from? And how does she work?*, 2019) introduced Alexa, which is built into devices for home automation and entertainment and making in this way the Internet of Things (IoT) more accessible to humans. An innovation is that developers can use Alexa Skills Kit (ASK) to create and publish free or paid Alexa skills. As we report in Section 7, Alexa introduces security issues.

Although personal voice assistants enable voice communication with their users, misunderstandings often occur, as they cannot understand the particular language people use in oral speech or fail to understand the whole context of the conversation.

Early in 2016, an evolution of Artificial Intelligence Technology occurred that changed dramatically the way people communicate with manufacturers. Social media platforms allowed developers to create chatbots for their brand or service to enable customers to perform specific daily actions within their messaging applications. At the end of 2016, 34.000 chatbots covered a wide range of uses (Wizu, 2018) in fields like Marketing, Supporting Systems, Health Care, Entertainment, Education, and Cultural Heritage. Thousands of text-based chatbots with specific features were developed for popular messaging platforms, industrial solutions, and research (Dale, 2016). Moreover, the Internet of Things (IoT) introduced a new era of connected smart objects where the use of chatbots improved communication between them (Kar & Haldar, 2016).

Worth mentioning is also the Microsoft XiaoIce, which is an AI chatbot that satisfies the human need for sociability. Apart from its personality, its contribution to the development of the chatbots is that it has intelligent and emotional quotient (IQ–EQ). It establishes long emotional relationships with its users, taking into account the cultural peculiarities and ethical issues (Zhou, Gao, Li, & Shum, 2019).

The way chatbots nowadays engage in a discussion is entirely different from their predecessor Eliza. They can share personal thoughts and family drama events, be relevant but also confusing, and deceive just as humans do (Shah, Warwick, Vallverdú, & Wu, 2016).

A growing increase in the use of chatbots was observed, especially after 2016 (Fig. 1). According to Fig. 2, the country that has shown more research interest in chatbots is the USA, while the United Kingdom and Japan follow with less than one-third of the number of papers published in the USA.

3. Categories of chatbots

The figure below (Fig. 3) shows the different categories of chatbots. Each category has been defined based on a simple criterion, and a chatbot can belong to more than one category at a time.

A chatbot can access a range of knowledge, which determines its Knowledge Domain. Chatbots that can answer any user question from whichever domain are called Generic chatbots. Chorus (*Chorus—A Crowd-Powered Conversational Agent on Google Hangouts*, 2020) is an example of a generic chatbot. Chatbots like Guardian (Good & Wilk, 2016), CRQA (Savenkov & Agichtein, 2016), or AskWiz that operate in more than one domain are Cross or Open-Domain chatbots. In contrast, Domain-Specific chatbots like InstuctableCrowd, Legion: Mobile, or SnapTravel can respond only to questions concerning a particular knowledge domain (Kucherbaev, Bozzon, & Houben, 2018).

Chatbots that offer services like booking services in restaurants, airlines, or searches in FAQ without being a friendly companion, belong to Interpersonal chatbots. Intrapersonal chatbots are close companions that live in the user's domain and understand his needs. They are usually connected to messenger applications like Slack and WhatsApp. Finally, Inter-agent chatbots provide communication with other chatbots. Alexa and Cortana are two chatbots that were integrated to communicate with each other (Nimavat & Champaneria, 2017).

The primary Goal a chatbot aims to achieve classifies them in Informative, Chat-based/Conversational, and Task-based chatbots. When users communicate with a chatbot to get specific information stored in a fixed source, Informative chatbots like Guardian, Facebook M, or FAQ chatbots are used. Chat-based/Conversational chatbots hold a natural conversation with the user like a real person would do. Finally, Task-based chatbots handle different functions, such as room booking, and are excellent at requesting information and responding to the user appropriately (Kucherbaev et al., 2018)(Nimavat & Champaneria, 2017).

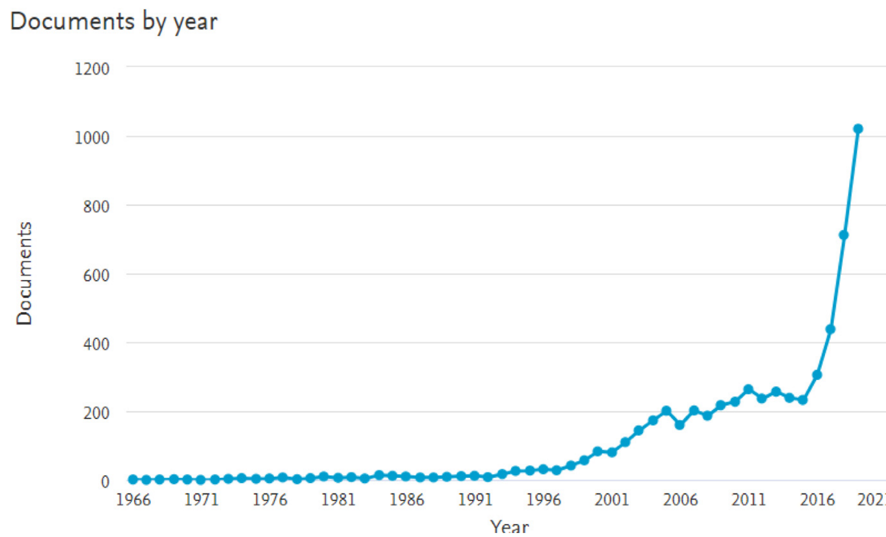

Fig. 1. Search Results in Scopus (*Scopus preview—Scopus—Welcome to Scopus*, 2020), from 1966 to 2019 for the keywords "chatbot" or "conversation agent" or "conversational interface".

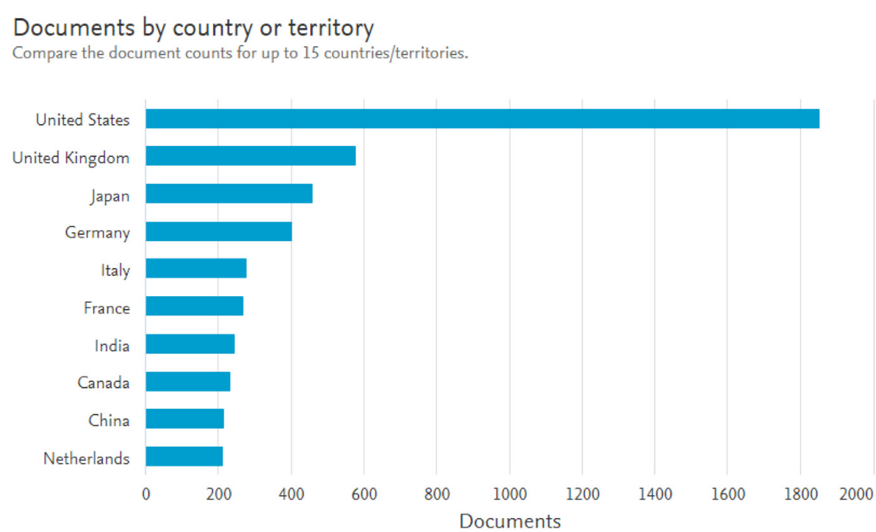

Fig. 2. Search results in Scopus (*Scopus preview—Scopus—Welcome to Scopus*, 2020) from 1966 to 2019 for the keywords "chatbot" or "conversation agent" or "conversational interface" by country or territory.

The Response Generation method separates chatbots into Rule-based, Retrieval based, and Generative based chatbots, which we analyze in the next section (Hien, Cuong, Nam, Nhung, & Thang, 2018).

In some cases where more flexibility is needed, the operation of a chatbot can be combined with human intervention. A Human-mediated chatbot utilizes human computation in at least one part of it. Fully autonomous chatbots may have weaknesses that can be overcome by staff working to integrate their intelligence into them. However, human computation lacks speed in information processing, and it is inevitable to cope with a vast amount of user requests (Kucherbaev et al., 2018).

Depending on the Permissions provided by the development platforms, chatbots can be divided into Open-source or Commercial, which we further discuss in Section 9.

Finally, another classification depends on the Communication channel that chatbots utilize, which can be text, voice, image, or all of them. The latest chatbots can now react to pictures, and aside from recognizing objects in the images, they can also comment on them and express their emotion (Shum, He, & Li, 2018).

4. Chatbot approaches

There are two approaches in developing a chatbot depending on the algorithms and the techniques adopted: pattern matching and machine learning approaches.

4.1. Pattern matching approaches

Rule-based chatbots match the user input to a rule pattern and select a predefined answer from a set of responses with the use of Pattern Matching algorithms. The context can also contribute to the rule selection and the format of the response (Marietto et al., 2013). ELIZA and its successor ALICE were the first chatbots to use pattern matching. At the same time, while PARRY, PC Therapist III, Chatterbot in "TinyMUD", TIPS, FRED, CONVERSE, HEX, Albert, and Jabberwacky (Bradeško & Mladenić, 2012; Masche & Le, 2018) use this technique.

Rule-based systems, typically, do not create new answers as the knowledge used is written by the developer in the form of conversational patterns (Ramesh, Ravishankaran, Joshi, & Chandrasekaran, 2017). The more extensive the database with the rules is, the more

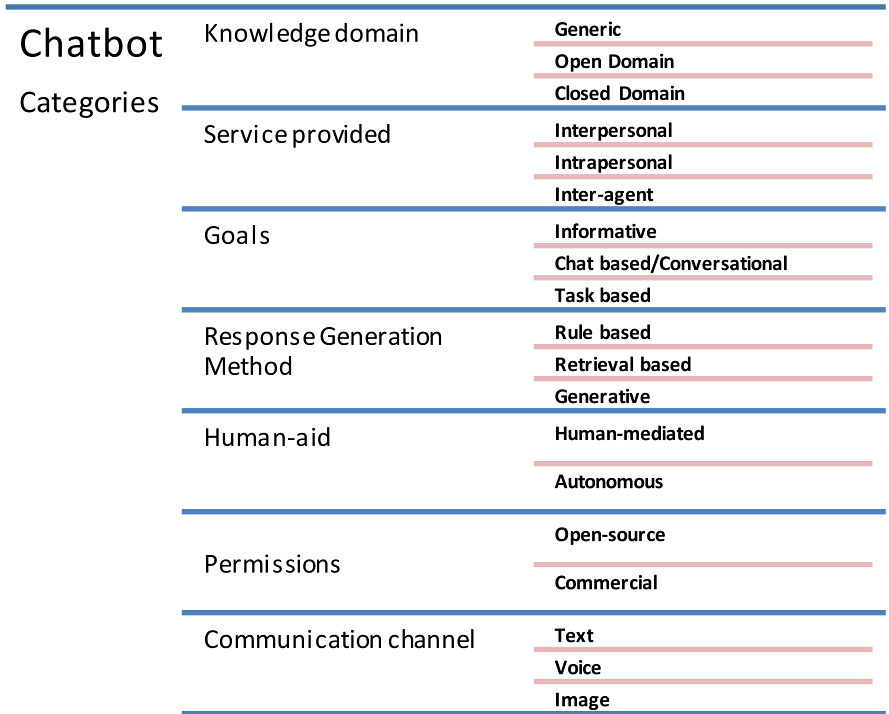

Fig. 3. Chatbots' Classification.

capable a chatbot is of answering the user's questions. As it takes thousands of rules for this type of chatbot to work correctly, it is difficult to deal with grammatical and syntactic errors in the user's responses. Cleverbot, Chatfuel, and Watson are some rule-based chatbots (Ramesh et al., 2017).

In most rule-based chatbots for single-turn communication, the answer is selected, taking into account only the last response. Human-like chatbots, use a multi-turn answer selection in which every response is used as feedback to choose an answer that is normal and appropriate to the entire context (Wu, Wu, Xing, Zhou, & Li, 2016).

The downside of the pattern matching approach is that the answers are automated, repeated, and do not have the originality and spontaneity of human response (Ramesh et al., 2017). On the other hand, there is fast response time, as a deeper syntactic or semantic examination of the input text is not performed (Jia, 2009).

In the following subsections, three of the most common languages for the implementation of chatbots with the pattern-matching approach are described and compared in their basic functionalities. These languages are AIML, Rivescript, and Chatscript.

Artificial Intelligence Markup Language (AIML)

During the years 1995 to 2000, developers created the Artificial Intelligence Markup Language (AIML) (Marietto et al., 2013), to build the Knowledge Base of the chatbots that adopt the Pattern Matching approach. It is based in XML, and it is open-source. ALICE was the first chatbot with a Knowledge Base implemented in the AIML language. Thanks to its usability, ease of learning and execution, and the availability of pre-authored AIML collections, AIML is the most used chatbot language (Arsovski, Muniru, & Cheok, 2017).

AIML data objects consist of topics, which include categories relevant to them. A category is a rule of the chatbot, which has a pattern to represent the user's input, and a template to describe the chatbot's response. The pattern can include words, wildcard symbols, and single spaces. All categories are stored in an object called Graphmaster, which has the form of a tree with its nodes representing the categories and its leaves representing the templates that are the chatbot responses. AIML uses a pattern matching technique that performs a first depth search in the Graphmaster to match the best pattern (Wallace, 2009).

A chatbot implemented with AIML is up to a degree context-aware. It can respond to user input in different ways randomly or based on the value of variables updated during the conversation.

In Fig. 4, we demonstrate a part of a Knowledge Base written in AIML.

Considering the AIML code in Fig. 4, a possible conversation between a user and a chatbot could be the following (Fig. 5):

It is worth mentioning that AIML, and other pattern-matching languages, sometimes are used together with Latent Semantic Analysis (LSA) or other techniques. For example, AIML may answer questions based on specific templates, while questions that cannot be answered in this way can use LSA for the production of responses (Nt, 2016).

RiveScript

RiveScript (*Artificial Intelligence Scripting Language—RiveScript.com*, 2019), created in 2009, is a line-based scripting language implementing the Knowledge Base in rule-based chatbots. It is open source and has interfaces available for many programming languages like Java and Python.

In RiveScript's syntax, the symbol "+" indicates a user input, while "-" denotes the chatbot response. The interpreter matches user input with the stored responses and determines the most suitable reaction to the user input. RiveScript also supports wildcards, conversational redirects, and it is context-aware as it supports user, and chatbot variables (Gupta, Borkar, Mello, & Patil, 2015).

An example of the RiveScript code is shown in Fig. 6.

Chatscript

ChatScript was released in 2011, and it is an expert system for developing rule-based chatbots with an open-source scripting language that is very compact. It matches user inputs to chatbot outputs using pattern matching. An embedded tagger and parser analyzes the user input and improves it in terms of grammar, syntax, and semantics (Wilcox & Wilcox, 2014). ChatScript uses concepts that are collections of similar words concerning the meaning and other parts of speech. Existing databases of concepts can be used directly by developers making the creation of a chatbot easier. It is also case-sensitive and thus, able to detect the emotion in the user's response when the capital or lowercase letters are used for this purpose. Apart from short-term

```
<category>
   <pattern> HELLO </pattern>
    <template>
      <random>
          <li> Hi!  What's your name? </li>
         <li> Hello, How are you? </li>
          <li> Hello! </li>
      </random>
   </template>
 </category>

<category>
   <pattern> MY NAME IS * </pattern>
   <template>Nice to meet you <set name="nameUser"> <star/> </set></template>
</category>

<category>
   <pattern> NIGHT </pattern>
    <template>Good night <get name="nameUser"/> </template>
</category>

<category>
   <pattern> _ NIGHT </pattern>
   <template><srai> NIGHT </srai> </template>
</category>

<category>
   <pattern> NIGHT * </pattern>
   <template> <srai> NIGHT  </srai> </template>
</category>

<category>
   <pattern> _ NIGHT * </pattern>
   <template><srai> NIGHT </srai> </template>
</category>
```

Fig. 4. Example of AIML code.

```
User: Hello
Chatbot: Hi! What's your name?
User: My name is Eleni
Chatbot: Nice to meet you Eleni
User: Goodnight
Chatbot: Goodnight Eleni
```

Fig. 5. Example of conversation.

memory, ChatScript also includes long-term memory using variables that store specific user information and can be used directly or in conjunction with conditionals to produce chatbot responses (Ramesh et al., 2017).

Some chatbots implemented with Chatscript are Suzette, Rosette, Chip Vivant, and Mistsuku (Bradeško & Mladenić, 2012).

The conversation in Fig. 5 is implemented in ChatScript, as shown in Fig. 7.

```
+Hello
- Hi!  What's your name?
- Hello, How are you?
-Hello!

+ my name is  *
-<set name=<formal>>Nice to meet you, <get name>!

+ Goodnight
- Goodnight <get name>
```

Fig. 6. Example of Rivescript code.

Discussion about Pattern Matching Languages

There are advantages and disadvantages when using AIML, Rive-Script, or Chatscript to implement a chatbot (Arsovski et al., 2017).

```
concept:~greeting [hello hi hey]

concept:~goodnight [night goodnight ]
Topic:~chitchat( ~greeting)
t: ~greeting

#! Hello
u: ( * ~greeting * ) [Hi!] [Hello], [What's your name?] [How are you?]

#! My name is
u: (* name * is _* ) Nice to meet you _0!
$username = '_0

#! I wish you a goodnight
u: ( <<!not * ~goodnight * >>) Goodnight $username
```

Fig. 7. Example of Chatscript code.

The main drawback to AIML is that the author must write a pattern for every possible response of the user. However, it helps the chatbot to respond quickly and easily. AIML is easy to learn and implement, but knowledge is presented as an instance in AIML files; therefore, there is no reliable provision for large-scale data management. Additionally, AIML is a rule-based matching word series that produces either a fully contained answer or an input word substitution and is highly inefficient when addressing large Knowledge Bases. (Trivedi, Gor, & Thakkar, 2019). If knowledge is created based on data obtained from the Internet, it must be updated periodically, as automatic updates cannot be conducted.

Moreover, the original version of AIML does not have the option of extension. AIML has poor matching patterns and is challenging to manage. While the content is straightforward to enter, the key challenge is the large amount of data the developer has to enter manually to construct a working chatbot (Arsovski et al., 2017).

RiveSript offers extra built-in features and more tags than AIML. More specifically, there is no need for additional configuration files to define information about the chatbot, for example, its name, which is necessary for AIML. Moreover, it applies the principle of inheritance in its topics, there are weighted random responses, and object macros.

In AIM, each user input is transformed to uppercase to minimize the overhead matching pattern. Nevertheless, this is to the detriment of the emotion expressed when the user uses capital letters. To address this downside and extend the potentials of given responses to the same user input, the ChatScript is case-sensitive. However, ChatScript's scripting language is more complicated than RiveScript and AIML, parsed line-delimited languages. It is also meaning-based, and its support of concepts means that apart from creating a chatbot, it is a system to manipulate human language. Therefore, the developer does not need to write many rules. ChatScript can combine various rules in complicated ways and provide responses that cannot be equally expressed in AIML or RiveScript. Moreover, the RiveScript code can be translated into ChatScript but not the other way around.

4.2. Machine learning approaches

Chatbots that adopt Machine Learning Approaches instead of Pattern Matching extract the content from the user input using Natural Language Processing (NLP), and dispose of the ability to learn from conversations. They consider the whole dialog context, not just the current turn, and do not require a predefined response for each possible user input. Typically, they need an extensive training set, the finding of which may constitute a crucial difficulty as available datasets may be inadequate. For example, movie scripts corpus may be too broad, or an IT helpline may be too specific (Lin, D'Haro, & Banchs, 2016).

Often, Artificial Neural Networks (ANNs) are used for the implementation of these chatbots. Retrieval-based models use a neural network to assign scores and select the most likely response from a set of responses. In contrast, Generative models synthesize the reply, usually using deep learning techniques.

Natural Language Processing (NLP)

Natural Language Processing (NLP) (Khurana, Koli, Khatter, & Singh, 2017) is an Artificial Intelligence field that examines how computer systems can interpret and control the natural language regarding text or speech. Information is collected on the comprehension and usage of human language to create the appropriate techniques for computer systems to manage human language and carry out a variety of tasks (Jung, 2019). Most NLP techniques rely on machine learning. They consist of Natural Language Understanding, which evolves the mission to understand a text, and Natural Language Generation (Langner, Vogel, & Black, 2010; Perera & Nand, 2017), which introduces the responsibility to generate the text commonly conducted by ANNs.

Natural Language Understanding (NLU)

Chatbots use Natural Language Understanding (NLU) (McShane, 2017) to retrieve context from the unstructured user input in human language and respond based on the current user's intention (Jung, 2019). The three major problems raised during the NLU process are the mechanisms of thought, the interpretation, and the general knowledge of the user (Chowdhury, 2003).

NLU supports intent classification and entity extraction, taking into account the context information. Entities may be system-defined or user-defined. Contexts are strings that store the object the user refers to (Ramesh et al., 2017). The intent classification model may be a classifier like, for example, a linear SVM algorithm, or it can be a pre-trained model that was created by the manual classification of collected text messages from users into topics (intents).

Similarly, the entity extraction model can be pre-trained by manually annotating entities in the user's text messages. For this reason, an annotated training corpus may be created by matching a label to each word block. After the models have been trained, they can automatically classify the user's new text messages in intents and extract entities (Hien et al., 2018).

Artificial Neural Networks

Retrieval and Generative-based chatbots use several kinds of Artificial Neural Networks. The system takes the user's input, computes its vector representations, feeds it as features to the neural network, and produces the response. The process of mapping words into vectors is called word embedding, and several methods may be simple or use deep learning techniques like Word2vec (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013).

In Retrieval-based systems, Artificial Neural networks are trained using both user's input and intents vectors. They take as input the user's input vector and give a probability of every intent. The classification of entities in the user input is done by Named-Entity Recognition (NER) systems, which can also use deep learning techniques.

While Generative-based chatbots are useful for involving a person in informal open-domain conversations, they are not ideal for closed-domain communications in which Ruled-based chatbots are suitable. Hybrid chatbots use a Retrieval-based approach and a Generative approach to respond to user input if there is no match to any of the rules (Mathur & Lopez, 2019).

Recurrent Neural Networks

Developers of chatbots use Recurrent Neural Networks (RNN) (Chung, Gulcehre, Cho, & Bengio, 2014) to take account of the previous context in a conversation. In RNN, the new user input and information of earlier data feed the neurons. In this way, a loop passes knowledge from one part of the network to the next.

Moreover, when chatbots developers need to refer to previous information and learn long-term dependencies, they use Long Short Term Memory networks (LSTMs), a particular type of RNN (Xu et al., 2017). How far back the developer wants to go back in the information of a discussion can be a configurable parameter.

A chatbot for answering Frequently Asked Questions (FAQs) was developed with the use of LSTM Recurrent Neural Networks, and experimental results showed that the chatbot could recognize questions and answer them with a very high level of accuracy (Muangkammuen, Intiruk, & Saikaew, 2018). Also, in Kim, Kwon, and Kim (2020), RNN and bidirectional Gated Recurrent Units (GRUs) were used in addition to an attention mechanism to generate the appropriate responses by combining the context of the conversation and the given knowledge.

Sequence-to-Sequence model

A typical Generative-based model is Sequence-to-Sequence (Seq2Seq), which generates a target sequence by looking at the source sequence. The source sequence is the user's input, and the target sequence is the chatbot's response (Ramesh et al., 2017). Two RNNs can be used as the encoder and the decoder, which is the most basic and original version of the model, or LSTMs and GRUs (Chung et al., 2014), can be used to model longer sentences. Such Neural Networks are trained using Backpropagation Through Time, a modified version of backpropagation (Mathur & Lopez, 2019).

In Zhou, Huang, Zhang, Zhu, and Liu (2017) and Zhou et al. (2019), emotionally aware chatbots were implemented using Seq2Seq with GRU. In another research (Cho, van Merrienboer, Gulcehre, Bahdanau, Bougares, Schwenk, & Bengio, 2014), an RNN Encoder–Decoder was proposed to generate a target sequence from a source sequence using a hidden unit with a reset and update unit to control the information it remembers. Also, in Shaikh, More, Puttoo, Shrivastav, and Shinde (2019), a chatbot that enrolls a virtual friend was proposed using Seq2Seq.

Deep Seq2seq Models

Generative chatbots can have a better and more human-like performance when the model is more-in-depth and has more parameters, as in the case of deep Seq2seq models containing multiple layers of LSTM networks (Csaky, 2017).

In Serban et al. (2016), a latent variable hierarchical recurrent encoder–decoder (VHRED) model is introduced, which maintains the long term context of the conversation and generates extended outputs. In another research, an open domain chatbot was developed with the use of Bidirectional Recurrent Neural Networks (BRNN) and attention layers to give the appropriate answers when the user input consists of more than 20 words (Dhyani & Kumar, 2020). A multilingual chatbot based on deep Seq2seq models with GRU cells was proposed to support people in their Cultural Heritage activities (Sperlí, 2020).

5. General architecture

An appropriate chatbot architectural design is useful to the study of chatbots and the aspiring chatbots creator. Therefore, several architectural designs have been proposed (Table 1). In Khanna et al. (2015), the authors propose an architecture, but it is specific for Rule-based chatbots. Similarly, Wu et al. (2016) present an architecture, which is specific for Retrieval-based chatbots. Moreover, they do not include a design. In Zumstein and Hundertmark (2017), the chatbot's Knowledge Base is connected to other databases and information systems that give answers to the user's queries. However, there is an absence of vital services like sentiment analysis and ambiguity handling. An interesting design by Khan (2017) presents an architecture structured in layers. However, it is quite abstract as it does not provide essential details for each layer. Also, it does not discuss how the user's input is analyzed and which are the components of the Dialog Management Component. Besides (Nimavat & Champaneria, 2017), present a simplistic design that lacks many essential details; for example, there is no information about the backend. In Zhou et al. (2017), a chatbot is proposed that generates emotionally consistent responses, but the architecture design focuses only on the Response Generation Module. Hahm works on architecture (Hahm, Kim, An, Lee, & Choi, 2018) too, but it gives no architectural design. The architecture in S. and Balakrishnan (2018) serves the integration of big data with the Knowledge Base of a chatbot. However, it concerns only Rule-based models implemented with AIML. In Mislevics, Grundspeņķis, and Rollande (2018), a detailed architectural design is presented for a chatbot that supports students. This design lacks a sentiment analysis component. In Zhou et al. (2019), the detailed architecture of Microsoft XiaoIce is presented. It includes many interesting characteristics, but it is specific to XiaoIce. In Villegas, Arias-Navarrete, and Palacios (2020), the research focuses on the layers of campus and how they are connected to a chatbot, but it gives no architectural design of the chatbot.

In this section, we compose an architectural design (Fig. 8) that is general and, at the same time, includes all the details we consider crucial. We describe the main components and the individual parts they contain and focus on what we consider most important. The developer can decide which parts he/she is going to implement depending on the type the chatbot.

In Table 1 there is a summary of the architectural features of the chatbots of the before mentioned studies and the proposed architecture.

5.1. User Interface Component

The operation of the chatbot begins when it receives the user's request through an application using text or speech input, such as a messenger application like Facebook, Slack, WhatsApp, WeChat, Viber, or Skype.

5.2. User Message Analysis Component

The User Interface Controller drives the user's request to the User Message Analysis Component to find the user's intention and extracts entities following pattern matching or machine learning approaches. The user's message can be retained as plain text, which keeps all the grammatical and syntactical structures of the input unchanged or processed by Natural Language Processing (NLP) (Chowdhury, 2003; Khurana et al., 2017).

More precisely, through their input to the chatbot, users express their purpose, which is the intent. The chatbot must understand the user's intent and perform the required actions. Different user inputs trigger different intents and may include parameters, called entities, to determine precise details about them (Ramesh et al., 2017).

Certain cognitive services can be linked to the User Message Analysis Component to improve accuracy.

Table 1
Architectural features of existing models and our model.

Study	Abstract Architectural design	Detailed Architectural design	Rule-based	Retrieval-based	Generative-based	User interface	User Message analysis Component	Spell checker	Translator	Sentiment Analysis	Dialog Management Component	Ambiguity Handling	Data Handling	Error Handling	Backend	Response Generation Component
Proposed Architecture		v	v	v	v	v	v	v	v	v	v	v	v	v	v	v
Khanna et al. (2015)	v		v			v						v	v	v	v	
Wu et al. (2016)				v												v
Zumstein and Hundertmark (2017)		v		v			v				v				v	v
Khan (2017)		v	v	v	v	v	v			v	v				v	v
Nimavat and Champaneria (2017)	v		v	v	v		v			v	v	v				v
Zhou et al. (2017)		v			v											v
Hahm et al. (2018)			v				v				v		v	v	v	v
S. and Balakrishnan (2018)		v	v			v	v				v				v	v
Mislevics et al. (2018)		v		v	v	v	v				v				v	v
Zhou et al. (2019)		v		v	v	v	v	v		v	v	v	v	v	v	v
Villegas et al. (2020)		v		v	v	v	v				v				v	v

The proposed architecture consists of a User Interface Component, a User Message Analysis Component, a Dialog Management Component, a Backend, and a Response Generation Component.

- A spell checker corrects the user spelling mistakes as the "purified" input usually results in better intent identification.
- A machine translator is used in the case of multilingual chatbot users. The user's language is identified and translated into the language of the chatbot's NLU.
- Sentiment analysis is applied to the user input to see how satisfied or irritated the user seems to be. In some cases, a real person may need to connect to the discussion if the user looks very frustrated.

Sentiment analysis detects a positive or negative opinion within a text. In Tatai, Csordás, Kiss, Szaló, and Laufer (2003), user reactions were recorded while discussing with chatbots that used different sentiments. Positive chatbot responses, regardless of the user's input sentiment, were preferred. Therefore, chatbots should analyze the sense of the input and learn only from the positive ignoring inputs with a negative or harmful view (Bird, Ekart, & Faria, 2019). A complete expression–emotion mapping database that maps words and expressions into emotions support an emotionally capable chatbot (Tatai et al., 2003). Cloud-based sentiment analysis services are already top-rated and can also be used in chatbots to obtain sentiment analysis from the user input (Keijsers, Bartneck, & Kazmi, 2019).

5.3. Dialog Management Component

The Dialog Management Component controls and updates the conversation context. It keeps the current intent and the identified entities until that point of the conversation. If the chatbot is unable to collect the necessary context information, it asks for additional context information from the user to fulfill missing entities. It also asks follow-up questions after the intent is recognized (Kucherbaev et al., 2018).

The Dialog Management Component typically includes the following modules.

Ambiguity Handling. This module gives answers when the chatbot cannot find the intent from the user's request or if no input is recognized. The chatbot may indicate that it did not have an answer, ask for clarification, start a new discussion (Heller et al., 2005) or give a general answer that covers a variety of issues so that the user is satisfied even if he has asked the most unforeseeable question.

Data Handling. User information is stored in a file. In this way, the chatbot can modify its answers depending on the user giving the impression of being more intelligent.

Error Handling. The Error Handling Module copes with unexpected errors to ensure the proper chatbot operation (Khanna et al., 2015).

After the intent identification, the chatbot proceeds to the next actions, which may be information retrieval from the Backend or responding to the user. In the first case, the control flow handle remains inside the Dialog Management Component, which uses it to determine the next action. In contrast, in the latter case, the control passes to the Backend.

5.4. Backend

The chatbot retrieves the information needed to fulfill the user's intent from the Backend through external APIs calls or Database requests. Once the appropriate information is extracted, it is forwarded to the Dialog Management Module and then to the Response Generation Module.

When Rule-based chatbots are used, there is a Knowledge Base (KB). It includes a list of hand-written responses that correspond to the user's inputs. For a chatbot to stand firm, the Knowledge Base must cover a wide variety of user's queries and contain a variety of responses to the same user input to avoid redundancy of replies (Khanna et al., 2015).

A Relational DataBase (RDB) may be used so that the chatbot can recall past conversations, making in this way the communication more consistent and relevant. This approach brings consistency and precision to the dialog as it allows the chatbot to access previous information history (A & John, 2015).

Creating the Knowledge Base of a chatbot is a necessary but often demanding and time-consuming task because it is manually developed. A method was proposed that automatically builds the Knowledge Base of a new chatbot from an existing one (Arsovski, Osipyan, Oladele, & Cheok, 2019). Also, a program transforms a corpus to the AIML Knowledge Base of a chatbot (Shawar & Atwell, 2004)(AbuShawar & Atwell, 2015). Often, Rule-based chatbots complete their Knowledge Base by asking the user questions and encouraging him for long conversations (Hahm et al., 2018).

The Knowledge Base may also support Ontologies (Semantic Nets) like Wordnet or OpenCyc (Al-Zubaide & Issa, 2011). The chatbot updates the conversation state and searches the nodes of the Knowledge Graph to make connections for the concepts used in the conversation (Chowdhury, 2003). Also, in S. and Balakrishnan (2018), the AIML Knowledge Base was enriched from a big data Knowledge Base with the chatbot's connection to the big data environment.

The use of language tricks in a chatbot's Knowledge Base makes it more human-like. These tricks simulate people's behavior in a conversation, such as stereotyped responses, typing errors, the manner of typing, the existence of a personality, and even irrational responses (aza, muha, zura, & Ahmad, 2018).

5.5. Response Generation Component

The Response Generation Component produces responses using one or more of three available models: Rule-based, Retrieval based, and Generative-based models.

The Rule-based model selects the response from a set of rules without generating new text responses. The Dialog Management component passes the placeholder values that may be needed to fill the template for the response to the Response Generation Module. Rule-based models use a Knowledge Base (KB) organized with conversational patterns (Ramesh et al., 2017).

The Retrieval-based model is more flexible as it selects the most suitable response with the check and analysis of available resources using APIs (Hien et al., 2018).

The Generative model uses Natural Language Generation (NLG) (Singh, Darbari, Bhattacharjee, & Verma, 2016) to respond in a human-like natural language based on the last and previous inputs. However, developing and training such a model is challenging because it needs an extensive set of data for training to establish a fruitful conversation (Hien et al., 2018). When the training corpus is small, grammatical errors are made, particularly in long sentences (Kim, Lee, Kim, Lee, & Kim, 2018).

There have also been hybrid approaches that compare the retrieved to the generated response and choose the better (Song et al., 2018).

When the chatbot produces a response, it presents it to the user and waits until it has feedback (Kompella, 2018).

6. Development

In this section, we discuss critical issues beyond the architectural design, relating to the construction of chatbots.

6.1. The process of creating a chatbot

The creation of a chatbot starts with planning the aims, procedures, and user requirements of it. Developing the chatbot using a programming language or a chatbot development platform follows, and afterward, the chatbot functionalities are tested locally. The chatbot is then published on an online website or a data center, and it is connected to one or more channels to send and receive messages. At this point, proper integration of a chatbot in an application is crucial, and there are three methodologies to do this: integration using API, manual integration, or third-party integration (Trivedi et al., 2019). Finally, evaluation is conducted by gathering data during a fixed time

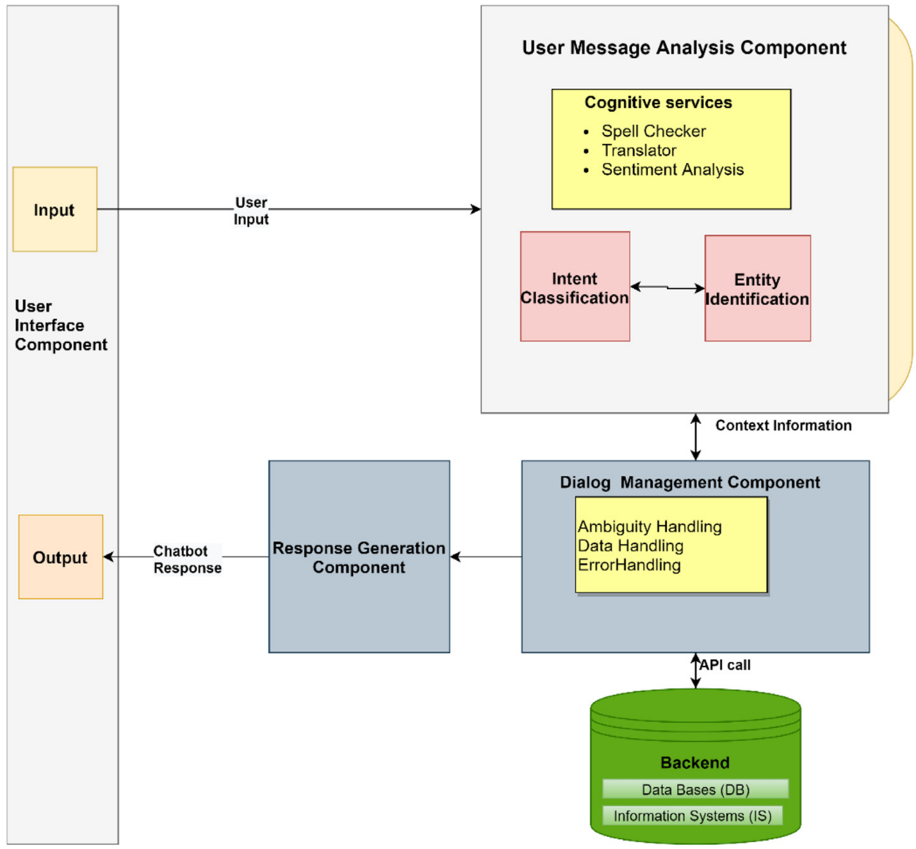

Fig. 8. General Chatbot architecture.

after deployment so that developers can detect errors and use the input to enhance the chatbot's performance and capabilities.

The selection of algorithms or platforms to build chatbots depends on what the chatbot should provide to the users and what category it falls into (Nimavat & Champaneria, 2017). The proper selection may result in the benefits of connectivity, efficiency, quick and uncomplicated production revisions, and minimal effort for the designer. Note that a chatbot is considered more efficient when the user can connect directly without downloading and installing it.

A chatbot can be developed using programming languages like Java and Python or a chatbot development platform that may be commercial or open-source (Nayyar, 2019). Open-source platforms include RASA (*Rasa*, 2019), Botkit *Botkit: Building Blocks for Building Bots* (2019), Chatterbot *About ChatterBot—ChatterBot 1.0.2 documentation* (2019), Pandorabots (*Pandorabots: Home*, 2019), Botlytics (*Botlytics*, 2019) and Microsoft Bot Framework (*Microsoft Bot Framework*, 2019) while commercial platforms include Botsify (Botsify, 2019), Chatfuel (*Chatfuel*, 2020), Manychat (*ManyChat – Chat Marketing Made Easy*, 2019) and Flow XO (*AI Online Chatbot Software, Live Chat on Websites*, 2019). Some NLU cloud platforms (Braun, Hernandez-Mendez, Matthes, & Langen, 2017) powered by machine learning are Google DialogFlow (*Dialogflow*, 2019), Facebook wit.ai (*Wit.ai.*, 2019), Microsoft LUIS (*LUIS (Language Understanding) – Cognitive Services – Microsoft Azure*, 2019), IBM Watson Conversation (*IBM Watson*, 2019), Amazon Lex (*Amazon Lex – Build Conversation Bots*, 2019), and SAP Conversation AI (*SAP Conversational AI | Automate Customer Service With AI Chatbots*, 2019).

Open-source platforms make their code available, and the developer can have full control of the implementation. Although commercial platforms do not give full control to developers, developers are usually benefiting from data efficient to train their chatbots. Moreover, commercial platforms facilitate easy integration to other products of the platform. For example, a chatbot built with Dialogflow may be easily interfaced with Google Assistant and other Google devices. Other facilities that many development platforms may offer include:

1. Easy addition of intents and easy testing of intent triggering (Greyling, 2020)
2. Context management. Most of the development platforms, although they use technically different approaches, support context management (*Contexts | Dialogflow Documentation*, 2020; *Importance of context in a chatbot conversation | IBM*, 2020). Proper context management is critical for intent classification. For example, if the user utterance is "In Kavala, today?" it is most likely that the response will come as the result of an unclassified intent. However, if a context like "weather forecasting" is activated, a successful answer will report the weather forecast in Kavala on the day of the conversation.
3. Support of both predefined and user-defined entities and automatic detection of entity values.
4. Support searching in collections of knowledge documents, like FAQs or articles, to find responses (Charrison, 2020; *Knowledge connectors | Dialogflow Documentation | Google Cloud*, 2020).

Individual platforms may also offer additional facilities. For example, Google Dialogflow supports events, both platform-defined and custom-defined. Instead of user input, events may also trigger intents. For example, an event occurred at a specific time may trigger an intent to alert the user that his/her break time has ended (*Events | Dialogflow Documentation | Google Cloud*, 2020).

Finally, development platforms may be combined with programming languages if the project requirements demand so.

Therefore, development platforms may be an integral part of implementing a chatbot. Of course, before choosing a commercial platform, one should consider the costs involved in using it.

6.2. Training a chatbot

There are many available corpora for training chatbots (Serban, Lowe, Charlin, & Pineau, 2015). Chitchat neural network approaches are usually trained on movie scripts corpora or dialogs from web platforms. A chatbot trained on such datasets, formed by discussions between different speakers, often lacks a specific personality (Zhang et al., 2018). Moreover, various problems may arise when using a dialog corpora with human dialog examples to retrain a chatbot (Atwell & Shawar, 2003). When the dialog corpus is not large enough for training, the chatbot's responses may have syntactic or semantic errors. Therefore, methods have been proposed to pre-train a chatbot using a large non-dialog corpus, and to retrain it using a small dialog corpus (Tammewar, Pamecha, Jain, Nagvenkar, & Modi, 2018).

Sometimes, dialog corpora and machine learning approaches are used to generate the AIML rules and therefore train Rule-based chatbots (Shawar & Atwell, 2010). Moreover, various versions of a chatbot in different languages have been created automatically (Shawar & Atwell, 2005).

Chatbot developers usually keep files from conversations after the chatbot is deployed. It helps them to understand the user requests better and improve the chatbot, which is continuously learning through the analysis of these conversations (et al. 2018). New training examples are extracted from the conversations a chatbot takes part in when the conversation goes on well, or feedback is requested case (Hancock, Bordes, Mazare, & Weston, 2019).

6.3. Connect the chatbot to a channel

Over the last few years, chatbots have gained popularity and are used in various messenger applications or connected to websites.

Facebook is a popular social network for chatbots, mainly for performing transactions or services rather than discussing with users. Facebook Messenger chatbots are usually part of group chats and perform functions like providing statistics around a sports match, creating a music playlist, or giving Smart Replies like informing about business hours or making a reservation without leaving the chat window. Therefore, Facebook chatbots primarily support users in a secretarial way rather than communicate with them.

Skype chatbots, similarly to Facebook ones, are commonly used in group chats for functional purposes. Skype chatbots become more interactive when they give users the alternative of voice chat instead of typing.

Twitter takes a new approach for its chatbots, offering a platform for businesses to connect with their customers and provides an enjoyable rather than a transactional interaction.

Companies use Slack chatbots internally to maximize efficiency, boost connectivity, or perform tasks. There are two categories of Slack chatbots, those that send notifications and those that perform specific transactions initiated by the user.

The benefits of chatbot encapsulation in messaging applications are included: The interaction with the chatbot can quickly be distributed across the user's social network without leaving the messaging app, which ensures the user's identity. Chatbots can be inserted into group chats or exchanged like any contact, whereas a notification system re-engages inactive users. Moreover, payment systems are built into the messaging application and can be used safely and efficiently (Klopfenstein et al., 2017).

In contrast to the channels mentioned above, website based chatbots provide developers with full control. On the website, one can say precisely how the chatbot works, including its purpose, user interface, and experience. Moreover, the users can engage in conversation without leaving the current page, having an easy way to ask questions.

6.4. Conversational mode of chatbots

When a chatbot takes the persona of a famous person, the interaction with its user seems to be improved. Freudbot, a chatbot developed with AIML, was used in distance and online education, and it proved to be a useful teaching and learning tool. It was programmed according to the conversational rules related to turn-taking, providing answers with implications that invited the user to request more information making the conversation longer (Heller et al., 2005).

Chatbots should also offer users a more accessible and realistic experience to ask about their privacy settings in applications or websites (Harkous, Fawaz, & Aberer, 2016).

Moreover, chatbots may draw interest and involve users in activities such as completing questionnaires. Responding to chatbot questions would be a fun and desirable solution to somebody because it does not take as much time as traditional questionnaires (Biduski, Bellei, Rodriguez, Zaina, & Bertoletti De Marchi, 2020).

Users and chatbots interact with each other, and it is interesting to examine "Symbiotic agency", a term that was initially used for the proxy agency where users and software enact in human–technology interaction. Currently, the term expands this concept of the proxy system to consider both how technology mediates somebody's thoughts, beliefs, and attitudes, and how human agency impacts the use of technological products (Neff & Nagy, 2016).

6.5. Other implementation considerations

When designing a chatbot, we must first determine the goals that it will serve. Based on the objectives, we will evaluate the primary character of the chatbot, for example, if we need a generic, cross-domain, or closed-domain chatbot. In the first two cases, we will probably need to use NLP techniques. Before deciding, we should take into account whether the necessary data for chatbot training is available. Conversely, for closed-domain chatbots, the use of a scripting language may be preferable. In general, it is not easy to find suitable training data for specific purposes.

On the other hand, a closed-domain chatbot based on scripting language can if properly designed, effectively guide the user towards achieving specific goals. For example, a vocabulary learning assistant can easily redirect the conversation when it fails to classify the user's intention by saying, 'I did not understand exactly. Do you want me to help you study some vocabulary?'. As long as the discussion remains in the context of vocabulary learning, the dialogs are limited and can be precisely designed so that the result is satisfactory. Also, such a chatbot can be an initial approach that will help in collecting data from users' phrases and will enrich the experience of its manufacturer so that if it is deemed useful to proceed to the second phase using NLP technology.

Whichever approach we follow, we must anticipate that our software will be measuring the user experience so that, based on the measurements, we will be able to make appropriate improvements. However, if we choose NLP technology, we must consider that the choice of an implementation platform that supports cognitive abstraction will provide us, to some extent, automatic updates on the ever-evolving aspects of Artificial Intelligence (Shaw, 2020).

Also, it is an excellent practice to provide the ability to deal with deadlocks with our staff members to avoid unpleasant user experiences (*What Are the Key Considerations When Implementing a Chatbot*, 2020).

A decision that is also important is related to the languages we need to support. Special care is required in voice communication as not all Text-To-Speech (TTS) and Automatic Speech Recognition (ASR) systems support all languages (*5 Core Considerations For Choosing Your Chatbot | by Ruth Zive | Chatbots Magazine*, 2020).

7. Weaknesses and threats of chatbots

In addition to their significant advantages, chatbots are not free of drawbacks and threats.

Customers are quite familiar with communicating with companies using their phones, email, newsletters, or websites while they use messengers applications mainly for their private communication. The new way companies interact with customers is through chatbots in Messengers or stand-alone apps. There is always time for customers to adapt to a new way of communicating, and this is something companies should take into account. In the transformation process, both traditional (offline and online) platforms should be supported, and consumers should be encouraged to use emerging technologies and tools (Zumstein & Hundertmark, 2017).

Data security is a significant concern for both providers and users. Companies are responsible for the proper protection and handling of customer data if they provide a stand-alone chatbot application. However, as companies make their chatbot available on third-party sites, data is often delivered to them. Privacy and data security must be maintained, especially when it comes to authentication and payment systems where confidential, sensitive, or financial details are accessed. Moreover, customers must be aware that when companies communicate with them, gather, store, and use personal data for commercial and marketing purposes (Zumstein & Hundertmark, 2017).

7.1. Failure in intent understanding

Despite their spectacular development, chatbots often fail to recognize the intent of their interlocutor. It is probably their most important weakness that manifests itself quite often. Failure to verify the user's intent creates frustration for him/her. Depending on the scope of the chatbot, this vulnerability may prove detrimental to the chatbot owner, for example, a frustrating conversation with a chatbot serving as a sales assistant can drive the customer away.

7.2. Toxic content in chabot's user inputs

Toxic content may be a severe drawback for chatbot providers and users. For example, personal information recording by unreliable services constitutes toxic content and should be avoided. Besides, utterances aimed at exploiting chatbot or breaching of confidentiality or to copyright theft are considered toxic content. A solution based on homomorphic redaction, for the secure handling of Personal Identifiable Information, proposed by Baudart, Dolby, Duesterwald, Hirzel, and Shinnar (2018). More precisely, an interference frame is activated that diverts the conversational flow to a sub dialog when hazardous material is identified, or other protection needs are present. This approach creates defenses for programming languages, cloud computing, and chatbots. It is non-intrusive, as it does not require updates to current chatbots or underlying conversational platforms. Tay, a chatbot that improves through conversation, sparked huge controversy when attacked by internet trolls on Twitter. After 16 h of its exploitation, it started sending highly abusive Tweets to users. Therefore, although learning from experience is emphasized as a successful strategy, the chatbot must have protection to avoid misuse of it (Neff & Nagy, 2016).

7.3. Deception towards chatbots

Detecting deception is critical in some applications where chatbots are used. Characteristics that make the interactions of chatbots more human-like induce undesired strategic behaviors of human deceivers to hide their deception. To better understand this relationship, the impact of chatbot conversational skills on behavioral manipulation measures was explored (Schuetzler, Grimes, & Giboney, 2019). The findings revealed that the cues of deception differ depending on chatbots' conversational skills. That improved conversational skills led users to engage in strategic activities that are counterproductive to fraud detection. Therefore, using more human-like chatbots can be inefficient for applications in which it is beneficial to identify when individuals are lying.

The convenience that home digital voice assistants offer to their users by monitoring smart devices and using voice commands to receive live assistance, sometimes come with security threats. In Lei et al. (2019), Amazon Alexa was considered as a case study, and several vulnerabilities to protection were identified. The Alexa service relies only on weak single-factor authentication, which can be broken because it follows voice commands without access control dependent on physical presence. This research proposes an additional authentication factor, physical presence.

7.4. Additional failure factors

Long replies where the vital information consists of one or two sentences hiding among several others may result in user discouragement and conversation vacation. Therefore, short and clear messages should be preferred. User errors in spelling may also fail the intent classification (*The Worst Chatbot Fails | Conversational AI vs. Chatbots*, 2019). A spelling checking mechanism employed as a preprocessing step may be useful here.

Other failures initiated from user input may be due to misused phrases, lousy intonation, poor pronunciation, use of subtle humor, speech impairments, use of slang, syntax errors (*The Chat Crash – When a Chatbot Fails*, 2020). Lack of personality on the part of the chatbot can also push a user away from the dialog. One can reduce this risk by giving the chatbot a name and an avatar.

In other cases, user disappointment may be due to a lack of clear chatbot strategy and consequent ineffective guidance of the user towards the communication goals (*#ChatbotFail: 4 Chatbot Customer Experience Fails (And How To Avoid Them)*, 2020). Sometimes a chatbot is designed to serve a specific purpose and then revised to fit an additional need. In this case, the revised chatbot will likely present a dual personality that will create negative feelings in the user (*4 Reasons for Enterprise Chatbot Failure and How to Overcome them using a Multi-Bot Approach*, 2020).

7.5. Mitigate the risks

Researchers around the world hope that further advances in technology will provide satisfactory answers to the weaknesses of chatbots. At this stage, however, few comments may help reduce these weaknesses based on available technology.

Integration with a live chat service, which will be activated for unidentified inputs, may prevent customer disappointment. Also, designing a chatbot that will reclaim personal information gained from other applications (*Three Ways To Mitigate Chatbot Risks*, 2020) may be useful in intent identification. A chatbot with built-in quality assurance tools allows development teams to improve the system continually (How to reduce risk in deploying conversational AI solutions, 2020). Moreover, the specification of approved wording for chatbot utterances may avoid offensive messages (*Three Ways To Mitigate Chatbot Risks*, 2020). Finally, to protect critical transactions like a payment one, integration of the chatbot with a system that performs such operations accurately and efficiently is essential.

8. Applications

8.1. Education environments

The growing demand for learning leads to high competition in higher education institutions. One of the critical reasons for sparse learning and high dropout rates is the fact that, when the number of students grows, the assistance the students get from their teacher is reduced. Chatbots, with their ability to provide educational content and personal assistance, come to support other e-learning practices (Colace, De Santo, Lombardi, Pascale, Pietrosanto, & Lemma, 2018).

Chatbots for learning support can preserve information by repeating old lessons when the students miss them. They also gather information during a course, which helps the improvement of the learning process and teaching. Students are facilitated in the study as chatbots can answer questions concerning the educational material. A chatbot can also help students with school administration issues, such as enrolling in a course, the exam schedule, their grades, and other related details to their studies so that the pressure on the school departments is considerably reduced. In research (Hien et al., 2018), the number of students participating in a University course was growing because a chatbot helped students register.

Students in Foreign Language Learning (FLL) still have little chances to use their target language. Teachers in FLL courses try to build openings during classes through peer or group training, but this is hampered by several reasons such as inadequate resources, insufficient ability for professional input, or lack of self-confidence (Fryer, 2006). Research results have shown that language learners prefer chatbots than human tutors because they feel more confident and can use them anytime (Haristiani, 2019). In Jia (2004), a chatbot was introduced to chat with English learners, which uses a simplistic approach to logical reasoning and inference primarily through syntactic and semantic analysis. It produces a response taking into account the context of the dialog, the user's knowledge and personality, and the knowledge of common sense and inference experience.

8.2. Customer service

The development of new technologies has made people interact with each other differently, and so has their interaction with businesses. E-commerce has evolved and completely changed the way companies sell their products, but there are some problems related to the quality of customer service. Especially in live chats, the waiting time for a business employee to respond may long, and the answers may not always be relevant (Nuruzzaman & Hussain, 2018).

Many companies use chatbots to support customers (Johannsen, Leist, Konadl, & Basche, 2018). Customer care is available 24 h a day via the chatbot, enabling consumers to post their request regardless of the standard operating hours, which enhances user satisfaction. In Gupta et al. (2015), a website based chatbot written in RiveScript helps customers decide which product is suitable for them. Another chatbot implemented with AIML and LSA uses a dataset of Frequently Asked Questions (FAQs) to respond to the users (Nt, 2016). In S. and Balakrishnan (2018), a framework was proposed with a big data interface to provide customer service chatbots with the ability to analyze knowledge from distributed environments.

The continuous use and development of chatbots are expected to make them dominate soon in the field of the customer service industry.

8.3. Health

In health care, chatbots are designed to provide patients with customized health and therapy information, patient-related products and services, and offer diagnosis and suggest treatments based on patient symptoms (*The Top 12 Healthcare Chatbots*, 2020). OneRemission (*OneRemission*, 2020) is a chatbot that helps people being informed about cancer. Youper (*Youper—Emotional Health Assistant Powered by AI*, 2020) takes care of its user's emotional health, and Florence (Florence—Your health assistant, 2020) reminds patients to take their pills. There are many other health care chatbots; some of them are Your.Md (*Your.MD - Health Guide and Self-Care Checker*, 2020), Ada-Health (*Ada*, 2020), Sensely (*Sensely: Character-based Enterprise Virtual Assistant Platform*, 2020) and Buoy Health (*Buoy Health: Check Symptoms & Find the Right Care*, 2020). Moreover, many chatbots were deployed to provide information during the COVID-19 pandemic, such as the HealthBuddy (*HealthBuddy: A new chatbot to engage with communities in Europe and central Asia on COVID-19*, 2020).

The advantages of using health care chatbots include encouraging medical decision-making and supporting, improving physical exercise, support of cognitive-behavioral therapy, and somatic disorders that deliver efficient health treatment with precision equal to that of human doctors (Palanica, Flaschner, Thommandram, Li, & Fossat, 2019).

Patients find that chatbots are more reliable contact partners than human physicists; they share more patient knowledge and disclose more symptoms. However, chatbots in healthcare, are generally associated with weak patient adherence because of the perceived lack of consistency or transparency represented by chatbots, as opposed to regular meetings with human doctors.

On the other hand, physicians believe that chatbots are more effective in administrative activities such as arranging appointments, finding hospitals, and delivering prescription reminders. Still, they are associated with significant risks, including incorrect medical knowledge. Therefore, physicians do not trust chatbots to replace complicated decision-making tasks that require professional medical advice. Especially in the field of psychiatry, chatbots offer the potential of a new and impactful tool. They are used for suicide prevention and cognitive-behavioral intervention, and they are adapted to different populations. A chatbot that delivers therapy can make mental health care system more accessible and more successful with people who are reluctant to talk to a doctor because they feel uncomfortable revealing their feelings. ELIZA served as a Rogerian psychotherapist in 1964, and in 1972 PARRY became a software capable of simulating a human's behavior with schizophrenia that was also "counseled" by ELIZA many times (Vaidyam, Wisniewski, Halamka, Kashavan, & Torous, 2019).

In some cases, chatbots may be better suited to meet patients' wishes than human physicians because they are not biased towards patients, and patients are not biased against chatbots due to gender, age, or race. Also, chatbots do not get exhausted or ill; they are cost-effective and may work uninterruptedly throughout the day, which is especially helpful for people who may have health issues beyond their doctors' working hours. They can also interact in different languages to help respond to specific patient needs.

8.4. Robotics

The most crucial area of research on chatbots is the natural language interface, which is a critical area for physical robots too. Therefore, in the field of physical robots, we find abundant applications of natural language. For example, a novel natural language interface is developed for the autonomous robot called KAMRO (Lueth, Laengle, Herzog, Stopp, & Rembold, 1994). Another natural language interface for instructions to a vision-based robot is designed by (S. (Lauria, Bugmann, Kyriacou, Bos, & Klein, 2001)). In (Selfridge & Vannoy, 1986), a natural language interface allows users to teach vision knowledge and assembly plans to a physical robot. The user can ask questions on vehicle behavior (Garcia et al., 2018). Even, programming a robot using natural language is being explored by Lauria, Bugmann, Kyriacou, and Klein (2002).

Also, in the field of education, the bibliography is quite impressive. For example, in Conti, Di Nuovo, Cirasa, and Di Nuovo (2017), humanoid narrates stories to students. Reduced anxiety in foreign language learning students when a physical robot acts as a tutor is reported by So et al. (2018); Vogt, de Haas, de Jong, Baxter, and Krahmer (2017). In another case, a physical robot help toddlers in vocabulary learning (Roy, Kieson, Abramson, & Crick, 2018).

8.5. Industrial use cases

At the current stage of technological evolution, chatbots are already widely used by many companies and organizations. Here are some examples to give the reader an idea of what is going on in practice.

In the banking sector, chatbots talk to customers and, among other services, provide information about their account balances, facilitate their bill payments, suggest ways to save resources and help activate cards. At the same time, they assist the Bank in collecting feedback from customers. Examples of such chatbots (k, 2020) are Bank Of America's Erika, HDFC's EVA, and Bank of Australia's Ceba.

In the food industry, chatbots accept and track orders, arrange delivery details, make reservations, ask for customer feedback, inform customers on offers and discounts and answer customer questions based on the company's FAQs (*Enterprise Chatbot Platform | AI and NLP Enabled Bots*, 2020). Examples of such chatbots are the Pizza-bot of Dominos (*Domino's introduces 'Dom the Pizza Bot' for Facebook Messenger – Econsultancy*, 2020), the Messenger bot of Whole Food (*Whole Foods Bot*, 2020), the SubWay's chatbot (Subway Bot, 2020), and the Burger King's chatbot (*Burger-Ordering Chatbots: Ai chatbot*, 2020).

IKEA, a leading furniture brand, has launched the so-called ORC chatbot. Zalando, a fashion brand, uses chatbot for order tracking. PVR Cinemas, a large chain of movie theaters in India, uses chatbot for ticket booking. USA's National Railroad Passenger Corporation uses a chatbot called Julie for ticket booking. World Health Organization uses a chatbot called the WHO Health Alert to provide information related to coronavirus.

Many other industry use cases can be found (*Chatbot use cases: What bots can do per industry and function*, 2020; *Top 10 Chatbot Use Cases That Really Work*, 2020; *Top chatbot use cases in different industries*, 2018), but if somebody only likes to chat, he/she can pick up one of Kuki at https://www.pandorabots.com/mitsuku/, Replika at https://my.replika.ai/, Evie at https://www.eviebot.com/en/, Elbot at https://www.elbot.com/, or Cleverbot at https://www.cleverbot.com/.

9. Discussion

As a literature review, this study has been mainly based on bibliographic research. In many cases, we transfer views of authors that have been published in scientific journals or conferences without systematically evaluating them as the size of the literature we have included is prohibitive for systematic evaluation. Besides, we used only English bibliography while works in other languages may contain useful information. Without prejudice to the above limitations, we proceed to a discussion of what has already been presented.

In recent years there has been a significant improvement in the development and use of chatbots with substantial benefits in many domains. At customer service centers, they work 24 h a day, 7 days a week, while managing many customers concurrently, improving payroll costs dramatically. Similarly, in education, they support an increased number of students to whom they provide educational content and personal assistance. In some cases, as in the case where they reduce language anxiety to foreign language students, they even surpass human-teachers. In the field of health care, they provide patients with various services. They also contribute to the development of natural language interfaces for the benefit of Robotics.

However, the usefulness of chatbots is not limited to the areas mentioned. It may not be an exaggeration to say that on most occasions where communication takes place through natural language is also an application opportunity for a chatbot.

As we have already mentioned, the use of chatbots carries risks mainly about personal data security. In this area, however, protection technologies are being developed.

The most crucial problem that chatbots face today is their limitation in understanding and producing natural speech. Sometimes, they cannot understand a phrase, resulting in inconsistencies in communication and unpleasant experiences with their interlocutor. Improving language comprehension and production is perhaps the most critical step in the future development of chatbots. We do not know to what extent the processing of natural language can be developed based on today's technology. Advantages in semantics, context, and knowledge will play a crucial role in NLP development (Hirschberg & Manning, 2015). Nevertheless, we have optimistic expectations, based on the rapid developments in the NLP in recent years.

However, a fundamental issue sets the limits of our optimism. What is the difference between human speech and the speech of a chatbot? Today's technology aims to build chatbots that can learn to talk, but they cannot learn to think. Different approaches in the field of psycholinguistics describe a different relationship between language and thought. We will not expand here, but we will point out that all views accept a close relationship between language and thought (*Language and Thought*, 2020). This finding may indicate a new direction that research should follow to produce chatbots that have a human-like speech. A prerequisite for this new direction is the clarification, by psycholinguistics and other related sciences, of the complex mechanisms of thought, language, and their relation.

CRediT authorship contribution statement

Eleni Adamopoulou: Investigation, Validation, Writing - original draft, Visualization. **Lefteris Moussiades:** Conceptualization, Methodology, Supervision, Writing - review & editing, Project administration.

Acknowledgment

This work is partially supported by the MPhil program "Advanced Technologies in Informatics and Computers", hosted by the Department of Computer Science, International Hellenic University.

References

4 reasons for enterprise chatbot failure and how to overcome them using a multi-bot approach. (2020). https://servisbot.com/four-reasons-for-enterprise-chatbot-failure/. (Retrieved 30 September 2020).

5 core considerations for choosing your chatbot | by Ruth Zive | chatbots magazine. (2020). https://chatbotsmagazine.com/5-core-considerations-for-choosing-your-chatbot-4f4bb0856ead. (Retrieved 30 September 2020).

A, S., & John, D. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications*, *6*, http://dx.doi.org/10.14569/IJACSA.2015.060712.

About ChatterBot—ChatterBot 1.0.2 documentation. (2019). https://chatterbot.readthedocs.io/en/stable/#. (Retrieved 25 November 2019).

AbuShawar, B., & Atwell, E. (2015). ALICE Chatbot: Trials and outputs. *Computación y Sistemas*, *19*, http://dx.doi.org/10.13053/cys-19-4-2326.

Ada: Your health companion. (2020). Ada website: https://ada.com/. (Retrieved 8 June 2020).

AI Online chatbot software, live chat on websites. (2019). Flow XO website: https://flowxo.com/. (Retrieved 25 November 2019).

Al-Zubaide, H., & Issa, A. (2011). Ontbot: Ontology based chatbot. http://dx.doi.org/10.1109/ISIICT.2011.6149594.

Amazon lex – build conversation bots. (2019). Amazon Web Services, Inc. website: https://aws.amazon.com/lex/. (Retrieved 25 November 2019).

Arsovski, S., Muniru, I., & Cheok, A. (2017). Analysis of the chatbot open source languages aiml and chatscript: A review. http://dx.doi.org/10.13140/RG.2.2.34160.15367.

Arsovski, S., Osipyan, H., Oladele, M. I., & Cheok, A. D. (2019). Automatic knowledge extraction of any chatbot from conversation. *Expert Systems with Applications*, *137*, 343–348. http://dx.doi.org/10.1016/j.eswa.2019.07.014.

Artificial intelligence scripting language—Rivescript.com. (2019). https://www.rivescript.com/. (Retrieved 17 November 2019).

Atwell, E., & Shawar, B. (2003). Using dialogue corpora to train a chatbot. https://doi.org/10.13140/2.1.1455.7122.

aza, muha, zura, & Ahmad, N. A. (2018). Review of chatbots design techniques. *International Journal of Computer Applications*, *181*, 7–10.

Bansal, H., & Khan, R. (2018). A review paper on human computer interaction. *International Journal of Advanced Research in Computer Science and Software Engineering*, *8*(53), http://dx.doi.org/10.23956/ijarcsse.v8i4.630.

Baudart, G., Dolby, J., Duesterwald, E., Hirzel, M., & Shinnar, A. (2018). Protecting chatbots from toxic content. (pp. 99–110). https://doi.org/10.1145/3276954.3276958.

Biduski, D., Bellei, E. A., Rodriguez, J., Zaina, L., & Bertoletti De Marchi, A. (2020). Assessing long-term user experience on a mobile health application through an in-app embedded conversation-based questionnaire. *Computers in Human Behavior*, *104*, Article 106169. http://dx.doi.org/10.1016/j.chb.2019.106169.

Bird, J., Ekart, A., & Faria, D. (2019). *Learning from interaction: an intelligent networked-based human-bot and bot-bot chatbot system: contributions presented at the 18th UK workshop on computational intelligence*. https://doi.org/10.1007/978-3-319-97982-3_15.

Botkit: Building blocks for building bots. (2019). https://botkit.ai/. (Retrieved 25 November 2019).

Botlytics. (2019). https://www.botlytics.co/. (Retrieved 25 November 2019).

Botsify (2019). Botsify—Create automated chatbots online for facebook messenger or website. Botsify website: https://botsify.com. (Retrieved 25 November 2019).

Bradeško, L., & Mladenić, D. (2012). A survey of chatbot systems through a loebner prize competition. C, 34.

Brandtzaeg, P., & Følstad, A. (2017). Why people use chatbots. In I. Kompatsiaris, & et al. (Eds.), *Internet science: Vol. 10673*. Cham: Springer, http://dx.doi.org/10.1007/978-3-319-70284-1_30.

Braun, D., Hernandez-Mendez, A., Matthes, F., & Langen, M. (2017). Evaluating natural language understanding services for conversational question answering systems. Undefined website: /paper/Evaluating-Natural-Language-Understanding-Services-Braun-Hernandez-Mendez/ab8c725e04fc25dc03e96332e4490573cd87abd8. (Retrieved 4 June 2019).

Buoy health: Check symptoms & find the right care. (2020). https://www.buoyhealth.com. (Retrieved 8 June 2020).

Burger-ordering chatbots: ai chatbot. (2020). https://www.trendhunter.com/trends/ai-chatbot. (Retrieved 20 August 2020).

Charrison (2020). Creating search skills in watson assistant. https://docs-prod.vmware.com/en/VMware-Workspace-ONE/services/iintelligent-hub_IDM/GUID-85052A5D-405C-441A-BC72-2E47D6ACCF94.html. (Retrieved 20 August 2020).

Chatbot | definition of chatbot in english by Lexico Dictionaries, . (2019). Lexico Dictionaries | English website: https://www.lexico.com/en/definition/chatbot. (Retrieved 16 July 2019).

#Chatbotfail: 4 chatbot customer experience fails (and how to avoid them). (2020). https://www.coredna.com/blogs/chatbot-fail#1. (Retrieved 30 September 2020).

Chatbot use cases: what bots can do per industry and function. (2020). Acquire website: https://acquire.io/blog/chatbot-use-cases/. (Retrieved 20 August 2020).

Chatfuel. (2020). https://chatfuel.com/. (Retrieved 31 March 2020).

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using rnn encoder–decoder for statistical machine translation. *Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation*, 1724–1734, https://doi.org/10.3115/v1/D14-1179.

Chorus—a crowd-powered conversational agent on google hangouts. (2020). http://talkingtothecrowd.org/index.html. (Retrieved 8 June 2020).

Chowdhury, G. G. (2003). Natural language processing. *Annual Review of Information Science and Technology*, *37*(1), 51–89. http://dx.doi.org/10.1002/aris.1440370103.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). *Empirical evaluation of gated recurrent neural networks on sequence modeling.*

Colace, F., De Santo, M., Lombardi, M., Pascale, F., Pietrosanto, A., & Lemma, S. (2018). Chatbot for e-learning: A case of study. *International Journal of Mechanical Engineering and Robotics Research*, *7*, 528–533. http://dx.doi.org/10.18178/ijmerr.7.5.528-533.

Colby, K. M., Hilf, F. D., Weber, S., & Kraemer, H. C. (1972). Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes. *Artificial Intelligence*, *3*, 199–221. http://dx.doi.org/10.1016/0004-3702(72)90049-5.

Colby, K. M., Weber, S., & Hilf, F. D. (1971). Artificial paranoia. *Artificial Intelligence*, *2*(1), 1–25. http://dx.doi.org/10.1016/0004-3702(71)90002-6.

Contexts | dialogflow documentation. (2020). Google Cloud website: https://cloud.google.com/dialogflow/docs/contexts-overview?. (Retrieved 20 August 2020).

Conti, D., Di Nuovo, A., Cirasa, C., & Di Nuovo, S. (2017). A comparison of kindergarten storytelling by human and humanoid robot with different social behavior. In *Proceedings of the companion of the 2017 ACM/IEEE international conference on human-robot interaction* (pp. 97–98). Vienna, Austria: Association for Computing Machinery, http://dx.doi.org/10.1145/3029798.3038359.

Cortana security flaw means your pc may be compromised. (2018). Panda Security Mediacenter website: https://www.pandasecurity.com/mediacenter/mobile-news/cortana-security-flaw/. (Retrieved 27 August 2020).

Costa, P. C. F. da. (2018). Conversing with personal digital assistants: On gender and artificial intelligence. *Journal of Science and Technology of the Arts*, *10*(3), http://dx.doi.org/10.7559/citarj.v10i3.563, 2-59–79.

Csaky, R. (2017). Deep learning based chatbot models. http://dx.doi.org/10.13140/RG.2.2.21857.40801.

Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, *22*, 811–817. http://dx.doi.org/10.1017/S1351324916000243.

Dhyani, M., & Kumar, R. (2020). An intelligent chatbot using deep learning with bidirectional RNN and attention model. *Materials Today: Proceedings*, http://dx.doi.org/10.1016/j.matpr.2020.05.450, S221478532034030X.

Dialogflow. (2019). Dialogflow website: https://dialogflow.com/. (Retrieved 25 November 2019).

Domino's introduces 'Dom the Pizza Bot' for Facebook Messenger – Econsultancy. (2020). https://econsultancy.com/domino-s-introduces-dom-the-pizza-bot-for-facebook-messenger/. (Retrieved 20 August 2020).

Dr. Sbaitso (2019). Wikipedia. https://en.wikipedia.org/w/index.php?title=Dr._Sbaitso&oldid=922071070.

Enterprise chatbot platform | ai and nlp enabled bots. (2020). https://surbo.io/. (Retrieved 20 August 2020).

Events | dialogflow documentation | google cloud. (2020). https://cloud.google.com/dialogflow/docs/events-overview. (Retrieved 20 August 2020).

Fernandes, A. (2018). Nlp, nlu, nlg and how chatbots work. Medium website: https://chatbotslife.com/nlp-nlu-nlg-and-how-chatbots-work-dd7861dfc9df. (Retrieved 8 November 2019).

Florence—your health assistant. (2020). https://florence.chat/. (Retrieved 9 June 2020).

Følstad, A., Bertinussen Nordheim, C., & Alexander Bjørkli, C. (2018). In Bodrunova S. (Ed.), *Internet science: vol. 11193*, *What makes users trust a chatbot for customer service? an exploratory interview study* (pp. 194–208). Cham: Springer, http://dx.doi.org/10.1007/978-3-030-01437-7_16.

Fryer, L. (2006). *Bots as language learning tools. language, learning and technology: Vol. 10.*

Garcia, F. J. C., Robb, D. A., Liu, X., Laskov, A., Patron, P., & Hastie, H. (2018). Explain yourself: A natural language interface for scrutable autonomous robots. ArXiv:1803.02088 [Cs]. Retrieved from http://arxiv.org/abs/1803.02088.

Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, *97*, 304–316. http://dx.doi.org/10.1016/j.chb.2019.01.020.

Good, N., & Wilk, C. (2016). Introducing the guardian chatbot. The guardian. Retrieved from https://www.theguardian.com/help/insideguardian/2016/nov/07/introducing-the-guardian-chatbot.

Google now. (2020). https://www.google.co.uk/landing/now/. (Retrieved 29 February 2020).

Google assistant, your own personal google. (2019). Assistant website: https://assistant.google.com/. (Retrieved 30 August 2019).

Greyling, C. (2020). How to create a chatbot with google dialogflow. Medium website: https://medium.com/@CobusGreyling/how-to-create-a-chatbot-with-google-dialogflow-60616c2b802f. (Retrieved 20 August 2020).

Gupta, S., Borkar, D., Mello, C. D., & Patil, S. (2015). An e-commerce website based chatbot. 6, 3..

Hahm, Y., Kim, J., An, S., Lee, M., & Choi, K.-S. (2018). Chatbot who wants to learn the knowledge: kb-agent. Semdeep/NLIWoD@ISWC, 4..

Hancock, B., Bordes, A., Mazare, P.-E., & Weston, J. (2019). Learning from dialogue after deployment: Feed yourself, chatbot!. ArXiv:1901.05415 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1901.05415.

Haristiani, N. (2019). Artificial intelligence (AI) chatbot as language learning medium: An inquiry. *Journal of Physics: Conference Series*, *1387*, Article 012020. http://dx.doi.org/10.1088/1742-6596/1387/1/012020.

Harkous, H., Fawaz, K., & Aberer, K. (2016). PriBots: conversational privacy with chatbots.

Healthbuddy: a new chatbot to engage with communities in europe and central asia on covid-19. (2020). http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/5/healthbuddy-a-new-chatbot-to-engage-with-communities-in-europe-and-central-asia-on-covid-19. (Retrieved 20 May 2020).

Heiser, J. F., Colby, K. M., Faught, W. S., & Parkison, R. C. (1979). Can psychiatrists distinguish a computer simulation of paranoia from the real thing?: The limitations of turing-like tests as measures of the adequacy of simulations. *Journal of Psychiatric Research*, *15*(3), 149–162. http://dx.doi.org/10.1016/0022-3956(79)90008-6.

Heller, B., Procter, M., Mah, D., Jewell, L., & Cheung, B. (2005). *Freudbot: An investigation of chatbot technology in distance education.*

Hien, H. T., Cuong, P.-N., Nam, L. N. H., Nhung, H. L. T. K., & Thang, L. D. (2018). Intelligent assistants in higher-education environments: The FIT-EBot, a chatbot for administrative and learning support. In *Proceedings of the ninth international symposium on information and communication technology* (pp. 69–76). New York, NY, USA: ACM, http://dx.doi.org/10.1145/3287921.3287937.

Hill, J., Randolph Ford, W., & 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. http://dx.doi.org/10.1016/j.chb.2015.02.026.

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, *349*(6245), 261–266. http://dx.doi.org/10.1126/science.aaa8685.

How to reduce risk in deploying conversational ai solutions. (2020). Boost.ai website: https://www.boost.ai/articles/reducing-risk-in-conversational-ai-deployments. (Retrieved 20 August 2020).

Hoy, M. (2018). Alexa, siri, cortana, and more: An introduction to voice assistants. *Medical Reference Services Quarterly*, *37*, 81–88. http://dx.doi.org/10.1080/02763869.2018.1404391.

Ibm watson. (2019). https://www.ibm.com/watson. (Retrieved 30 August 2019).

Importance of context in a chatbot conversation | IBM. (2020). https://www.ibm.com/cloud/blog/importance-context-chatbot-conversation. (Retrieved 20 August 2020).

Jabberwacky (2019). Wikipedia. Retrieved from https://en.wikipedia.org/w/index.php?title=Jabberwacky&oldid=926273345.

Jia, J. (2004). NLML–A markup language to describe the unlimited english grammar. ArXiv:Cs/0404018. Retrieved from http://arxiv.org/abs/cs/0404018.

Jia, J. (2009). CSIEC: A computer assisted english learning chatbot based on textual knowledge and reasoning. *Knowledge-Based Systems*, *22*(4), 249–255. http://dx.doi.org/10.1016/j.knosys.2008.09.001.

Johannsen, F., Leist, S., Konadl, D., & Basche, M. (2018). *Comparison of commercial chatbot solutions for supporting customer interaction*. ECIS.

Jung, S. (2019). Semantic vector learning for natural language understanding. *Computer Speech & Language*, *56*, 130–145. http://dx.doi.org/10.1016/j.csl.2018.12.008.

Jwala, K. (2019). Developing a Chatbot using Machine Learning. 8(1) 4.

k, venkat (2020). Examples and best use cases of chatbots in banking industry. Medium website: https://medium.com/voice-tech-podcast/examples-and-best-use-cases-of-chatbots-in-banking-industry-1a46cebdb0cd. (Retrieved 20 August 2020).

Kar, R., & Haldar, R. (2016). Applying chatbots to the internet of things: Opportunities and architectural elements. *International Journal of Advanced Computer Science and Applications*, *7*, http://dx.doi.org/10.14569/IJACSA.2016.071119.

Keijsers, M., Bartneck, C., & Kazmi, H. S. (2019). Cloud-based sentiment analysis for interactive agents. In *Proceedings of the 7th international conference on human-agent interaction* (pp. 43–50). Kyoto, Japan: Association for Computing Machinery, http://dx.doi.org/10.1145/3349537.3351883.

Khan, R. (2017). Standardized architecture for conversational agents a.k.a. chatbots. *International Journal of Computer Trends and Technology*, *50*, 114–121. http://dx.doi.org/10.14445/22312803/IJCTT-V50P120.

Khanna, A., Pandey, B., Vashishta, K., Kalia, K., Bhale, P., & Das, T. (2015). A study of today's A.I. through chatbots and rediscovery of machine intelligence. *International Journal of U- and e-Service, Science and Technology*, *8*, 277–284. http://dx.doi.org/10.14257/ijunesst.2015.8.7.28.

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2017). Natural language processing: State of the art, current trends and challenges. ArXiv:1708.05148 [Cs]. Retrieved from http://arxiv.org/abs/1708.05148.

Kim, S., Kwon, O.-W., & Kim, H. (2020). Knowledge-grounded chatbot based on dual wasserstein generative adversarial networks with effective attention mechanisms. *Applied Sciences*, *10*(9), 3335. http://dx.doi.org/10.3390/app10093335.

Kim, J., Lee, H.-G., Kim, H., Lee, Y., & Kim, Y.-G. (2018). *Two-step training and mixed encoding-decoding for implementing a generative chatbot with a small dialogue corpus* (pp. 31–35). http://dx.doi.org/10.18653/v1/W18-6707.

Klopfenstein, L., Delpriori, S., Malatini, S., & Bogliolo, A. (2017). The rise of bots: A survey of conversational interfaces, patterns, and paradigms. In *Proceedings of the 2017 conference on designing interactive systems* (pp. 555–565). Association for Computing Machinery, http://dx.doi.org/10.1145/3064663.3064672.

Knowledge connectors | dialogflow documentation | google cloud. (2020). https://cloud.google.com/dialogflow/docs/knowledge-connectors. (Retrieved 20 August 2020).

Kompella, R. (2018). Conversational AI chat-bot—Architecture overview. Medium website: https://towardsdatascience.com/architecture-overview-of-a-conversational-ai-chat-bot-4ef3dfefd52e. (Retrieved 28 October 2019).

Kucherbaev, P., Bozzon, A., & Houben, G.-J. (2018). Human-aided bots. *IEEE Internet Computing*, *22*(6), 36–43. http://dx.doi.org/10.1109/MIC.2018.252095348.

Langner, B., Vogel, S., & Black, A. (2010). Evaluating a dialog language generation system: Comparing the MOUNTAIN system to other NLG approaches. pp. 1109–1112.

Language and thought. (2020). http://srwyatt.tripod.com/languagethought.htm. (Retrieved 20 August 2020).

Lauria, S., Bugmann, G., Kyriacou, T., Bos, J., & Klein, A. (2001). Training personal robots using natural language instruction. *IEEE Intelligent Systems*, *16*(5), 38–45. http://dx.doi.org/10.1109/MIS.2001.956080.

Lauria, Stanislao, Bugmann, G., Kyriacou, T., & Klein, E. (2002). Mobile robot programming using natural language. *Robotics and Autonomous Systems*, *38*(3), 171–181. http://dx.doi.org/10.1016/S0921-8890(02)00166-5.

Lei, X., Tu, G.-H., Liu, A. X., Ali, K., Li, C.-Y., & Xie, T. (2019). The insecurity of home digital voice assistants—Amazon alexa as a case study. ArXiv:1712.03327 [Cs]. Retrieved from http://arxiv.org/abs/1712.03327.

Lin, L., D'Haro, L., & Banchs, R. (2016). A web-based platform for collection of human-chatbot interactions. (pp. 363–366). http://dx.doi.org/10.1145/2974804.2980500.

Lueth, T. C., Laengle, T., Herzog, G., Stopp, E., & Rembold, U. (1994). KANTRA-human-machine interaction for intelligent robots using natural language. In *Proceedings of 1994 3rd IEEE international workshop on robot and human communication* (pp. 106–111). https://doi.org/10.1109/ROMAN.1994.365947.

Luis (language understanding) – cognitive services – microsoft azure. (2019). https://www.luis.ai/home. (Retrieved 25 November 2019).

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. Humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, *38*(6), 937–947. http://dx.doi.org/10.1287/mksc.2019.1192.

Manychat – chat marketing made easy. (2019). https://manychat.com/. (Retrieved 25 November 2019).

Marietto, M., Varago de Aguiar, R., Barbosa, G., Botelho, W., Pimentel, E., Franca, R., & Silva, V. (2013). Artificial intelligence markup language: A brief tutorial. *International Journal of Computer Science and Engineering Survey*, *04*, http://dx.doi.org/10.5121/ijcses.2013.4301.

Masche, J., & Le, N.-T. (2018). *A review of technologies for conversational systems* (pp. 212–225). http://dx.doi.org/10.1007/978-3-319-61911-8_19.

Mathur, S., & Lopez, D. (2019). A scaled-down neural conversational model for chatbots. *Concurrency Computations: Practice and Experience*, *31*(10), Article e4761. http://dx.doi.org/10.1002/cpe.4761.

Mauldin, M. L. (1994). *CHATTERBOTS, TINYMUDS, and the turing test: entering the loebner prize competition*. AAAI.

McShane, M. (2017). Natural language understanding (NL, unot NLP) in cognitive systems. *AI Magazine*, *38*(4), 43–56. http://dx.doi.org/10.1609/aimag.v38i4.2745.

Microsoft bot framework. (2019). https://dev.botframework.com/. (Retrieved 25 November 2019).

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. s., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, *26*.

Mislevics, A., Grundspeņķis, J., & Rollande, R. (2018). *A systematic approach to implementing chatbots in organizations—RTU leo showcase*. BIR workshops.

Molnár, G., & Zoltán, S. (2018). The role of chatbots in formal education. https://doi.org/10.1109/SISY.2018.8524609.

Mori, M., MacDorman, K., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, *19*, 98–100. http://dx.doi.org/10.1109/MRA.2012.2192811.

Muangkammuen, P., Intiruk, N., & Saikaew, K. R. (2018). Automated thai-FAQ chatbot using RNN-LSTM. In *2018 22nd international computer science and engineering conference* (pp. 1–4). Chiang Mai, Thailand: IEEE, http://dx.doi.org/10.1109/ICSEC.2018.8712781.

Nayyar, D. A. (2019). Chatbots and the open source tools you can use to develop them. Open Source For You website: https://opensourceforu.com/2019/01/chatbots-and-the-open-source-tools-you-can-use-to-develop-them/. (Retrieved 28 August 2019).

Neff, G., & Nagy, P. (2016). Talking to bots: Symbiotic agency and the case of tay. *International Journal of Communication*, *10*, 4915–4931.

Nimavat, K., & Champaneria, T. (2017). Chatbots: An overview types, architecture, tools and future possibilities. *International Journal for Scientific Research and Development*, *5*(7), 1019–1024, Retrieved from http://www.ijsrd.com/.

Nt, T. (2016). An e-business chatbot using AIML and LSA. (pp. 2740–2742). https://doi.org/10.1109/ICACCI.2016.7732476.

Nuruzzaman, M., & Hussain, O. K. (2018). A survey on chatbot implementation in customer service industry through deep neural networks. In *2018 IEEE 15th international conference on e-business engineering* (pp. 54–61). Xi'an: IEEE, http://dx.doi.org/10.1109/ICEBE.2018.00019.

Oneremission, . (2020). https://keenethics.com/project-one-remission. (Retrieved 8 June 2020).

Palanica, A., Flaschner, P., Thommandram, A., Li, M., & Fossat, Y. (2019). Physicians' perceptions of chatbots in health care: Cross-sectional web-based survey. *Journal of Medical Internet Research*, *21*(4), Article e12887. http://dx.doi.org/10.2196/12887.

Pamungkas, E. W. (2019). Emotionally-aware chatbots: A survey. ArXiv:1906.09774 [Cs]. Retrieved from http://arxiv.org/abs/1906.09774.

Pandorabots: home. (2019). https://home.pandorabots.com/home.html. (Retrieved 25 November 2019).

Perera, R., & Nand, P. (2017). Recent advances in natural language generation: A survey and classification of the empirical literature. *Computing and Informatics*, *36*, 1–31. http://dx.doi.org/10.4149/cai_2017_1_1.

Personal digital assistant—cortana home assistant—microsoft. (2019). Microsoft Cortana, your intelligent assistant website: https://www.microsoft.com/en-us/cortana. (Retrieved 30 August 2019).

R. Ranoliya, B., Raghuwanshi, N., & Singh, S. (2017). Chatbot for university related FAQs. In *2017 international conference on advances in computing, communications and informatics* (pp. 1525–1530). Udupi. https://doi.org/10.1109/ICACCI.2017.8126057.

Ramesh, K., Ravishankaran, S., Joshi, A., & Chandrasekaran, K. (2017). A survey of design techniques for conversational agents. In *Information, communication and computing technology: Vol. 750* (pp. 336–350). Singapore: Springer, http://dx.doi.org/10.1007/978-981-10-6544-6_31.

Rasa: open source conversational A.I.. (2019). Rasa website: https://rasa.com/. (Retrieved 18 November 2019).

Roy, S., Kieson, E., Abramson, C., & Crick, C. (2018). Using human reinforcement learning models to improve robots as teachers. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 225–226). Chicago, IL, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3173386.3177074.

S., R., & Balakrishnan, K. (2018). Empowering chatbots with business intelligence by big data integration. *International Journal of Advanced Research in Computer Science*, *9*(1), 627–631. http://dx.doi.org/10.26483/ijarcs.v9i1.5398.

Sannon, S., Stoll, B., DiFranzo, D., Jung, M., & Bazarova, N. N. (2018). How personification and interactivity influence stress-related disclosures to conversational agents. In *Companion of the 2018 ACM conference on computer supported cooperative work and social computing* (pp. 285–288). New York, NY, USA: ACM, http://dx.doi.org/10.1145/3272973.3274076.

Sap conversational ai | automate customer service with ai chatbots. (2019). SAP Conversational AI website: https://cai.tools.sap. (Retrieved 25 November 2019).

Savenkov, D., & Agichtein, E. (2016). CRQA: crowd-powered real-time automatic question answering system. In *Fourth AAAI conference on human computation and crowdsourcing. presented at the fourth aaai conference on human computation and crowdsourcing.* Retrieved from https://aaai.org/ocs/index.php/HCOMP/HCOMP16/paper/view/14049.

Schuetzler, R. M., Grimes, G. M., & Giboney, J. S. (2019). The effect of conversational agent skill on user behavior during deception. *Computers in Human Behavior*, *97*, 250–259. http://dx.doi.org/10.1016/j.chb.2019.03.033.

Scopus preview—scopus—welcome to scopus. (2020). https://www.scopus.com/home.uri. (Retrieved 14 June 2020).

Selfridge, M., & Vannoy, W. (1986). A natural language interface to a robot assembly system. *IEEE Journal of Robotics and Automation*, *2*(3), 167–171. http://dx.doi.org/10.1109/JRA.1986.1087055.

Sensely: character-based enterprise virtual assistant platform. (2020). Sensely website: https://www.sensely.com/. (Retrieved 8 June 2020).

Serban, I., Lowe, R., Charlin, L., & Pineau, J. (2015). A survey of available corpora for building data-driven dialogue systems. *Dialogue and Discourse*, *9*, http://dx.doi.org/10.5087/dad.2018.101.

Serban, I., Sordoni, A., Lowe, R., Charlin, L., Pineau, J., Courville, A., & Bengio, Y. (2016). A hierarchical latent variable encoder-decoder model for generating dialogues.

Shah, H., Warwick, K., Vallverdú, J., & Wu, D. (2016). Can machines talk? Comparison of Eliza with modern dialogue systems. *Computers in Human Behavior*, *58*, 278–295. http://dx.doi.org/10.1016/j.chb.2016.01.004.

Shaikh, A., More, D., Puttoo, R., Shrivastav, S., & Shinde, S. (2019). A survey paper on chatbots. 06(04) 4.

Shaw, G. (2020). 10 key chatbot implementation considerations you should know. Medium website: https://chatbotsjournal.com/10-key-chatbot-implementation-considerations-you-should-know-306d5f27044e. (Retrieved 30 September 2020).

Shawar, B., & Atwell, E. (2004). Accessing an information system by chatting. 3136, pp. 407–412. https://doi.org/10.1007/978-3-540-27779-8_39.

Shawar, B., & Atwell, E. (2005). Using corpora in machine-learning chatbot systems. *International Journal of Corpus Linguistics*, *10*, 489–516. http://dx.doi.org/10.1075/ijcl.10.4.06sha.

Shawar, B., & Atwell, E. (2007). Chatbots: Are they really useful?. *LDV Forum*, *22*, 29–49.

Shawar, B., & Atwell, E. (2010). Chatbots: Can they serve as natural language interfaces to QA corpus? https://doi.org/10.2316/P.2010.689-050.

Shum, H., He, X., & Li, D. (2018). From eliza to xiaoice: Challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, *19*(1), 10–26. http://dx.doi.org/10.1631/FITEE.1700826.

Singh, S., Darbari, H., Bhattacharjee, K., & Verma, S. (2016). Open source NLG systems: A survey with a vision to design a true NLG system. 9, pp. 4409–4421.

Siri Apple website: https://www.apple.com/siri/. (Retrieved 30 August 2019).

Siri (2020). Wikipedia. Retrieved from https://en.wikipedia.org/w/index.php?title=Siri&oldid=942524254.

Skjuve, M., Haugstveit, I., Følstad, A., & Brandtzaeg, P. (2019). Help! is my chatbot falling into the uncanny valley? An empirical study of user experience in human-chatbot interaction. *Human Technology*, *15*, 30–54. http://dx.doi.org/10.17011/ht/urn.201902201607.

So, W.-C., Wong, M. K.-Y., Lam, W.-Y., Cheng, C.-H., Yang, J.-H., Huang, Y., & … Lee, C.-C. (2018). Robot-based intervention may reduce delay in the production of intransitive gestures in chinese-speaking preschoolers with autism spectrum disorder. *Molecular Autism*, *9*(1), 34. http://dx.doi.org/10.1186/s13229-018-0217-5.

Soffar, H. (2019). Apple siri features, use, advantages, disadvantages & using of siri for learning. Science online website: https://www.online-sciences.com/technology/apple-siri-features-use-advantages-disadvantages-using-of-siri-for-learning/. (Retrieved 27 August 2020).

Song, Y., Li, C.-T., Nie, J.-Y., Zhang, M., Zhao, D., & Yan, R. (2018). An ensemble of retrieval-based and generation-based human-computer conversation systems. (pp. 4382–4388). http://dx.doi.org/10.24963/ijcai.2018/609.

Sperlí, G. (2020). A deep learning based chatbot for cultural heritage. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing* (pp. 935–937). Brno Czech Republic: ACM, http://dx.doi.org/10.1145/3341105.3374129.

Subway bot. (2020). ChatbotGuide.org website: https://www.chatbotguide.org/subway-bot. (Retrieved 20 August 2020).

Tammewar, A., Pamecha, M., Jain, C., Nagvenkar, A., & Modi, K. (2018). Production ready chatbots: generate if not retrieve. AAAI Workshops.

Tatai, G., Csordás, A., Kiss, Á., Szaló, A., & Laufer, L. (2003). Happy chatbot, happy user. In T. Rist, R. S. Aylett, D. Ballin, & J. Rickel (Eds.), *Intelligent Virtual Agents* (pp. 5–12). Berlin Heidelberg: Springer.

The chat crash – when a chatbot fails. (2020). Toptal Design Blog website: https://www.toptal.com/designers/ux/chatbot-fails. (Retrieved 17 September 2020).

The top 12 healthcare chatbots. (2020). The Medical Futurist website: https://medicalfuturist.com/top-12-health-chatbots. (Retrieved 20 May 2020).

The worst chatbot fails | conversational ai vs. chatbots. (2019). Netomi website: https://www.netomi.com/the-worst-chatbot-fails-and-how-to-avoid-them. (Retrieved 30 September 2020).

Three ways to mitigate chatbot risks. (2020). https://www.cioapplications.com/cxoinsights/three-ways-to-mitigate-chatbot-risks-nid-2221.html. (Retrieved 20 August 2020).

Top 10 chatbot use cases that really work. (2020). https://www.revechat.com/blog/chatbots-use-cases/. (Retrieved 20 August 2020).

Top chatbot use cases in different industries. (2018). BotCore website: https://botcore.ai/blog/top-chatbot-use-cases-in-different-industries/. (Retrieved 20 August 2020).

Trivedi, A., Gor, V., & Thakkar, Z. (2019). Chatbot generation and integration: A review. *International Journal of Advance Research, Ideas and Innovations in Technology*, *5*(2), 1308–1311, XX.XXX/IJARIIT-V5I2-1840.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *LIX*(236), 433–460. http://dx.doi.org/10.1093/mind/LIX.236.433.

Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and conversational agents in mental health: A review of the psychiatric landscape. *The Canadian Journal of Psychiatry*, *64*(7), 456–464. http://dx.doi.org/10.1177/0706743719828977.

Villegas, W., Arias-Navarrete, A., & Palacios, X. (2020). Proposal of an architecture for the integration of a chatbot with artificial intelligence in a smart campus for the improvement of learning. *Sustainability*, *12*(1500), http://dx.doi.org/10.3390/su12041500.

Vogt, P., de Haas, M., de Jong, C., Baxter, P., & Krahmer, E. (2017). Child-robot interactions for second language tutoring to preschool children. *Frontiers in Human Neuroscience*, *11*, http://dx.doi.org/10.3389/fnhum.2017.00073.

Wallace, R. S. (2009). The anatomy of a.l.I.C.e. In R. Epstein, G. Roberts, & G. Beber (Eds.), *Parsing the turing test: philosophical and methodological issues in the quest for the thinking computer* (pp. 181–210). Dordrecht: Springer Netherlands, http://dx.doi.org/10.1007/978-1-4020-6710-5_13.

Watson assistant | ibm cloud. (2020). https://www.ibm.com/cloud/watson-assistant/. (Retrieved 29 February 2020).

Weizenbaum, J. (1966). ELIZA—A computer program for the study of natural language communication between man and machine. *Commun. ACM*, *9*(1), 36–45. http://dx.doi.org/10.1145/365153.365168.

What are the key considerations when implementing a chatbot. (2020). Engage Hub website: https://engagehub.com/blog/what-are-the-key-considerations-when-implementing-a-chatbot. (Retrieved 30 September 2020).

What exactly is alexa? where does she come from? and how does she work?. (2019). Digital Trends website: https://www.digitaltrends.com/home/what-is-amazons-alexa-and-what-can-it-do/. (Retrieved 30 August 2019).

Whole foods bot.. (2020). ChatbotGuide.org website: https://www.chatbotguide.org/whole-foods-bot. (Retrieved 20 August 2020).

Wilcox, B., & Wilcox, S. (2014). Making it real: Loebner-winning chatbot design. *Arbor*, *189*, Article a086. http://dx.doi.org/10.3989/arbor.2013.764n6009.

Wit.ai.. (2019). https://wit.ai/. (Retrieved 25 November 2019).

Wizu (2018). A visual history of chatbots. Medium website: https://chatbotsmagazine.com/a-visual-history-of-chatbots-8bf3b31dbfb2. (Retrieved 24 February 2020).

Wu, Y., Wu, W., Xing, C., Zhou, M., & Li, Z. (2016). Sequential matching network: A new architecture for multi-turn response selection in retrieval-based chatbots. ArXiv:1612.01627 [Cs]. Retrieved from http://arxiv.org/abs/1612.01627.

Xu, A., Liu, Z., Guo, Y., Sinha, V., & Akkiraju, R. (2017). A new chatbot for customer service on social media. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 3506–3510). New York, NY, USA: ACM, http://dx.doi.org/10.1145/3025453.3025496.

Youper—emotional health assistant powered by ai. (2020). Youper website: https://www.youper.ai/. (Retrieved 8 June 2020).

Your.md - health guide and self-care checker. (2020). Your.MD website: https://www.your.md/. (Retrieved 8 June 2020).

Zemčík, T. (2019). A brief history of chatbots. *DEStech Transactions on Computer Science and Engineering*, http://dx.doi.org/10.12783/dtcse/aicae2019/31439.

Zhang, S., Dinan, E., Urbanek, J., Szlam, A., Kiela, D., & Weston, J. (2018). Personalizing dialogue agents: I have a dog, do you have pets too?. ArXiv:1801.07243 [Cs]. Retrieved from http://arxiv.org/abs/1801.07243.

Zhou, L., Gao, J., Li, D., & Shum, H.-Y. (2019). The design and implementation of xiaoice, an empathetic social chatbot. ArXiv:1812.08989 [Cs]. Retrieved from http://arxiv.org/abs/1812.08989.

Zhou, H., Huang, M., Zhang, T., Zhu, X., & Liu, B. (2017). Emotional chatting machine: emotional conversation generation with internal and external memory.

Zumstein, D., & Hundertmark, S. (2017). Chatbots – an interactive technology for personalized communication. *Transactions and Services. IADIS International Journal on WWW/Internet*, *15*, 96–109.

Further reading

https://promo.bankofamerica.com/Erica/. (Accessed 24 October 2020).

https://www.hdfcbank.com/personal/ways-to-bank/eva. (Accessed 24 October 2020).

https://www.commbank.com.au/support/ceba.html. (Accessed 24 October 2020).

https://www.smartinsights.com/marketplace-analysis/customer-analysis/how-ikea-are-innovating-in-customer-research/. (Accessed 24 October 2020).

https://annual-report.zalando.com/2017/choose-your-chatbot/. (Accessed 24 October 2020).

https://www.news18.com/news/tech/pvr-cinemas-rolls-out-ai-powered-chatbot-on-their-website-1326262.html. (Accessed 24 October 2020).

https://www.verint.com/engagement/our-story/customer-success/amtraks-intelligent-virtual-assistant-ask-julie-success-story-with-verint/. (Accessed 24 October 2020).

https://www.who.int/news-room/feature-stories/detail/who-health-alert-brings-covid-19-facts-to-billions-via-whatsapp. (Accessed 24 October 2020).

http://www.cleverbot.com/. (Accessed 24 October 2020).

https://www.pandorabots.com/mitsuku/. (Accessed 24 October 2020).